Part III

Contingent Claims Pricing

Chapter 7

Basics of Derivative Pricing

Chapter 4 showed how general pricing relationships for contingent claims could be derived in terms of an equilibrium stochastic discount factor or in terms of elementary securities. This chapter takes a more detailed look at this important area of asset pricing.[1] The field of contingent claims pricing experienced explosive growth following the seminal work on option pricing by Fischer Black and Myron Scholes (Black and Scholes 1973) and by Robert Merton (Merton 1973b). Research on contingent claims valuation and hedging continues to expand, with significant contributions coming from both academics and finance practitioners. This research is driving and is being driven by innovations in financial markets. Because research has given new insights into how potential contingent securities might be priced and hedged, financial service providers are more willing to introduce such securities to the market. In addition, existing contingent securities motivate further research by academics and practitioners whose goal is to improve the pricing and hedging of these securities.

[1] The topics in this chapter are covered in greater detail in undergraduate and masters-level financial derivatives texts such as (McDonald 2002) and (Hull 2000). Readers with a background in derivatives at this level may wish to skip this chapter. For others without this knowledge, this chapter is meant to present some fundamentals of derivatives that provide a foundation for more advanced topics covered in later chapters.

We begin by considering two major categories of contingent claims, namely, forward contracts and option contracts. These securities are called *derivatives* because their cashflows derive from another "underlying" variable, such as an asset price, interest rate, or exchange rate.[2] For the case of a derivative whose underlying is an asset price, we will show that the absence of arbitrage opportunities places restrictions on the derivative's value relative to that of its underlying asset.[3] In the case of forward contracts, arbitrage considerations alone may lead to an exact pricing formula. However, in the case of options, these no-arbitrage restrictions cannot determine an exact price for the derivative, but only bounds on the option's price. An exact option pricing formula requires additional assumptions regarding the probability distribution of the underlying asset's returns. The second section of this chapter illustrates how options can be priced using the well-known binomial option pricing technique. This is followed by a section covering different binomial model applications.

The next section begins with a reexamination of forward contracts and how they are priced. We then compare them to option contracts and analyze how the absence of arbitrage opportunities restricts option values.

7.1 Forward and Option Contracts

Chapter 3's discussion of arbitrage derived the link between spot and forward contracts for foreign exchange. Now we show how that result can be generalized to valuing forward contracts on any dividend-paying asset. Following this, we compare option contracts to forward contracts and see how arbitrage places limits on option prices.

[2]Derivatives have been written on a wide assortment of other variables, including commodity prices, weather conditions, catastrophic insurance losses, and credit (default) losses.

[3]Thus, our approach is in the spirit of considering the underlying asset as an elementary security and using no-arbitrage restrictions to derive implications for the derivative's price.

7.1.1 Forward Contracts on Assets Paying Dividends

Similar to the notation introduced previously, let $F_{0\tau}$ be the current date 0 forward price for exchanging one share of an underlying asset τ periods in the future. Recall that this forward price represents the price agreed to at date 0 but to be paid at future date $\tau > 0$ for delivery at date τ of one share of the asset. The long (*short*) party in a forward contract agrees to purchase (*deliver*) the underlying asset in return for paying (*receiving*) the forward price. Hence, the date $\tau > 0$ payoff to the long party in this forward contract is $S_\tau - F_{0\tau}$, where S_τ is the spot price of one share of the underlying asset at the maturity date of the contract.[4] The short party's payoff is simply the negative of the long party's payoff. When the forward contract is initiated at date 0, the parties set the forward price, $F_{0\tau}$, to make the value of the contract equal zero. That is, by setting $F_{0\tau}$ at date 0, the parties agree to the contract without one of them needing to make an initial payment to the other.

Let $R_f > 1$ be one plus the per-period risk-free rate for borrowing or lending over the time interval from date 0 to date τ. Also, let us allow for the possibility that the underlying asset might pay dividends during the life of the forward contract, and use the notation D to denote the date 0 present value of dividends paid by the underlying asset over the period from date 0 to date τ.[5] The asset's dividends over the life of the forward contract are assumed to be known at the initial date 0, so that D can be computed by discounting each dividend payment at the appropriate date 0 risk-free rate corresponding to the time until the dividend payment is made. In the analysis that follows, we also assume that risk-free interest rates are nonrandom, though most of our results in this section and the next continue to hold when interest rates are assumed to change

[4] Obviously S_τ is, in general, random as of date 0 while $F_{0\tau}$ is known as of date 0.

[5] In our context, "dividends" refer to any cashflows paid by the asset. For the case of a coupon-paying bond, the cashflows would be its coupon payments.

randomly through time.[6]

Now we can derive the equilibrium forward price, $F_{0\tau}$, to which the long and short parties must agree in order for there to be no arbitrage opportunities. This is done by showing that the long forward contract's date τ payoffs can be exactly replicated by trading in the underlying asset and the risk-free asset. Then we argue that in the absence of arbitrage, the date 0 values of the forward contract and the replicating trades must be the same.

The following table outlines the cashflows of a long forward contract as well as the trades that would exactly replicate its date τ payoffs.

Date 0 Trade	**Date 0 Cashflow**	**Date τ Cashflow**
Long Forward Contract	0	$S_\tau - F_{0\tau}$
Replicating Trades		
1) Buy Asset and Sell Dividends	$-S_0 + D$	S_τ
2) Borrow	$R_f^{-\tau} F_{0\tau}$	$-F_{0\tau}$
Net Cashflow	$-S_0 + D + R_f^{-\tau} F_{0\tau}$	$S_\tau - F_{0\tau}$

Note that the payoff of the long forward party involves two cashflows: a positive cashflow of S_τ, which is random as of date 0, and a negative cashflow equal to $-F_{0\tau}$, which is certain as of date 0. The former cashflow can be replicated by purchasing one share of the underlying asset but selling ownership of the dividends paid by the asset between dates 0 and τ.[7] This would cost $S_0 - D$, where S_0 is the date 0 spot price of one share of the underlying asset.

[6] This is especially true for cases in which the underlying asset pays no dividends over the life of the contract, that is, $D = 0$. Also, some results can generalize to cases where the underlying asset pays dividends that are random, such as the case when dividend payments are proportional to the asset's value.

[7] In the absence of an explict market for selling the assets' dividends, the individual could borrow the present value of dividends, D, and repay this loan at the future dates when the dividends are received. This will generate a date 0 cashflow of D, and net future cashflows of zero since the dividend payments exactly cover the loan repayments.

The latter cashflow can be replicated by borrowing the discounted value of $F_{0\tau}$. This would generate current revenue of $R_f^{-\tau} F_{0\tau}$. Therefore, the net cost of replicating the long party's cashflow is $S_0 - D -$ $R_f^{-\tau} F_{0\tau}$. In the absence of arbitrage, this cost must be the same as the cost of initiating the long position in the forward contract, which is zero.[8] Hence, we obtain the no-arbitrage condition

$$S_0 - D - R_f^{-\tau} F_{0\tau} = 0 \tag{7.1}$$

or

$$F_{0\tau} = (S_0 - D)\, R_f^{\tau} \tag{7.2}$$

Equation (7.2) determines the equilibrium forward price of the contract. Note that if this contract had been initiated at a previous date, say, date -1, at the forward price $F_{-1\tau} = X$, then the date 0 value (replacement cost) of the long party's payoff, which we denote as f_0, would still be the cost of replicating the two cashflows:

$$f_0 = S_0 - D - R_f^{-\tau} X \tag{7.3}$$

[8] If $S_0 - D - R_f^{-\tau} F_{0\tau} < 0$, the arbitrage would be to perform the following trades at date 0: 1) purchase one share of the stock and sell ownership of the dividends; 2) borrow $R_f^{-\tau} F_{0\tau}$; 3) take a short position in the forward contract. The date 0 net cashflow of these three transactions is $-(S_0 - D) + R_f^{-\tau} F_{0\tau} + 0 > 0$, by assumption. At date τ the individual would: 1) deliver the one share of the stock to satisfy the short forward position; 2) receive $F_{0\tau}$ as payment for delivering this one share of stock; 3) repay borrowing equal to $F_{0\tau}$. The date τ net cashflow of these three transactions is $0 + F_{0\tau} - F_{0\tau} = 0$. Hence, this arbitrage generates a positive cashflow at date 0 and a zero cashflow at date τ. Conversely, if $S_0 - D - R_f^{-\tau} F_{0\tau} > 0$, an arbitrage would be to perform the following trades at date 0: 1) short- sell one share of the stock and purchase rights to the dividends to be paid to the lender of the stock (in the absence of an explict market for buying the assets' dividends, the individual could lend out the present value of dividends, D, and receive payment on this loan at the future dates when the dividends are to be paid); 2) lend $R_f^{-\tau} F_{0\tau}$; 3) take a long position in the forward contract. The date 0 net cashflow of these three transactions is $(S_0 - D) - R_f^{-\tau} F_{0\tau} + 0 > 0$, by assumption. At date τ the individual would: 1) obtain one share of the stock from the long forward position and deliver it to satisfy the short sale obligation; 2) pay $F_{0\tau}$ to short party in forward contract; 3) receive $F_{0\tau}$ from lending agreement. The date τ net cashflow of these three transactions is $0 - F_{0\tau} + F_{0\tau} = 0$. Hence, this arbitrage generates a positive cashflow at date 0 and a zero cashflow at date τ.

However, as long as date 0 is following the initiation of the contract, the value of the payoff would not, in general, equal zero. Of course, the replacement cost of the short party's payoff would be simply $-f_0 = R_f^{-\tau} X + D - S_0$.

It should be pointed out that our derivation of the forward price in equation (7.2) did not require any assumption regarding the random distribution of the underlying asset price, S_τ. The reason for this is due to our ability to replicate the forward contract's payoff using a *static replication* strategy: all trades needed to replicate the forward contract's date τ payoff were done at the initial date 0. As we shall see, such a static replication strategy is not possible, in general, when pricing other contingent claims such as options. Replicating option payoffs will entail, in general, a *dynamic replication* strategy: trades to replicate an option's payoff at date τ will involve trades at multiple dates during the interval between dates 0 and τ. As will be shown, such a dynamic trading strategy requires some assumptions regarding the stochastic properties of the underlying asset's price. Typically, assumptions are made that result in the markets for the contingent claim and the underlying asset being dynamically complete.

As a prerequisite to these issues of option valuation, let us first discuss the basic features of option contracts and compare their payoffs to those of forward contracts.[9]

7.1.2 Basic Characteristics of Option Prices

The owner of a *call* option has the right, but not the obligation, to buy a given asset in the future at a pre-agreed price, known as the *exercise* price, or *strike* price. Similarly, the owner of a *put* option has the right, but not the obligation,

[9] Much of the next section's results are due to Robert C. Merton (Merton 1973b). For greater details see this article.

to sell a given asset in the future at a preagreed price. For each owner (buyer) of an option, there is an option seller, also referred to as the option *writer*. If the owner of a call (*put*) option chooses to exercise, the seller must deliver (*receive*) the underlying asset or commodity in return for receiving (*paying*) the pre-agreed exercise price. Since an option always has a non-negative payoff to the owner, this buyer of the option must make an initial payment, called the option's premium, to the seller of the option.[10]

Options can have different features regarding which future date(s) that exercise can occur. A *European* option can be exercised only at the maturity of the option contract, while an *American* option can be exercised at any time prior to the maturity of the contract.

Let us define the following notation, similar to that used to describe a forward contract. Let S_0 denote the current date 0 price per share of the underlying asset, and let this asset's price at the maturity date of the option contract, τ, be denoted as S_τ. We let X be the exercise price of the option and denote the date t price of European call and put options as c_t and p_t, respectively. Then based on our description of the payoffs of call and put options, we can write the maturity values of European call and put options as

$$c_\tau = \max\left[S_\tau - X,\, 0\right] \tag{7.4}$$

$$p_\tau = \max\left[X - S_\tau,\, 0\right] \tag{7.5}$$

Now we recall that the payoffs to the long and short parties of a forward contract are $S_\tau - F_{0\tau}$ and $F_{0\tau} - S_\tau$, respectively. If we interpret the pre-agreed forward price, $F_{0\tau}$, as analogous to an option's preagreed exercise price, X, then we see that a call option's payoff equals that of the long forward payoff whenever the

[10] The owner of an option will choose to exercise it only if it is profitable to do so. The owner can always let the option expire unexercised, in which case its resulting payoff would be zero.

long forward payoff is positive, and it equals 0 when the long forward payoff is negative. Similarly, the payoff of the put option equals the short forward payoff when this payoff is positive, and it equals 0 when the short forward payoff is negative. Hence, assuming $X = F_{0\tau}$, we see that the payoff of a call option weakly dominates that of a long forward position, while the payoff of a put option weakly dominates that of a short forward position.[11] This is due to the consequence of option payoffs always being nonnegative whereas forward contract payoffs can be of either sign.

Lower Bounds on European Option Values

Since a European call option's payoff is at least as great as that of a comparable long forward position, this implies that the current value of a European call must be at least as great as the current value of a long forward position. Hence, because equation (7.3) is the current value of a long forward position contract, the European call's value must satisfy

$$c_0 \geq S_0 - D - R_f^{-\tau} X \tag{7.6}$$

Furthermore, because the call option's payoff is always nonnegative, its current value must also be nonnegative; that is, $c_0 \geq 0$. Combining this restriction with (7.6) implies

$$c_0 \geq \max\left[S_0 - D - R_f^{-\tau} X,\, 0\right] \tag{7.7}$$

By comparing a European put option's payoff to that of a short forward position, a similar argument can be made to prove that

$$p_0 \geq \max\left[R_f^{-\tau} X + D - S_0,\, 0\right] \tag{7.8}$$

[11] A payoff is said to dominate another when its value is strictly greater in all states of nature. A payoff weakly dominates another when its value is greater in some states of nature and the same in other states of nature.

An alternative proof is as follows. Consider constructing two portfolios at date 0:

Date 0:

- Portfolio A = a put option having value p_0 and a share of the underlying asset having value S_0
- Portfolio B = a bond having initial value of $R_f^{-\tau} X + D$

Then at date τ, these two portfolios are worth:

Date τ:

- Portfolio A = $\max\left[X - S_\tau, 0\right] + S_\tau + DR_f^\tau = \max\left[X, S_\tau\right] + DR_f^\tau$
- Portfolio B = $X + DR_f^\tau$

Since portfolio A's value at date τ is always at least as great as that of portfolio B, the absence of arbitrage implies that its value at date 0 must always be at least as great as that of portfolio B at date 0. Hence, $p_0 + S_0 \geq R_f^{-\tau} X + D$, proving result (7.8).

Put-Call Parity

Similar logic can be used to derive an important relationship that links the value of European call and put options that are written on the same underlying asset and that have the same maturity date and exercise price. This relationship is referred to as *put-call parity*:

$$c_0 + R_f^{-\tau} X + D \,=\, p_0 + S_0 \tag{7.9}$$

To show this, consider forming the following two portfolios at date 0:

Date 0:

- Portfolio A = a put option having value p_0 and a share of the underlying asset having value S_0
- Portfolio B = a call option having value c_0 and a bond with initial value of $R_f^{-\tau}X + D$

Then at date τ, these two portfolios are worth:

Date τ:

- Portfolio A $= \max[X - S_\tau, 0] + S_\tau + DR_f^\tau = \max[X, S_\tau] + DR_f^\tau$
- Portfolio B $= \max[0, S_\tau - X] + X + DR_f^\tau = \max[X, S_\tau] + DR_f^\tau$

Since portfolios A and B have exactly the same payoff, in the absence of arbitrage their initial values must be the same, proving the put-call parity relation (7.9). Note that if we rearrange (7.9) as $c_0 - p_0 = S_0 - R_f^{-\tau}X - D = f_0$, we see that the value of a long forward contract can be replicated by purchasing a European call option and writing (selling) a European put option.

American Options

Relative to European options, American options have the additional right that allows the holder (owner) to exercise the option prior to the maturity date. Hence, all other things being equal, an American option must be at least as valuable as a European option. Thus, if we let the uppercase letters C_0 and P_0 be the current values of American call and put options, respectively, then comparing them to European call and put options having equivalent underlying asset, maturity, and exercise price features, it must be the case that $C_0 \geq c_0$ and $P_0 \geq p_0$.

There are, however, cases where an American option's early exercise feature has no value, because it would not be optimal to exercise the option early. This situation occurs for the case of an American call option written on an asset that pays no dividends over the life of the option. To see this, note that inequality (7.7) says that prior to maturity, the value of a European call option must satisfy $c_0 \geq S_0 - R_f^{-\tau} X$. However, if an American call option is exercised prior to maturity, its value equals $C_0 = S_0 - X < S_0 - R_f^{-\tau} X < c_0$. This contradicts the condition $C_0 \geq c_0$. Hence, if a holder of an American call option wished to liquidate his position, it would always be better to sell the option, receiving C_0, rather than exercising it for the lower amount $S_0 - X$. By exercising early, the call option owner loses the time value of money due to paying X now rather than later. Note, however, that if the underlying asset pays dividends, early exercise of an American call option just prior to a dividend payment may be optimal. In this instance, early exercise would entitle the option holder to receive the asset's dividend payment, a payment that would be lost if exercise were delayed.

For an American put option that is sufficiently *in the money*, that is, S_0 is significantly less than X, it may be optimal to exercise the option early, selling the asset immediately and receiving \$$X$ now, rather than waiting and receiving \$$X$ at date τ (which would have a present value of $R_f^{-\tau} X$). Note that this does not necessarily violate inequality (7.8), since at exercise $P_0 = X - S_0$, which could be greater than $R_f^{-\tau} X + D - S_0$ if the remaining dividends were sufficiently small.

7.2 Binomial Option Pricing

The previous section demonstrated that the absence of arbitrage restricts the price of an option in terms of its underlying asset. However, the no-arbitrage assumption, alone, cannot determine an exact option price as a function of the underlying asset price. To do so, one needs to make an additional assumption regarding the distribution of returns earned by the underlying asset. As we shall see, particular distributional assumptions for the underlying asset can lead to a situation where the option's payoff can be replicated by trading in the underlying asset and a risk-free asset and, in general, this trading occurs at multiple dates. When such a dynamic replication strategy is feasible, the option market is said to be dynamically complete. Assuming the absence of arbitrage then allows us to equate the value of the option's payoff to the prices of more primitive securities, namely, the prices of the underlying asset and the risk-free asset. We now turn to a popular discrete-time, discrete-state model that produces this result.

The model presented in this section was developed by John Cox, Stephen Ross, and Mark Rubinstein (Cox, Ross, and Rubinstein 1979). It makes the assumption that the underlying asset, hereafter referred to as a stock, takes on one of only two possible values each period. While this may seem unrealistic, the assumption leads to a formula that often can accurately price options. This *binomial* option pricing technique is frequently applied by finance practitioners to numerically compute the prices of complex options. Here, we start by considering the pricing of a simple European option written on a non-dividend-paying stock.

In addition to assuming the absence of arbitrage opportunities, the binomial model assumes that the current underlying stock price, S, either moves up, by a proportion u, or down, by a proportion d, each period. The probability of an up move is π, so that the probability of a down move is $1 - \pi$. This two-state stock price process can be illustrated as

$$\begin{array}{ccl} & & uS \quad \text{with probability } \pi \\ & \nearrow & \\ S & & \\ & \searrow & \\ & & dS \quad \text{with probability } 1-\pi \end{array} \tag{7.10}$$

Denote R_f as one plus the risk-free interest rate for the period of unit length. This risk-free return is assumed to be constant over time. To avoid arbitrage between the stock and the risk-free investment, we must have $d < R_f < u$.[12]

7.2.1 Valuing a One-Period Option

Our valuation of an option whose maturity can span multiple periods will use a backward dynamic programming approach. First, we will value the option when it has only one period left until maturity; then we will value it when it has two periods left until maturity; and so on until we establish an option formula for an arbitrary number of periods until maturity.

Let c equal the value of a European call option written on the stock and having a strike price of X. At maturity, $c = \max[0, S_\tau - X]$. Thus, *one period prior to maturity*:

$$
\begin{array}{lll}
 & & c_u \equiv \max[0, uS - X] \;\; \text{with probability } \pi \\
 & \nearrow & \\
c & & \\
 & \searrow & \\
 & & c_d \equiv \max[0, dS - X] \quad \text{with probability } 1 - \pi
\end{array}
\tag{7.11}
$$

What is c one period before maturity? Consider a portfolio containing Δ shares of stock and \$$B$ of bonds. It has current value equal to $\Delta S + B$. Then the value of this portfolio evolves over the period as

[12] If $R_f < d$, implying that the return on the stock is always higher than the risk-free return, an arbitrage would be to borrow at the risk-free rate and use the proceeds to purchase the stock. A profit is assured because the return on the stock would always exceed the loan repayment. Conversely, if $u < R_f$, implying that the return on the stock is always lower than the risk-free return, an arbitrage would be to short-sell the stock and use the proceeds to invest at the risk-free rate. A profit is assured because the risk-free return will always exceed the value of the stock to be repaid to the stock lender.

$$
\Delta S + B \begin{array}{l} \nearrow \quad \Delta uS + R_f B \text{ with probability } \pi \\ \\ \searrow \quad \Delta dS + R_f B \quad \text{with probability } 1-\pi \end{array} \tag{7.12}
$$

With two securities (the bond and stock) and two states of nature (up or down), Δ and B can be chosen to replicate the payoff of the call option:

$$
\Delta uS + R_f B = c_u \tag{7.13}
$$

$$
\Delta dS + R_f B = c_d \tag{7.14}
$$

Solving for Δ and B that satisfy these two equations, we have

$$
\Delta^* = \frac{c_u - c_d}{(u-d)\,S} \tag{7.15}
$$

$$
B^* = \frac{uc_d - dc_u}{(u-d)\,R_f} \tag{7.16}
$$

Hence, a portfolio of Δ^* shares of stock and $\$B^*$ of bonds produces the same cashflow as the call option.[13] This is possible because the option market is complete. As was shown in Chapter 4, in this situation there are equal numbers of states and assets having independent returns so that trading in the stock and bond produces payoffs that span the two states. Now since the portfolio's return replicates that of the option, the absence of arbitrage implies

[13] Δ^*, the number of shares of stock per option contract needed to replicate (or hedge) the option's payoff, is referred to as the option's *hedge ratio*. It can be verified from the formulas that for standard call options, this ratio is always between 0 and 1. For put options, it is always between -1 and 0. B^*, the investment in bonds, is negative for call options but positive for put options. In other words, the replicating trades for a call option involve buying shares in the underlying asset partially financed by borrowing at the risk-free rate. The replicating trades for a put option involve investing at the risk-free rate partially financed by short-selling the underlying asset.

$$c = \Delta^* S + B^* \tag{7.17}$$

This analysis provides practical insights for option traders. Suppose an option writer wishes to hedge her position from selling an option, that is, insure that she will be able to cover her liability to the option buyer in all states of nature. Then her appropriate hedging strategy is to purchase Δ^* shares of stock and $\$B^*$ of bonds since, from equations (7.13) and (7.14), the proceeds from this hedge portfolio will cover her liability in both states of nature. Her cost for this hedge portfolio is $\Delta^* S + B^*$, and in a perfectly competitive options market, the premium received for selling the option, c, will equal this hedging cost.

Example: If $S = \$50$, $u = 2$, $d = .5$, $R_f = 1.25$, and $X = \$50$, then

$$uS = \$100,\ dS = \$25,\ c_u = \$50,\ c_d = \$0$$

Therefore,

$$\Delta^* = \frac{50 - 0}{(2 - .5)\,50} = \frac{2}{3}$$

$$B^* = \frac{0 - 25}{(2 - .5)\,1.25} = -\frac{40}{3}$$

so that

$$c = \Delta^* S + B^* = \frac{2}{3}\,(50) - \frac{40}{3} = \frac{60}{3} = \$20$$

If $c < \Delta^* S + B^*$, then an arbitrage is to short-sell Δ^* shares of stock, invest $\$-B^*$ in bonds, and buy the call option. Conversely, if $c > \Delta^* S + B^*$, then an arbitrage is to write the call option, buy Δ^* shares of stock, and borrow $\$-B^*$.

The resulting option pricing formula has an interesting implication. It can be rewritten as

$$\begin{aligned} c &= \Delta^* S + B^* = \frac{c_u - c_d}{(u-d)} + \frac{uc_d - dc_u}{(u-d)\, R_f} \\ &= \frac{\left[\frac{R_f - d}{u-d} \max\left[0, uS - X\right] + \frac{u - R_f}{u-d} \max\left[0, dS - X\right]\right]}{R_f} \end{aligned} \tag{7.18}$$

which *does not* depend on the probability of an up or down move of the stock, π.

Thus, given S, investors will agree on the no-arbitrage value of the call option even if they do not agree on π. The call option formula does not *directly* depend on investors' attitudes toward risk. It is a *relative* (to the stock) pricing formula. This is reminiscent of Chapter 4's result (4.44) in which contingent claims could be priced based on state prices but without knowledge of the probability of different states occurring. Since π determines the stock's expected rate of return, $u\pi + d(1-\pi) - 1$, this does not need to be known or estimated in order to solve for the no-arbitrage value of the option, c. However, we do need to know u and d, that is, the size of movements per period, which determine the stock's *volatility*.

Note also that we can rewrite c as

$$c = \frac{1}{R_f}\left[\widehat{\pi} c_u + (1 - \widehat{\pi})\, c_d\right] \tag{7.19}$$

where $\widehat{\pi} \equiv \frac{R_f - d}{u-d}$.

Since $0 < \widehat{\pi} < 1$, $\widehat{\pi}$ has the properties of a probability. In fact, this is the *risk-neutral* probability, as defined in Chapter 4, of an up move in the stock's price. To see that $\widehat{\pi}$ equals the true probability π if individuals are risk-neutral, note that if the expected return on the stock equals the risk-free return, R_f, then

$$[u\pi + d(1-\pi)]\,S = R_f S \tag{7.20}$$

which implies that

$$\pi = \frac{R_f - d}{u - d} = \widehat{\pi} \tag{7.21}$$

so that $\widehat{\pi}$ does equal π under risk neutrality. Thus, (7.19) can be expressed as

$$c_t = \frac{1}{R_f}\widehat{E}\,[c_{t+1}] \tag{7.22}$$

where, as in Chapter 4's equation (4.46), $\widehat{E}\,[\cdot]$ denotes the expectation operator evaluated using the risk-neutral probabilities $\widehat{\pi}$ rather than the true, or physical, probabilities π.

7.2.2 Valuing a Multiperiod Option

Next, consider the option's value with *two periods prior to maturity.* The stock price process is

$$\begin{array}{ccccc} & & & & u^2S \\ & & & \nearrow & \\ & & uS & & \\ & \nearrow & & \searrow & \\ S & & & & duS \\ & \searrow & & \nearrow & \\ & & dS & & \\ & & & \searrow & \\ & & & & d^2S \end{array} \tag{7.23}$$

so that the option price process is

$$
\begin{array}{ccccc}
 & & & & c_{uu} \equiv \max\left[0, u^2 S - X\right] \\
 & & & \nearrow & \\
 & & c_u & & \\
 & & & \searrow & \\
 & \nearrow & & & \\
c & & & & c_{du} \equiv \max\left[0, duS - X\right] \\
 & \searrow & & & \\
 & & & \nearrow & \\
 & & c_d & & \\
 & & & \searrow & \\
 & & & & c_{dd} \equiv \max\left[0, d^2 S - X\right]
\end{array}
\tag{7.24}
$$

Using the results from our analysis when there was only one period to maturity, we know that

$$c_u = \frac{\widehat{\pi} c_{uu} + (1 - \widehat{\pi})\, c_{du}}{R_f} \tag{7.25}$$

$$c_d = \frac{\widehat{\pi} c_{du} + (1 - \widehat{\pi})\, c_{dd}}{R_f} \tag{7.26}$$

With two periods to maturity, the one-period-to-go cashflows of c_u and c_d can be replicated once again by the stock and bond portfolio composed of $\Delta^* = \frac{c_u - c_d}{(u-d)S}$ shares of stock and $B^* = \frac{uc_d - dc_u}{(u-d)R_f}$ of bonds. No arbitrage implies

$$c = \Delta^* S + B^* = \frac{1}{R_f}\left[\widehat{\pi} c_u + (1 - \widehat{\pi})\, c_d\right] \tag{7.27}$$

which, as before says that $c_t = \frac{1}{R_f} \widehat{E}\left[c_{t+1}\right]$. The market is not only complete over the last period but over the second-to-last period as well. Substituting in for c_u and c_d, we have

$$\begin{aligned} c &= \frac{1}{R_f^2}\left[\widehat{\pi}^2 c_{uu} + 2\widehat{\pi}\left(1-\widehat{\pi}\right)c_{ud} + \left(1-\widehat{\pi}\right)^2 c_{dd}\right] \\ &= \frac{1}{R_f^2}\left[\widehat{\pi}^2 \max\left[0, u^2 S - X\right] + 2\widehat{\pi}\left(1-\widehat{\pi}\right)\max\left[0, duS - X\right]\right] \\ &\quad + \frac{1}{R_f^2}\left[\left(1-\widehat{\pi}\right)^2 \max\left[0, d^2 S - X\right]\right] \end{aligned} \tag{7.28}$$

which can also be interpreted as $c_t = \frac{1}{R_f^2}\widehat{E}\left[c_{t+2}\right]$. This illustrates that when a market is complete each period, it becomes complete over the sequence of these individual periods. In other words, the option market is said to be *dynamically complete*. Even though the tree diagrams in (7.23) and (7.24) indicate that there are four states of nature two periods in the future (and three different payoffs for the option), these states can be spanned by a dynamic trading strategy involving just two assets. That is, we have shown that by appropriate trading in just two assets, payoffs in greater than two states can be replicated.

Note that c depends only on S, X, u, d, R_f, and the time until maturity, two periods. Repeating this analysis for three, four, five, . . . , n periods prior to maturity, we always obtain

$$c = \Delta^* S + B^* = \frac{1}{R_f}\left[\widehat{\pi} c_u + \left(1-\widehat{\pi}\right)c_d\right] \tag{7.29}$$

By repeated substitution for c_u, c_d, c_{uu}, c_{ud}, c_{dd}, c_{uuu}, and so on, we obtain the formula, with *n periods prior to maturity*:

$$c = \frac{1}{R_f^n}\left[\sum_{j=0}^{n}\left(\frac{n!}{j!\left(n-j\right)!}\right)\widehat{\pi}^j\left(1-\widehat{\pi}\right)^{n-j}\max\left[0, u^j d^{n-j} S - X\right]\right] \tag{7.30}$$

Similar to before, equation (7.30) can be interpreted as $c_t = \frac{1}{R_f^n}\widehat{E}\left[c_{t+n}\right]$, implying that the market is dynamically complete over any number of periods prior to the option's expiration. The formula in (7.30) can be further simplified by defining "a" as the minimum number of upward jumps of S for it to exceed X. Thus a is the smallest nonnegative integer such that $u^a d^{n-a} S > X$. Taking the natural logarithm of both sides, a is the minimum integer $> ln(X/Sd^n)/ln(u/d)$.

Therefore, for all $j < a$ (the option matures out-of-the money),

$$\max\left[0, u^j d^{n-j} S - X\right] = 0 \tag{7.31}$$

while for all $j > a$ (the option matures in-the-money),

$$\max\left[0, u^j d^{n-j} S - X\right] = u^j d^{n-j} S - X \tag{7.32}$$

Thus, the formula for c can be rewritten:

$$c = \frac{1}{R_f^n}\left[\sum_{j=a}^{n}\left(\frac{n!}{j!\,(n-j)!}\right)\widehat{\pi}^j\,(1-\widehat{\pi})^{n-j}\left[u^j d^{n-j} S - X\right]\right] \tag{7.33}$$

Breaking up (7.33) into two terms, we have

$$\begin{aligned} c \quad = \quad & S\left[\sum_{j=a}^{n}\left(\frac{n!}{j!\,(n-j)!}\right)\widehat{\pi}^j\,(1-\widehat{\pi})^{n-j}\left[\frac{u^j d^{n-j}}{R_f^n}\right]\right] \\ & -XR_f^{-n}\left[\sum_{j=a}^{n}\left(\frac{n!}{j!\,(n-j)!}\right)\widehat{\pi}^j\,(1-\widehat{\pi})^{n-j}\right] \end{aligned} \tag{7.34}$$

The terms in brackets in (7.34) are complementary binomial distribution functions, so that we can write (7.34) as

$$c = S\phi[a; n, \widehat{\pi}'] - XR_f^{-n}\phi[a; n, \widehat{\pi}] \tag{7.35}$$

where $\widehat{\pi}' \equiv \left(\frac{u}{R_f}\right)\widehat{\pi}$ and $\phi[a; n, \widehat{\pi}]$ represents the probability that the sum of n random variables that equal 1 with probability $\widehat{\pi}$ and 0 with probability $1 - \widehat{\pi}$ will be $\geq a$. These formulas imply that c is the discounted expected value of the call's terminal payoff under the risk-neutral probability distribution.

If we define τ as the time until maturity of the call option and σ^2 as the variance per unit time of the stock's rate of return (which depends on u and d), then by taking the limit as the number of periods $n \to \infty$, but the length of each period $\frac{\tau}{n} \to 0$, the Cox-Ross-Rubinstein binomial option pricing formula converges to the well-known Black-Scholes-Merton option pricing formula:[14]

$$c = SN(z) - XR_f^{-\tau} N\left(z - \sigma\sqrt{\tau}\right) \tag{7.36}$$

where $z \equiv \frac{\left[\ln \left(\frac{S}{XR_f^{-\tau}}\right) + \frac{1}{2}\sigma^2\tau\right]}{\left(\sigma\sqrt{\tau}\right)}$ and $N(\cdot)$ is that cumulative standard normal distribution function.

7.3 Binomial Model Applications

Cox, Ross, and Rubinstein's binomial technique is useful for valuing relatively complicated options, such as those having American (early exercise) features. In this section we show how the model can be used to value an American put option and an option written on an asset that pays dividends.

Similar to our earlier presentation, assume that over each period of length Δt, stock prices follow the process

[14] The intuition for why (7.36) is a limit of (7.35) is due to the Central Limit Theorem. As the number of periods becomes large, the sum of binomially distributed, random stock rates of return becomes normally distributed. Note that in the Black-Scholes-Merton formula, R_f is now the risk-free return per unit time rather than the risk-free return for each period. The relationship between σ and u and d will be discussed shortly. The Cox-Ross-Rubinstein binomial model (7.35) also can have a different continuous-time limit, namely, the jump-diffusion model that will be presented in Chapter 11.

$$\begin{array}{ccl} & & uS \quad \text{with probability } \pi \\ & \nearrow & \\ S & & \\ & \searrow & \\ & & dS \quad \text{with probability } 1-\pi \end{array} \tag{7.37}$$

The results of our earlier analysis showed that the assumption of an absence of arbitrage allowed us to apply *risk-neutral* valuation techniques to derive the price of an option. Recall that, in general, this method of valuing a derivative security can be implemented by

1) setting the expected rate of return on all securities equal to the risk-free rate

2) discounting the expected value of future cashflows generated from (1) by this risk-free rate

For example, suppose we examine the value of the stock, S, in terms of the risk-neutral valuation method. Similar to the previous analysis, define R_f as the risk-free return per unit time, so that the risk-free return over a time interval Δt is $R_f^{\Delta t}$. Then we have

$$\begin{aligned} S &= R_f^{-\Delta t} \hat{E}\left[S_{t+\Delta t}\right] \\ &= R_f^{-\Delta t}\left[\widehat{\pi} uS + (1-\widehat{\pi}) dS\right] \end{aligned} \tag{7.38}$$

where $\hat{E}\left[\cdot\right]$ represents the expectations operator under the condition that the expected rates of return on all assets equal the risk-free interest rate, which is not necessarily the assets' true expected rates of return. Rearranging (7.38), we obtain

$$R_f^{\Delta t} = \widehat{\pi} u + (1-\widehat{\pi}) d \tag{7.39}$$

which implies

$$\widehat{\pi} = \frac{R_f^{\Delta t} - d}{u - d} \tag{7.40}$$

This is the same formula for $\widehat{\pi}$ as was derived earlier. Hence, risk-neutral valuation is consistent with this simple example.

7.3.1 Calibrating the Model

To use the binomial model to value actual options, the parameters u and d must be calibrated to fit the variance of the underlying stock. When estimating a stock's volatility, it is often assumed that stock prices are lognormally distributed. This implies that the continuously compounded rate of return on the stock over a period of length Δt, given by $\ln(S_{t+\Delta t}) - \ln(S_t)$, is normally distributed with a constant, per-period variance of $\Delta t \sigma^2$. As we shall see in Chapter 9, this constant variance assumption is also used in the Black-Scholes option pricing model. Thus, the sample standard deviation of a time series of historical log stock price changes provides us with an estimate of σ. Based on this value of σ, approximate values of u and d that result in the same variance for a binomial stock price distribution are[15]

[15] That the values of u and d in (7.41) result in a variance of stock returns given by $\sigma^2 \Delta t$ for sufficiently small Δt can be verified by noting that, in the binomial model, the variance of the end-of-period stock price is $E\left[S_{t+\Delta t}^2\right] - E\left[S_{t+\Delta t}\right]^2 = \pi u^2 S^2 + (1-\pi) d^2 S^2 - [\pi S + (1-\pi) dS]^2 = S^2 \left\{\pi u^2 + (1-\pi) d^2 - [\pi u + (1-\pi) d]^2\right\}$
$= S^2 \left[e^{\alpha \Delta t}\left(e^{\sigma\sqrt{\Delta t}} + e^{-\sigma\sqrt{\Delta t}}\right) - 1 - e^{2\alpha\Delta t}\right]$, where $\pi = e^{\alpha \Delta t}$ and α is the (continuously compounded) expected rate of return on the stock per unit time. This implies that the variance of the return on the stock is $\left[e^{\alpha \Delta t}\left(e^{\sigma\sqrt{\Delta t}} + e^{-\sigma\sqrt{\Delta t}}\right) - 1 - e^{2\alpha\Delta t}\right]$. Expanding this expression in a series using $e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \ldots$ and then ignoring all terms of order $(\Delta t)^2$ and higher, it equals $\Delta t \sigma^2$.

$$
\begin{aligned}
u &= e^{\sigma\sqrt{\Delta t}} \\
d &= \frac{1}{u} = e^{-\sigma\sqrt{\Delta t}}
\end{aligned}
\tag{7.41}
$$

Hence, condition (7.41) provides a simple way of calibrating u and d to the stock's volatility, σ.

Now consider the path of the stock price. Because we assumed $u = \frac{1}{d}$, the binomial process for the stock price has the simplified form:

$$
\begin{array}{ccccc}
 & & & & u^4S \\
 & & & u^3S \begin{smallmatrix}\nearrow\\ \searrow\end{smallmatrix} & \\
 & & u^2S \begin{smallmatrix}\nearrow\\ \searrow\end{smallmatrix} & & u^2S \\
 & uS \begin{smallmatrix}\nearrow\\ \searrow\end{smallmatrix} & & uS \begin{smallmatrix}\nearrow\\ \searrow\end{smallmatrix} & \\
S \begin{smallmatrix}\nearrow\\ \searrow\end{smallmatrix} & & S \begin{smallmatrix}\nearrow\\ \searrow\end{smallmatrix} & & S \\
 & dS \begin{smallmatrix}\nearrow\\ \searrow\end{smallmatrix} & & dS \begin{smallmatrix}\nearrow\\ \searrow\end{smallmatrix} & \\
 & & d^2S \begin{smallmatrix}\nearrow\\ \searrow\end{smallmatrix} & & d^2S \\
 & & & d^3S \begin{smallmatrix}\nearrow\\ \searrow\end{smallmatrix} & \\
 & & & & d^4S
\end{array}
\tag{7.42}
$$

Given the stock price, S, and its volatility, σ, the above tree or "lattice" can be calculated for any number of periods using $u = e^{\sigma\sqrt{\Delta t}}$ and $d = e^{-\sigma\sqrt{\Delta t}}$.

7.3.2 Valuing an American Option

We can numerically value an option on this stock by starting at the last period and working back toward the first period. Recall that an American put option that is not exercised early will have a final period (date τ) value

$$P_\tau = \max\left[0, X - S_\tau\right] \tag{7.43}$$

The value of the put at date $\tau - \Delta t$ is then the risk-neutral expected value discounted by $R_f^{-\Delta t}$:

$$\begin{aligned} P_{\tau-\Delta t} &= R_f^{-\Delta t}\hat{E}\left[P_\tau\right] \\ &= R_f^{-\Delta t}\left(\widehat{\pi}P_{\tau,u} + \left(1-\widehat{\pi}\right)P_{\tau,d}\right) \end{aligned} \tag{7.44}$$

where $P_{\tau,u}$ is the date τ value of the option if the stock price changes by proportion u, while $P_{\tau,d}$ is the date τ value of the option if the stock price changes by proportion d. However, with an American put option, we need to check whether this value exceeds the value of the put if it were exercised early. Hence, the put option's value can be expressed as

$$P_{\tau-\Delta t} = \max\left[X - S_{\tau-\Delta t},\ \ R_f^{-\Delta t}\left(\widehat{\pi}P_{\tau,u} + \left(1-\widehat{\pi}\right)P_{\tau,d}\right)\right] \tag{7.45}$$

Let us illustrate this binomial valuation technique with the following example:

A stock has a current price of $S = \$80.50$ and a volatility $\sigma = 0.33$. If $\Delta t = \frac{1}{9}$ year, then $u = e^{\frac{.33}{\sqrt{9}}} = e^{.11} = 1.1163$ and $d = \frac{1}{u} = 0.8958$.

Thus the three-period tree for the stock price is

$$
\begin{array}{ccccccc}
Date:0 & & 1 & & 2 & & 3 \\
 & & & & & & 111.98 \\
 & & & & & \nearrow & \\
 & & & & 100.32 & & \\
 & & & \nearrow & & \searrow & \\
 & & 89.86 & & & & 89.86 \\
 & \nearrow & & \searrow & & \nearrow & \\
S = 80.50 & & & & 80.50 & & \\
 & \searrow & & \nearrow & & \searrow & \\
 & & 72.12 & & & & 72.12 \\
 & & & \searrow & & \nearrow & \\
 & & & & 64.60 & & \\
 & & & & & \searrow & \\
 & & & & & & 57.86
\end{array}
$$

Next, consider valuing an American put option on this stock that matures in $\tau = \frac{1}{3}$ years (4 months) and has an exercise price of $X = \$75$. Assume that the risk-free return is $R_f = e^{0.09}$; that is, the continuously compounded risk-free interest rate is 9 percent. This implies

$$\widehat{\pi} = \frac{R_f^{\Delta t} - d}{u - d} = \frac{e^{\frac{0.09}{9}} - 0.8958}{1.1163 - 0.8958} = 0.5181$$

We can now start at date 3 and begin filling in the tree for the put option:

$$
\begin{array}{ccccccc}
Date: 0 & & 1 & & 2 & & 3 \\
 & & & & & & P_{uuu} \\
 & & & & & \nearrow & \\
 & & & & P_{uu} & & \\
 & & & \nearrow & & \searrow & \\
 & & P_u & & & & P_{uud} \\
 & \nearrow & & \searrow & & \nearrow & \\
P & & & & P_{ud} & & \\
 & \searrow & & \nearrow & & \searrow & \\
 & & P_d & & & & P_{udd} \\
 & & & \searrow & & \nearrow & \\
 & & & & P_{dd} & & \\
 & & & & & \searrow & \\
 & & & & & & P_{ddd}
\end{array}
$$

Using $P_3 = \max[0, X - S_3]$, we have

$$
\begin{array}{cccccccc}
Date:0 & & 1 & & 2 & & 3 \\
 & & & & & & 0.00 \\
 & & & & & \nearrow & \\
 & & & & P_{uu} & & \\
 & & & \nearrow & & \searrow & \\
 & & P_u & & & & 0.00 \\
 & \nearrow & & \searrow & & \nearrow & \\
P & & & & P_{ud} & & \\
 & \searrow & & \nearrow & & \searrow & \\
 & & P_d & & & & 2.88 \\
 & & & \searrow & & \nearrow & \\
 & & & & P_{dd} & & \\
 & & & & & \searrow & \\
 & & & & & & 17.14
\end{array}
$$

Next, using $P_2 = \max\left[X - S_2, \quad R_f^{-\Delta t}\left(\widehat{\pi} P_{3,u} + (1-\widehat{\pi})\, P_{3,d}\right)\right]$, we have

$Date: 0$	1	2	3
			0.00
		0.00	
	P_u		0.00
P		1.37	
	P_d		2.88
		10.40*	
			17.14

*Note that at P_{dd} the option is exercised early since

$$
\begin{aligned}
P_{dd} &= \max\left[X - S_2, \quad R_f^{-\Delta t}\left(\widehat{\pi} P_{3,u} + (1-\widehat{\pi})\, P_{3,d}\right)\right] \\
&= \max\left[75 - 64.60, 9.65\right] = \$10.40
\end{aligned}
$$

Next, using $P_1 = \max\left[X - S_1, \quad R_f^{-\Delta t}\left(\widehat{\pi} P_{2,u} + (1-\widehat{\pi})\, P_{2,d}\right)\right]$, we have

$$
\begin{array}{lllllll}
Date: 0 & & 1 & & 2 & & 3 \\
 & & & & & & 0.00 \\
 & & & & 0.00 & \begin{array}{c}\nearrow \\ \searrow\end{array} & \\
 & & 0.65 & \begin{array}{c}\nearrow \\ \searrow\end{array} & & & 0.00 \\
P & \begin{array}{c}\nearrow \\ \searrow\end{array} & & & 1.37 & \begin{array}{c}\nearrow \\ \searrow\end{array} & \\
 & & 5.66 & \begin{array}{c}\nearrow \\ \searrow\end{array} & & & 2.88 \\
 & & & & 10.40^{*} & \begin{array}{c}\nearrow \\ \searrow\end{array} & \\
 & & & & & & 17.14
\end{array}
$$

Note that the option is *not* exercised early at P_d since

$$
\begin{aligned}
P_d &= \max\left[X - S_1, \quad R_f^{-\Delta t}\left(\widehat{\pi} P_{2,u} + (1-\widehat{\pi}) P_{2,d}\right)\right] \\
&= \max\left[75 - 72.12, 5.66\right] = \$5.66
\end{aligned}
$$

Finally, we calculate the value of the put at date 0 using

$$
\begin{aligned}
P_0 &= \max\left[X - S_0, \quad R_f^{-\Delta t}\left(\widehat{\pi} P_{1,u} + (1-\widehat{\pi}) P_{1,d}\right)\right] \\
&= \max\left[-5.5, 3.03\right] = \$3.03
\end{aligned}
$$

and the final tree for the put is

$$
\begin{array}{ccccccc}
Date:0 & & 1 & & 2 & & 3 \\
 & & & & & & 0.00 \\
 & & & & & \nearrow & \\
 & & & & 0.00 & & \\
 & & & \nearrow & & \searrow & \\
 & & 0.65 & & & & 0.00 \\
 & \nearrow & & \searrow & & \nearrow & \\
3.03 & & & & 1.37 & & \\
 & \searrow & & \nearrow & & \searrow & \\
 & & 5.66 & & & & 2.88 \\
 & & & \searrow & & \nearrow & \\
 & & & & 10.40^{*} & & \\
 & & & & & \searrow & \\
 & & & & & & 17.14
\end{array}
$$

7.3.3 Options on Dividend-Paying Assets

One can generalize the procedure shown in section 7.3.2 to allow for the stock (or portfolio of stocks such as a stock index) to *continuously* pay dividends that have a per unit time yield equal to δ; that is, for Δt sufficiently small, the owner of the stock receives a dividend of $\delta S \Delta t$. For this case of a dividend-yielding asset, we simply redefine

$$\widehat{\pi} = \frac{\left(R_f e^{-\delta}\right)^{\Delta t} - d}{u - d} \tag{7.46}$$

This is because when the asset pays a dividend yield of δ, its expected risk-neutral appreciation is $\left(R_f e^{-\delta}\right)^{\Delta t}$ rather than $R_f^{\Delta t}$.

For the case in which a stock is assumed to pay a known dividend yield, δ, at a *single point in time*, then if date $i\Delta t$ is *prior* to the stock going ex-dividend, the nodes of the stock price tree equal

$$u^j d^{i-j} S \qquad j = 0, 1, \ . \ . \ . \ , i. \tag{7.47}$$

If the date $i\Delta t$ is *after* the stock goes ex-dividend, the nodes of the stock price tree equal

$$u^j d^{i-j} S \,(1-\delta) \qquad j = 0, 1, \ . \ . \ . \ , i. \tag{7.48}$$

The value of an option is calculated as before. We work backwards and again check for the optimality of early exercise.

7.4 Summary

In an environment where there is an absence of arbitrage opportunities, the price of a contingent claim is restricted by the price of its underlying asset. For some derivative securities, such as forward contracts, the contract's payoff can be replicated by the underlying asset and a riskless asset using a static trading strategy. In such a situation, the absence of arbitrage leads to a unique link between the derivative's price and that of its underlying asset without the need for additional assumptions regarding the asset's return distribution. For other types of derivatives, including options, static replication may not be possible. An additional assumption regarding the underlying asset's return distribution is necessary for valuing such derivative contracts. An example is the assumption that the underlying asset's returns are binomially distributed. In this case, an option's payoff can be dynamically replicated by repeated trading in a portfolio consisting of its underlying asset and a risk-free asset. Consistent with our earlier analysis, this situation of a dynamically complete market allows us to value derivatives using the risk-neutral approach. We also illustrated the flexibility of this binomial model by applying it to value options having an early exercise feature as well as options written on a dividend-paying asset.

As will be shown in Chapter 9, the binomial assumption is not the only way to obtain market completeness and a unique option pricing formula. If one assumes that investors can trade continuously in the underlying asset, and the underlying's returns follow a continuous-time diffusion process, then these alternative assumptions can also lead to market completeness. The next chapter prepares us for this important topic by introducing the mathematics of continuous-time stochastic processes.

7.5 Exercises

1. In light of this chapter's discussion of forward contracts on dividend-paying assets, reinterpret Chapter 3's example of a forward contract on a foreign currency. In particular, what are the "dividends" paid by a foreign currency?

2. What is the lower bound for the price of a three-month European put option on a dividend-paying stock when the stock price is \$58, the strike price is \$65, the annualized, risk-free return is $R_f = e^{0.05}$, and the stock is to pay a \$3 dividend two months from now?

3. Suppose that c_1, c_2, and c_3 are the prices of European call options with strike prices X_1, X_2, and X_3, respectively, where $X_3 > X_2 > X_1$ and $X_3 - X_2 = X_2 - X_1$. All options are written on the same asset and have the same maturity. Show that

$$c_2 \leq \frac{1}{2}(c_1 + c_3)$$

Hint: consider a portfolio that is long the option having a strike price of

X_1, long the option having the strike price of X_3, and short two options having the strike price of X_2.

4. Consider the binomial (Cox-Ross-Rubinstein) option pricing model. The underlying stock pays no dividends and has the characteristic that $u = 2$ and $d = 1/2$. In other words, if the stock increases (*decreases*) over a period, its value doubles (*halves*). Also, assume that one plus the risk-free interest rate satisfies $R_f = 5/4$. Let there be two periods and three dates: 0, 1, and 2. At the initial date 0, the stock price is $S_0 = 4$. The following option is a type of *Asian* option referred to as an average price call. The option matures at date 2 and has a terminal value equal to

$$c_2 = \max\left[\frac{S_1 + S_2}{2} - 5, 0\right]$$

where S_1 and S_2 are the prices of the stock at dates 1 and 2, respectively. Solve for the no-arbitrage value of this call option at date 0, c_0.

5. Calculate the price of a three-month American put option on a non-dividend-paying stock when the stock price is \$60, the strike price is \$60, the annualized, risk-free return is $R_f = e^{0.10}$, and the annual standard deviation of the stock's rate of return is $\sigma = .45$, so that $u = 1/d = e^{\sigma\sqrt{\Delta\tau}} = e^{.45\sqrt{\Delta\tau}}$. Use a binomial tree with a time interval of one month.

6. Let the current date be t and let $T > t$ be a future date, where $\tau \equiv T - t$ is the number of periods in the interval. Let $A(t)$ and $B(t)$ be the date t prices of single shares of assets A and B, respectively. Asset A pays no dividends but asset B does pay dividends, and the present (date t) value of asset B's known dividends per share paid over the interval from t to T equals D. The per-period risk-free return is assumed to be constant and equal to R_f.

a. Consider a type of forward contract that has the following features. At date t an agreement is made to exchange at date T one share of asset A for F shares of asset B. No payments between the parties are exchanged at date t. Note that F is negotiated at date t and can be considered a forward price. Give an expression for the equilibrium value of this forward price and explain your reasoning.

b. Consider a type of European call option that gives the holder the right to buy one share of asset A in exchange for paying X shares of asset B at date T. Give the no-arbitrage lower bound for the date t value of this call option, $c(t)$.

c. Derive a put-call parity relation for European options of the type described in part (b).

Chapter 8

Essentials of Diffusion Processes and Itô's Lemma

This chapter covers the basic properties of continuous-time stochastic processes having continuous sample paths, commonly referred to as diffusion processes. It describes the characteristics of these processes that are helpful for modeling many financial and economic time series. Modeling a variable as a continuous-time, rather than a discrete-time, random process can allow for different behavioral assumptions and sharper model results. A variable that follows a continuous-time stochastic process can display constant change yet be observable at each moment in time. In contrast, a discrete-time stochastic process implies that there is no change in the value of the variable over a fixed interval, or that the change cannot be observed between the discrete dates. If an asset price is modeled as a discrete-time process, it is natural to presume that no trading in the asset occurs over the discrete interval. Often this makes problems that involve hedging the asset's risk difficult, since portfolio allocations cannot be rebalanced over the nontrading period. Thus, hedging risky-asset returns may be less than perfect when a discrete-time process is assumed.[1]

Instead, if one assumes that asset prices follow continuous-time processes, prices can be observed and trade can take place continuously. When asset prices follow continuous sample paths, dynamic trading strategies that can fully hedge an asset's risk are possible. Making this continuous hedging assumption often simplifies optimal portfolio choice problems and problems of valuing contingent claims (derivative securities). It permits asset returns to have a continuous distribution (an infinite number of states), yet market completeness is possible because payoffs may be dynamically replicated through continuous trading. Such markets are characterized as dynamically complete.

The mathematics of continuous-time stochastic processes can be traced to Louis Bachelier's 1900 Sorbonne doctoral thesis, *Theory of Speculation*. He developed the mathematics of diffusion processes as a by-product of his modeling of option values. While his work predated Albert Einstein's work on Brownian motion by five years, it fell into obscurity until it was uncovered by Leonard J. Savage and Paul A. Samuelson in the 1950s. Samuelson (Samuelson 1965) used Bachelier's techniques to develop a precursor of the Black-Scholes option pricing model, but it was Robert C. Merton who pioneered the application of continuous-time mathematics to solve a wide variety of problems in financial economics.[2] The popularity of modeling financial time series by continuous-time processes continues to this day.

This chapter's analysis of continuous-time processes is done at an intuitive

[1] Imperfect hedging may, indeed, be a realistic phenomenon. However, in many situations it may not be caused by the inability to trade during a period of time but due to discrete movements (jumps) in asset prices. We examine how to model an asset price process that is a mixture of a diffusion process and a jump process in Chapter 11. Imperfect hedging can also arise because transactions costs lead an individual to choose not to trade to hedge small price movements. For models of portfolio choice in the presence of transactions costs, see work by George Constantinides (Constantinides 1986) and Bernard Dumas and Elisa Luciano (Dumas and Luciano 1991).

[2] See a collection of Merton's work in (Merton 1992).

level rather than a mathematically rigorous one.[3] The first section examines Brownian motion, which is the fundamental building block of diffusion processes. We show how Brownian motion is a continuous-time limit of a discrete-time random walk. How diffusion processes can be developed by generalizing a pure Brownian motion process is the topic of the second section. The last section introduces Itô's lemma, which tells us how to derive the stochastic process for a function of a variable that follows a diffusion process. Itô's lemma is applied extensively in continuous-time financial modeling. It will be used frequently during the remainder of this book.

8.1 Pure Brownian Motion

Here we show how a Brownian motion process can be defined as the limit of a discrete-time process.[4] Consider the following stochastic process observed at date t, $z(t)$. Let Δt be a discrete change in time, that is, some time interval. The change in $z(t)$ over the time interval Δt is given by

$$z(t+\Delta t) - z(t) \equiv \Delta z = \sqrt{\Delta t}\,\tilde{\epsilon} \tag{8.1}$$

where $\tilde{\epsilon}$ is a random variable with $E[\tilde{\epsilon}] = 0$, $Var[\tilde{\epsilon}] = 1$, and $Cov[z(t+\Delta t) - z(t),\, z(s+\Delta t) - z(s)] = 0$ if $(t,\, t+\Delta t)$ and $(s,\, s+\Delta t)$ are nonoverlapping intervals. $z(t)$ is an example of a "random walk" process. Its standard deviation equals the square root of the time between observations.

Given the moments of $\tilde{\epsilon}$, we have $E[\Delta z] = 0$, $Var[\Delta z] = \Delta t$, and $z(t)$ has

[3] The following books (in order of increasing rigor and difficulty) provide more in-depth coverage of the chapter's topics: (Neftci 1996), (Karlin and Taylor 1975), (Karlin and Taylor 1981), and (Karatzas and Shreve 1991).

[4] Brownian motion is named after botanist Robert Brown, who in 1827 observed that pollen suspended in a liquid moved in a continuous, random fashion. In 1905, Albert Einstein explained this motion as the result of random collisions of water molecules with the pollen particle.

serially uncorrelated (independent) increments. Now consider the change in $z(t)$ over a fixed interval, from 0 to T. Assume T is made up of n intervals of length Δt. Then

$$z(T) - z(0) = \sum_{i=1}^{n} \Delta z_i \tag{8.2}$$

where $\Delta z_i \equiv z(i \cdot \Delta t) - z([i-1] \cdot \Delta t) \equiv \sqrt{\Delta t}\, \tilde{\epsilon}_i$, and $\tilde{\epsilon}_i$ is the value of $\tilde{\epsilon}$ over the i^{th} interval. Hence (8.2) can also be written as

$$z(T) - z(0) = \sum_{i=1}^{n} \sqrt{\Delta t}\, \tilde{\epsilon}_i = \sqrt{\Delta t} \sum_{i=1}^{n} \tilde{\epsilon}_i \tag{8.3}$$

Now note that the first two moments of $z(T) - z(0)$ are

$$E_0[\, z(T) - z(0)\,] = \sqrt{\Delta t} \sum_{i=1}^{n} E_0[\, \tilde{\epsilon}_i \,] = 0 \tag{8.4}$$

$$Var_0[\, z(T) - z(0)\,] = \left(\sqrt{\Delta t}\right)^2 \sum_{i=1}^{n} Var_0[\tilde{\epsilon}_i] = \Delta t \cdot n \cdot 1 = T \tag{8.5}$$

where $E_t\,[\cdot]$ and $Var_t\,[\cdot]$ are the mean and variance operators, respectively, conditional on information at date t. We see that holding T (the length of the time interval) fixed, the mean and variance of $z(T) - z(0)$ are independent of n.

8.1.1 The Continuous-Time Limit

Now let us perform the following experiment. Suppose we keep T fixed but let n, the number of intervening increments of length Δt, go to infinity. Can we say something else about the distribution of $z(T) - z(0)$ besides what its first two moments are? The answer is yes. Assuming that the $\tilde{\epsilon_i}$ are independent and identically distributed, we can state

$$p\lim_{n\to\infty}(z(T)-z(0)) \;=\; p\lim_{\Delta\, t\to 0}(z(T)-z(0)) \;\sim\; N(0,\,T) \tag{8.6}$$

In other words, $z(T) - z(0)$ has a *normal distribution* with mean zero and variance T. This follows from the *Central Limit Theorem*, which states that the sum of n independent, identically distributed random variables has a distribution that converges to the normal distribution as $n \to \infty$. Thus, the distribution of $z(t)$ over any finite interval, $[\,0,\,T\,]$, can be thought of as the sum of infinitely many small independent increments, $\Delta\, z_i = \sqrt{\Delta t}\,\tilde{\epsilon}_i$, which are realizations from an arbitrary distribution. However, when added together, these increments result in a normal distribution. Therefore, without loss of generality, we can assume that each of the $\tilde{\epsilon}_i$ have a standard (mean 0, variance 1) normal distribution.[5]

The limit of one of these minute independent increments can be defined as

$$dz(t) \;\equiv\; \lim_{\Delta\, t\to 0}\Delta z \;=\; \lim_{\Delta t\to 0}\sqrt{\Delta\, t}\tilde{\epsilon} \tag{8.7}$$

where $\tilde{\epsilon} \sim N(0,\,1)$. Hence, $E[\,dz(t)\,] = 0$ and $Var[\,dz(t)\,] = dt$.[6] dz is referred to as a pure Brownian motion process, or a Wiener process, named after the mathematician Norbert Wiener, who in 1923 first proved its existence. We can now write the change in $z(t)$ over any finite interval $[\,0,\,T\,]$ as

$$z(T) - z(0) \;=\; \int_0^T dz(t) \;\sim\; N(0,\,T) \tag{8.8}$$

The integral in (8.8) is a *stochastic* or Itô integral, not the usual Riemann or

[5] Note that sums of normally distributed random variables are also normally distributed. Thus, the Central Limit Theorem also applies to sums of normals.

[6] That the $Var\,[dz\,(t)] = dt$ can be confirmed by noting that the sum of the variance over the interval from 0 to T is $\int_0^T dt = T$.

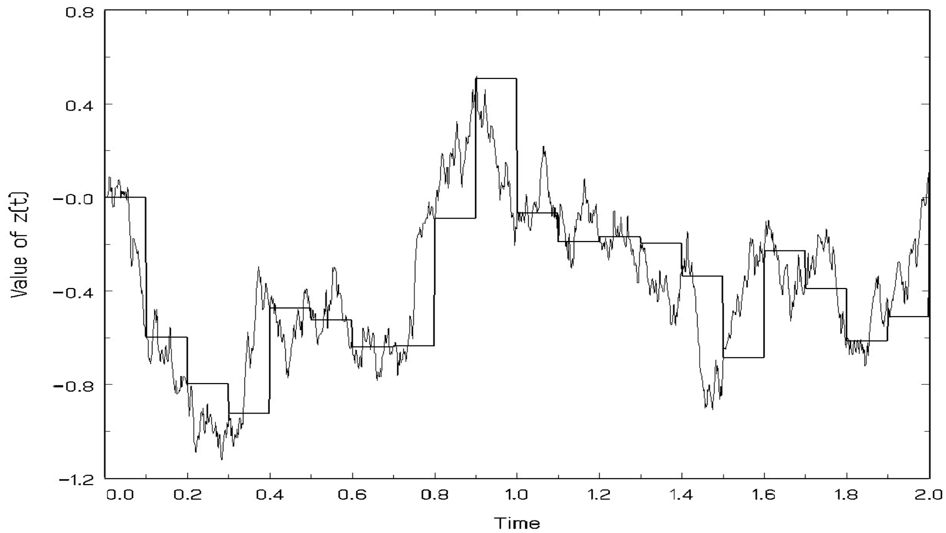

Figure 8.1: Random Walk and Brownian Motion

Lebesgue integrals that measure the area under deterministic functions.[7] Note that $z(t)$ is a continuous process but constantly changing (by $\tilde{\epsilon}$ over each infinitesimal interval Δt), such that over any finite interval it has unbounded variation.[8] Hence, it is nowhere differentiable (very jagged); that is, its derivative $dz(t)/dt$ does not exist.

The step function in Figure 8.1 illustrates a sample path for $z(t)$ as a discrete-time random walk process with $T = 2$ and $n = 20$, so that $\Delta t = 0.1$. As $n \to \infty$, so that $\Delta t \to 0$, this random walk process becomes the continuous-time Brownian motion process also shown in the figure.

Brownian motion provides the basis for more general continuous-time stochastic processes. We next analyze such processes known as *diffusion* processes. Diffusion processes are widely used in financial economics and are characterized as continuous-time Markov processes having continuous sample paths.[9]

[7] Kiyoshi Itô was a Japanese mathematician who developed the calculus of stochastic processes (Itô 1944), (Itô 1951).

[8] This means that if you measured the length of the continuous process's path over a finite interval, it would be infinitely long.

[9] A stochastic process is said to be Markov if the date t probability distribution of its future date $T > t$ value depends only on the process's date t value, and not values at prior dates $s < t$. In other words, the process's future states are conditionally independent of past

8.2 Diffusion Processes

To illustrate how we can build on the basic Wiener process, consider the process for dz multiplied by a constant, σ. Define a new process $x(t)$ by

$$dx(t) = \sigma\, dz(t) \tag{8.9}$$

Then over a discrete interval, $[0, T]$, $x(t)$ is distributed

$$x(T) - x(0) = \int_0^T dx = \int_0^T \sigma\, dz(t) = \sigma \int_0^T dz(t) \sim N(0, \sigma^2 T) \tag{8.10}$$

Next, consider adding a deterministic (nonstochastic) change of $\mu(t)$ per unit of time to the $x(t)$ process:

$$dx = \mu(t)dt + \sigma dz \tag{8.11}$$

Now over any discrete interval, $[0, T]$, we have

$$\begin{aligned} x(T) - x(0) &= \int_0^T dx = \int_0^T \mu(t)dt + \int_0^T \sigma\, dz(t) \\ &= \int_0^T \mu(t)dt + \sigma \int_0^T dz(t) \sim N(\int_0^T \mu(t)dt,\, \sigma^2 T) \end{aligned} \tag{8.12}$$

For example, if $\mu(t) = \mu$, a constant, then $x(T) - x(0) = \mu T + \sigma \int_0^T dz(t) \sim N(\mu T, \sigma^2 T)$. Thus, we have been able to generalize the standard trendless Wiener process to have a nonzero mean as well as any desired variance. The process $dx = \mu dt + \sigma dz$ is referred to as arithmetic Brownian motion.

In general, both μ and σ can be time varying. We permit them to be functions of calendar time, t, and/or functions of the contemporaneous value of the random variable, $x(t)$. In this case, the *stochastic differential equation*

---

states, given information on its current state.

describing $x(t)$ is

$$dx(t) = \mu[x(t), t]\,dt + \sigma[x(t), t]\,dz \tag{8.13}$$

and is a continuous-time Markov process, by which we mean that the instantaneous change in the process at date t has a distribution that depends only on t and the current level of the state variable $x(t)$, and not on prior values of the $x(s)$, for $s < t$. The function $\mu[x(t), t]$, which denotes the process's instantaneous expected change per unit time, is referred to as the process's *drift*, while the instantaneous standard deviation per unit time, $\sigma[x(t), t]$, is described as the process's *volatility*.

The process in equation (8.13) can also be written in terms of its corresponding integral equation:

$$x(T) - x(0) = \int_0^T dx = \int_0^T \mu[x(t), t]\,dt + \int_0^T \sigma[x(t), t]\,dz \tag{8.14}$$

In this general case, $dx(t)$ could be described as being *instantaneously* normally distributed with mean $\mu[x(t), t]\,dt$ and variance $\sigma^2[x(t), t]\,dt$, but over any finite interval, $x(t)$ generally will not be normally distributed. One needs to know the functional form of $\mu[x(t), t]$ and $\sigma[x(t), t]$ to determine the discrete-time distribution of $x(t)$ implied by its continuous-time process. Shortly, we will show how this discrete-time probability can be derived.

8.2.1 Definition of an Itô Integral

An Itô integral is formally defined as a mean-square limit of a sum involving the discrete Δz_i processes. For example, when $\sigma[x(t), t]$ is a function of $x(t)$ and t, the Itô integral in equation (8.14), $\int_0^T \sigma[x(t), t]\,dz$, is defined from the relationship

$$\lim_{n\to\infty} E_0\left[\left(\sum_{i=1}^{n}\sigma\left[x\left(\left[i-1\right]\cdot\Delta t\right),\left[i-1\right]\cdot\Delta t\right]\Delta z_i-\int_0^T\sigma[x(t),\,t]\,dz\right)^2\right]=0 \tag{8.15}$$

where we see that within the parentheses of (8.15) is the difference between the Itô integral and its discrete-time approximation. An important Itô integral that will be used next is $\int_0^T\,[dz\,(t)]^2$. In this case, (8.15) gives its definition as

$$\lim_{n\to\infty} E_0\left[\left(\sum_{i=1}^{n}\left[\Delta z_i\right]^2-\int_0^T\,\left[dz\,(t)\right]^2\right)^2\right]=0 \tag{8.16}$$

To better understand the properties of $\int_0^T\,[dz\,(t)]^2$, recall from (8.5) that

$$\begin{aligned} Var_0\left[z\left(T\right)-z\left(0\right)\right] &= Var_0\left[\sum_{i=1}^{n}\Delta z_i\right]=E_0\left[\left(\sum_{i=1}^{n}\Delta z_i\right)^2\right] \\ &= E_0\left[\sum_{i=1}^{n}\left[\Delta z_i\right]^2\right]=T \end{aligned} \tag{8.17}$$

because increments of z are serially uncorrelated. Further, straightforward algebra shows that[10]

$$E_0\left[\left(\sum_{i=1}^{n}\left[\Delta z_i\right]^2-T\right)^2\right]=2T\Delta t \tag{8.18}$$

Hence, taking the limit as $\Delta t\to 0$, or $n\to\infty$, of the expression in (8.18), one obtains

[10] This calculation uses the result that $E_0\left[\left(\Delta z_i\right)^2\left(\Delta z_j\right)^2\right]=\left(\Delta t\right)\left(\Delta t\right)=\left(\Delta t\right)^2$ for $i\neq j$ and $E_0\left[\left(\Delta z_i\right)^4\right]=3\left(\Delta t\right)^2$ because the fourth moment of a normally distributed random variable equals 3 times its squared variance.

$$\lim_{n\to\infty} E_0\left[\left(\sum_{i=1}^{n}[\Delta z_i]^2 - T\right)^2\right] = \lim_{\Delta t\to 0} 2T\Delta t = 0 \tag{8.19}$$

Comparing (8.16) with (8.19) implies that in the sense of mean-square convergence, we have the equality

$$\begin{aligned}\int_0^T [dz(t)]^2 &= T \\ &= \int_0^T dt\end{aligned} \tag{8.20}$$

Since $\int_0^T [dz(t)]^2$ converges to $\int_0^T dt$ for any T, we can see that over an infinitesimally short time period, $[dz(t)]^2$ converges to dt.

To further generalize continuous-time processes, suppose that we have some variable, F, that is a function of the current value of a diffusion process, $x(t)$, and (possibly) also is a direct function of time. Can we then characterize the stochastic process followed by $F(x(t), t)$, which now depends on the diffusion process, $x(t)$? The answer is yes, and Itô's lemma shows us how to do it.

8.3 Functions of Continuous-Time Processes and Itô's Lemma

Itô's lemma also is known as the *fundamental theorem of stochastic calculus.* It gives the rule for finding the differential of a function of variables that follow stochastic differential equations containing Wiener processes. Here we state Itô's lemma for the case of a function of a single variable that follows a diffusion process.

Itô's Lemma (*univariate case*): Let the variable $x(t)$ follow the stochastic differential equation $dx(t) = \mu(x,t)\,dt + \sigma(x,t)\,dz$. Also let $F(x(t),\,t)$ be at least a twice-differentiable function. Then the differential of $F(x,\,t)$ is given by

$$dF = \frac{\partial F}{\partial x}\,dx + \frac{\partial F}{\partial t}\,dt + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}\,(dx)^2 \tag{8.21}$$

where the product $(dx)^2 = \sigma(x,t)^2 dt$. Hence, substituting in for dx and $(dx)^2$, (8.21) can be rewritten:

$$dF = \left[\frac{\partial F}{\partial x}\,\mu(x,t) + \frac{\partial F}{\partial t} + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}\,\sigma^2(x,t)\right] dt + \frac{\partial F}{\partial x}\,\sigma(x,t)\,dz \tag{8.22}$$

Proof: A formal proof is rather lengthy and only a brief, intuitive outline of a proof is given here.[11] Let us first expand $F\left(x(t+\Delta t), t+\Delta t\right)$ in a Taylor series around date t and the value of x at date t:

$$\begin{aligned} F\left(x(t+\Delta t), t+\Delta t\right) &= F(x\left(t\right),\,t) + \frac{\partial F}{\partial x}\Delta x + \frac{\partial F}{\partial t}\Delta t + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}\left(\Delta x\right)^2 \\ &\quad + \frac{\partial^2 F}{\partial x \partial t}\Delta x \Delta t + \frac{1}{2}\frac{\partial^2 F}{\partial t^2}\left(\Delta t\right)^2 + H \end{aligned} \tag{8.23}$$

where $\Delta x \equiv x(t+\Delta t) - x\left(t\right)$ and H refers to terms that are multiplied by higher orders of Δx and Δt. Now a discrete-time approximation of Δx can be written as

$$\Delta x = \mu(x,t)\,\Delta t + \sigma(x,t)\,\sqrt{\Delta t}\tilde{\epsilon} \tag{8.24}$$

Defining $\Delta F \equiv F\left(x(t+\Delta t), t+\Delta t\right) - F(x\left(t\right),\,t)$ and substituting (8.24) in for

[11] For more details, see Chapter 3 of Merton (Merton 1992), Chapter 16 of Ingersoll (Ingersoll 1987), or Chapter 10 in Neftci (Neftci 1996). A rigorous proof is given in Karatzas and Shreve (Karatzas and Shreve 1991).

Δx, equation (8.23) can be rewritten as

$$\begin{aligned}
\Delta F &= \frac{\partial F}{\partial x}\left(\mu(x,t)\,\Delta t + \sigma(x,t)\,\sqrt{\Delta t}\tilde{\epsilon}\right) + \frac{\partial F}{\partial t}\Delta t \\
&+\frac{1}{2}\frac{\partial^2 F}{\partial x^2}\left(\mu(x,t)\,\Delta t + \sigma(x,t)\,\sqrt{\Delta t}\tilde{\epsilon}\right)^2 \\
&+\frac{\partial^2 F}{\partial x \partial t}\left(\mu(x,t)\,\Delta t + \sigma(x,t)\,\sqrt{\Delta t}\tilde{\epsilon}\right)\Delta t + \frac{1}{2}\frac{\partial^2 F}{\partial t^2}\left(\Delta t\right)^2 + H
\end{aligned} \tag{8.25}$$

The final step is to consider the limit of equation (8.25) as Δt becomes infinitesimal; that is, $\Delta t \rightarrow dt$ and $\Delta F \rightarrow dF$. Recall from (8.7) that $\sqrt{\Delta t}\tilde{\epsilon}$ becomes dz and from (8.20) that $\left[\sqrt{\Delta t}\tilde{\epsilon}\right]\left[\sqrt{\Delta t}\tilde{\epsilon}\right]$ becomes $\left[dz\,(t)\right]^2$ and converges to dt. Furthermore, it can be shown that all terms of the form $\left(\Delta t\right)^n$ where $n > 1$ go to zero as $\Delta t \rightarrow dt$. Hence, terms that are multiplied by $\left(\Delta t\right)^{\frac{3}{2}}, \left(\Delta t\right)^2, \left(\Delta t\right)^{\frac{5}{2}}$, . . . , including all of the terms in H, vanish. The result is equation (8.22). Similar arguments show that[12]

$$\begin{aligned}
(dx)^2 &= \left(\mu(x,t)\,dt + \sigma(x,t)\,dz\right)^2 \\
&= \sigma(x,t)^2\left(\,dz\right)^2 = \sigma(x,t)^2 dt
\end{aligned} \tag{8.26}$$

Note from (8.22) that the dF process is similar to the dx process in that both depend on the same Brownian motion dz. Thus, while dF will have a mean (drift) and variance (volatility) that differs from dx, they both depend on the same source of uncertainty.

[12] Thus, it may be helpful to remember that in the continuous-time limit, $(dz)^2 = dt$ but $dzdt = 0$ and $dt^n = 0$ for $n > 1$. This follows from thinking of the discrete approximation of dz as being proportional to $\sqrt{\Delta t}$, and any product that results in $(\Delta t)^n$ will go to zero as $\Delta t \rightarrow dt$ when n is strictly greater than 1.

8.3.1 Geometric Brownian Motion

A process that is used in many applications is the geometric Brownian motion process. It is given by

$$dx \;=\; \mu x\, dt + \sigma x\, dz \tag{8.27}$$

where μ and σ are constants. It is an attractive process because if x starts at a positive value, it always remains positive. This is because its mean and variance are both proportional to its current value, x. Hence, a process like dx is often used to model the price of a limited-liability security, such as a common stock. Now consider the following function $F(x,t) = \ln(x)$. For example, if x is a security's price, then $dF = d(\ln x)$ represents this security's continuously compounded rate of return. What type of process does $dF = d(\ln x)$ follow? Applying Itô's lemma, we have

$$\begin{aligned} dF \;&=\; d(\ln x) \;=\; \left[\frac{\partial(\ln x)}{\partial x}\,\mu x \;+\; \frac{\partial(\ln x)}{\partial t} \;+\; \frac{1}{2}\frac{\partial^2(\ln x)}{\partial x^2}\,(\sigma x)^2\right] dt \\ &\quad + \frac{\partial(\ln x)}{\partial x}\,\sigma x\, dz \\ &=\; \left[\mu \;+\; 0 \;-\; \frac{1}{2}\sigma^2\right] dt \;+\; \sigma\, dz \end{aligned} \tag{8.28}$$

Thus, we see that if x follows *geometric* Brownian motion, then $F = \ln x$ follows *arithmetic* Brownian motion. Since we know that

$$F(T) \;-\; F(0) \;\sim\; N\left((\mu - \frac{1}{2}\sigma^2)\,T,\; \sigma^2 T\right) \tag{8.29}$$

then $x(t) = e^{F(t)}$ has a lognormal distribution over any discrete interval (by the definition of a lognormal random variable). Hence, geometric Brownian motion is lognormally distributed over any time interval.

Figure 8.2 illustrates 300 simulated sample paths of geometric Brownian

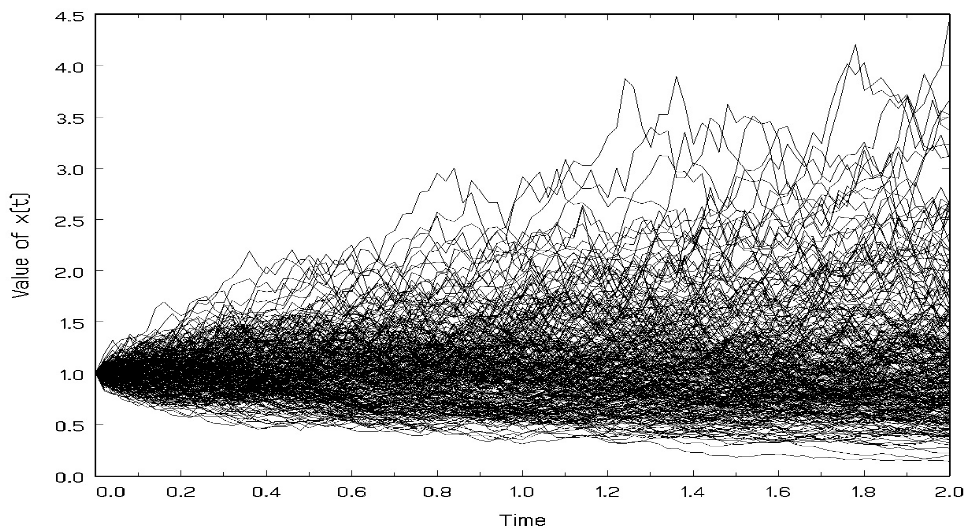

Figure 8.2: Geometric Brownian Motion Sample Paths

motion for $\mu = 0.10$ and $\sigma = 0.30$ over the period from $t = 0$ to 2, with $x(0) = 1$. These drift and volatility values are typical for a U.S. common stock. As the figure shows, the sample paths determine a frequency distribution at $T = 2$, which is skewed upward, as it should be since the discrete-time distribution is lognormal and bounded at zero.

8.3.2 Kolmogorov Equation

There are many instances where knowledge of a diffusion process's discrete-time probability distribution is very useful. As we shall see in future chapters, valuing a contingent claim often entails computing an expected value of its discounted terminal payoff at a specific future date. This discounted terminal payoff frequently depends on the value of a diffusion process, so that computing its expected value requires knowledge of the process's discrete-time probability distribution. Another situation where it is helpful to know a diffusion process's discrete-time distribution occurs when one wishes to estimate the process's drift and volatility parameters using time series data. Because time series data is typically sampled discretely rather than continuously, empirical techniques such as maximum likelihood estimation often require use of the process's discrete-time distribution.

A general method for finding the implied discrete-time probability distribution for a continuous-time process is to use the backward Kolmogorov equation. A heuristic derivation of this condition is as follows. Let $x(t)$ follow the general diffusion process given by equation (8.13). Also let $p(x,T;x_t,t)$ be the probability density function for x at date T given that it equals x_t at date t, where $T \geq t$. Applying Itô's lemma to this density function, one obtains[13]

$$dp = \left[\frac{\partial p}{\partial x_t}\,\mu(x_t,t) + \frac{\partial p}{\partial t} + \frac{1}{2}\frac{\partial^2 p}{\partial x_t^2}\,\sigma^2(x_t,t)\right]dt + \frac{\partial p}{\partial x_t}\,\sigma(x_t,t)\,dz \qquad (8.30)$$

Intuitively, one can see that only new information that was unexpected at date t should change the probability density of x at date T. In other words, for small $\Delta < T-t$, $E\left[p\left(x,T;x_{t+\Delta},t+\Delta\right)|x(t)=x_t\right] = p(x,T;x_t,t)$.[14] This implies that the expected change in p should be zero; that is, the drift term in (8.30) should be zero:

$$\frac{1}{2}\sigma^2\left(x_t,t\right)\frac{\partial^2 p}{\partial x_t^2} + \mu[x_t,\,t]\frac{\partial p}{\partial x_t} + \frac{\partial p}{\partial t} = 0 \qquad (8.31)$$

Condition (8.31) is referred to as the backward Kolmogorov equation. This partial differential equation for $p(x,T;x_t,t)$ can be solved subject to the boundary condition that when t becomes equal to T, then x must equal x_t with probability 1. Formally, this boundary condition can be written as $p(x,t;x_t,t) = \delta(x-x_t)$, where $\delta(\cdot)$ is the Dirac delta function, which is defined as $\delta(0)=\infty$, $\delta(y)=0$

[13] In order to invoke Itô's lemma, we assume that the density function $p(x,T,x_t,t)$ is differentiable in t and twice differentiable in x_t. Under particular conditions, the differentiability of p can be proved, but this issue will not be dealt with here.

[14] Essentially, this result derives from the *Law of Iterated Expectations*.

for all $y \neq 0$, and $\int_{-\infty}^{\infty} \delta(y)\, dy = 1$.

For example, recall that if x_t follows geometric Brownian motion, then $\mu[x_t, t] = \mu x_t$ and $\sigma^2(x_t, t) = \sigma^2 x_t^2$ where μ and σ are constants. In this case the Kolmogorov equation becomes

$$\frac{1}{2}\sigma^2 x_t^2 \frac{\partial^2 p}{\partial x_t^2} + \mu x_t \frac{\partial p}{\partial x_t} + \frac{\partial p}{\partial t} = 0 \tag{8.32}$$

By substitution into (8.32), it can be verified that the solution to this partial differential equation subject to the boundary condition that $p(x, t; x_t, t) = \delta(x - x_t)$ is[15]

$$p(x, T, x_t, t) = \frac{1}{x\sqrt{2\pi\sigma^2 (T-t)}} \exp\left[-\frac{\left(\ln x - \ln x_t - \left(\mu - \frac{1}{2}\sigma^2\right)(T-t)\right)^2}{2\sigma^2 (T-t)} \right] \tag{8.33}$$

which is the lognormal probability density function for the random variable $x \in (0, \infty)$. Hence, the backward Kolmogorov equation verifies that a variable following geometric Brownian motion is lognormally distributed. For a diffusion process with general drift and volatility functions, $\mu(x,t)$ and $\sigma(x,t)$, it may not be easy or possible to find a closed-form expression solution for $p(x, T, x_t, t)$ such as in (8.33). Still, there are a number of instances where the Kolmogorov equation is valuable in deriving or verifying a diffusion's discrete-time distribution.[16]

[15] Methods for solving partial differential equations are beyond the scope of this book. However, if one makes the change in variable $y_t = \ln(x_t)$, then equation (8.32) can be transformed to a more simple partial differential equation with constant coefficients. Its solution is the probability density function of a normally distributed random variable. Reversing the change in variables to $x_t = e^{y_t}$ results in the lognormal density function.

[16] Andrew Lo (Lo 1988) provides additional examples where the backward Kolmogorov equation is used to derive discrete-time distributions. These examples include not only diffusion processes but the type of mixed jump-diffusion processes that we will examine in Chapter 11.

8.3.3 Multivariate Diffusions and Itô's Lemma

In a number of portfolio choice and asset pricing applications that we will encounter in future chapters, one needs to derive the stochastic process for a function of several variables, each of which follows a diffusion process. So suppose we have m different diffusion processes of the form[17]

$$dx_i = \mu_i \, dt + \sigma_i \, dz_i \qquad i = 1, \ldots, m, \tag{8.34}$$

and $dz_i dz_j = \rho_{ij} dt$, where ρ_{ij} has the interpretation of a correlation coefficient of the two Wiener processes. What is meant by this correlation? Recall that $dz_i dz_i = (dz_i)^2 = dt$. Now the Wiener process dz_j can be written as a linear combination of two other Wiener processes, one being dz_i, and another process that is uncorrelated with dz_i, call it dz_{iu}:

$$dz_j = \rho_{ij} dz_i + \sqrt{1 - \rho_{ij}^2} dz_{iu} \tag{8.35}$$

Then from this interpretation of dz_j, we have

$$\begin{aligned} dz_j dz_j &= \rho_{ij}^2 (dz_i)^2 + \left(1 - \rho_{ij}^2\right)(dz_{iu})^2 + 2\rho_{ij}\sqrt{1 - \rho_{ij}^2} dz_i dz_{iu} \qquad (8.36) \\ &= \rho_{ij}^2 dt + \left(1 - \rho_{ij}^2\right) dt + 0 \\ &= dt \end{aligned}$$

and

[17] Note μ_i and σ_i may be functions of calendar time, t, and the current values of x_j, $j = 1, ..., m$.

$$\begin{aligned} dz_i dz_j &= dz_i \left(\rho_{ij} dz_i + \sqrt{1-\rho_{ij}^2} dz_{iu} \right) \\ &= \rho_{ij} \left(dz_i \right)^2 + \sqrt{1-\rho_{ij}^2} dz_i dz_{iu} \\ &= \rho_{ij} dt + 0 \end{aligned} \tag{8.37}$$

Thus, ρ_{ij} can be interpreted as the proportion of dz_j that is perfectly correlated with dz_i.

We can now state, without proof, a multivariate version of Itô's lemma.

Itô's Lemma (multivariate version): Let $F(x_1, \ldots, x_m, t)$ be at least a twice-differentiable function. Then the differential of $F(x_1, \ldots, x_m, t)$ is given by

$$dF = \sum_{i=1}^{m} \frac{\partial F}{\partial x_i} dx_i + \frac{\partial F}{\partial t} dt + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 F}{\partial x_i \, \partial x_j} dx_i \, dx_j \tag{8.38}$$

where $dx_i \, dx_j = \sigma_i \sigma_j \rho_{ij} \, dt$. Hence, (8.38) can be rewritten

$$\begin{aligned} dF &= \left[\sum_{i=1}^{m} \left(\frac{\partial F}{\partial x_i} \mu_i + \frac{1}{2} \frac{\partial^2 F}{\partial x_i^2} \sigma_i^2 \right) + \frac{\partial F}{\partial t} + \sum_{i=1}^{m} \sum_{j>i}^{m} \frac{\partial^2 F}{\partial x_i \, \partial x_j} \sigma_i \sigma_j \rho_{ij} \right] dt \\ &\quad + \sum_{i=1}^{m} \frac{\partial F}{\partial x_i} \sigma_i \, dz_i \end{aligned} \tag{8.39}$$

Equation (8.39) generalizes our earlier statement of Itô's lemma for a univariate diffusion, equation (8.22). Notably, we see that the process followed by a function of several diffusion processes inherits each of the processes' Brownian motions.

8.4 Summary

Diffusion processes and Itô's lemma are important tools for modeling financial time series, especially when individuals are assumed to be able to trade continuously. Brownian motion is the foundation of diffusion processes and is a continuous-time limit of a particular discrete-time random walk process. By modifying Brownian motion's instantaneous mean and variance, a wide variety of diffusion processes can be created. Itô's lemma tells us how to find the differential of a function of a diffusion process. As we shall see in the next chapter, Itô's lemma is essential for valuing a contingent claim when its payoff depends on the price of an underlying asset that follows a diffusion. This is because the contingent claim's value becomes a function of the underlying asset's value.

This chapter also showed that Itô's lemma could be used to derive the Kolmogorov equation, an important relation for finding the discrete-time distribution of a random variable that follows a diffusion process. Finally, we saw that multivariate diffusions are natural extensions of univariate ones and that the process followed by a function of several diffusions can be derived from a multivariate version of Itô's lemma.

8.5 Exercises

1. A variable, $x(t)$, follows the process

$$dx = \mu dt + \sigma dz$$

   where μ and σ are constants. Find the process followed by $y(t) = e^{\alpha x(t) - \beta t}$.

2. Let P be a price index, such as the Consumer Price Index (CPI). Let M equal the nominal supply (stock) of money in the economy. For example, M might be designated as the amount of bank deposits and currency in

circulation. Assume P and M each follow geometric Brownian motion processes

$$\begin{aligned}\frac{dP}{P} &= \mu_p dt + \sigma_p dz_p \\ \frac{dM}{M} &= \mu_m dt + \sigma_m dz_m\end{aligned}$$

with $dz_p dz_m = \rho dt$. Monetary economists define real money balances, m, to be $m = \frac{M}{P}$. Derive the stochastic process for m.

3. The value (price) of a portfolio of stocks, $S(t)$, follows a geometric Brownian motion process:

$$dS/S = \alpha_s dt + \sigma_s dz_s$$

while the dividend yield for this portfolio, $y(t)$, follows the process

$$dy = \kappa\left(\gamma S - y\right)dt + \sigma_y y^{\frac{1}{2}} dz_y$$

where $dz_s dz_y = \rho dt$ and κ, γ, and σ_y are positive constants. Solve for the process followed by the portfolio's dividends paid per unit time, $D(t) = yS$.

4. The Ornstein-Uhlenbeck process can be useful for modeling a time series whose value changes stochastically but which tends to revert to a long-run value (its unconditional or steady state mean). This continuous-time process is given by

$$dy(t) = \left[\alpha - \beta y(t)\right]dt + \sigma dz(t)$$

The process is sometimes referred to as an *elastic random walk*. $y(t)$ varies stochastically around its unconditional mean of α/β, and β is a measure of the strength of the variable's reversion to this mean. Find the distribution of $y(t)$ given $y(t_0)$, where $t > t_0$. In particular, find $E[y(t) \mid y(t_0)]$ and $Var[y(t) \mid y(t_0)]$. Hint: make the change in variables:

$$x(t) = \left(y(t) - \frac{\alpha}{\beta} \right) e^{\beta(t-t_0)}$$

and apply Itô's lemma to find the stochastic process for $x(t)$. The distribution and first two moments of $x(t)$ should be obvious. From this, derive the distribution and moments of $y(t)$.

Chapter 9

Dynamic Hedging and PDE Valuation

Having introduced diffusion processes and Itô's lemma in the previous chapter, we now apply these tools to derive the equilibrium prices of contingent claims. In this chapter asset prices are modeled as following diffusion processes. Because prices are permitted to vary continuously, it is feasible to also assume that individuals can choose to trade assets continuously. With the additional assumption that markets are "frictionless," this environment can allow the markets for a contingent claim, its underlying asset, and the risk-free asset to be dynamically complete.[1] Although the returns of the underlying asset and its contingent claim have a continuous distribution over any finite time interval, implying an infinite number of states for their future values, the future values of the contingent claim can be replicated by a dynamic trading strategy involving its underlying asset and the risk-free asset.

[1]Frictionless markets are characterized as having no direct trading costs or restrictions, that is, markets for which there are no transactions costs, taxes, short sales restrictions, or indivisibilities when trading assets.

By way of three examples, we illustrate the Black-Scholes-Merton portfolio hedging argument that results in a partial differential equation (PDE) for a contingent claim's price. Solving this PDE subject to the appropriate boundary condition determines a unique price for the contingent security. Our first example is the well-known Fischer Black-Myron Scholes (Black and Scholes 1973) option pricing model. The second is the equilibrium term structure model of Oldrich Vasicek (Vasicek 1977). The final example combines aspects of the first two. It is Robert Merton's (Merton 1973b) option pricing model with stochastic interest rates.

As the next chapter will show, contingent claims prices also can be derived using alternative solution techniques: the martingale pricing approach, which involves computing expectations of a risk-neutral probability distribution; and the stochastic discount factor (pricing kernel) approach, where expectations are computed for the physical probability distribution. In some situations, it may be easier to derive contingent claims prices by solving the equilibrium PDE. In others, the martingale technique or stochastic discount factor approach may be simplest. All of these methods should be in a financial economist's toolbox.

9.1 Black-Scholes Option Pricing

The major insight of Black and Scholes (Black and Scholes 1973) is that when assets follow diffusion processes, an option's payoff can be replicated by continuous trading in its underlying asset and a risk-free asset. In the absence of arbitrage, the ability to replicate or "hedge" the option with the underlying stock and a risk-free asset restricts the option's value to bear a particular relationship to its underlying asset and the risk-free return. The Black-Scholes hedging argument is similar to that presented earlier in the context of the binomial option pricing model. The main difference is that the appropriate replicating portfolio changed only once per period in the binomial model, whereas in the Black-Scholes environment the replicating portfolio changes continuously. In the binomial model, market completion resulted from the assumption that at the end of each period there were only two states for the underlying asset's value. Under Black-Scholes assumptions, markets become dynamically complete due to the ability to trade continuously in the underlying asset whose price follows a continuous sample path.

9.1.1 Portfolio Dynamics in Continuous Time

A prerequisite for analyzing the Black-Scholes hedging of contingent claims is to consider the dynamics of a security portfolio in continuous time. The Black-Scholes hedge portfolio consists of a position in the contingent claim and its underlying asset, but we will begin by examining the general problem of an investor who can trade in any n different assets whose prices follow diffusion processes. Let us define $S_i(t)$ as the price per share of asset i at date t, where $i = 1, ..., n$. The instantaneous rate of return on the i^{th} asset is assumed to satisfy the process

$$dS_i(t) / S_i(t) = \mu_i \, dt + \sigma_i \, dz_i \tag{9.1}$$

where its instantaneous expected return and variance, μ_i and σ_i^2, may be functions of time and possibly other asset prices or state variables that follow diffusion processes. For simplicity, assets are assumed to pay no cashflows (dividends or coupon payments), so that their total returns are given by their price changes.[2] An investor is assumed to form a portfolio of these assets and, in general, the portfolio may experience cash inflows and outflows. Thus, let $F(t)$

[2] This is not a critical assumption. What matters is the assets' expected rates of return and covariances, rather than their price changes per se. If an asset, such as a common stock or mutual fund, paid a dividend that was reinvested into new shares of the asset, then equation (9.1) would represent the percentage change in the value of the asset holding and thus the total rate of return.

be the net cash outflow per unit time from the portfolio at date t. For example, $F(t)$ may be positive because the individual chooses to liquidate some of the portfolio to pay for consumption expenditures. Alternatively, $F(t)$ may be negative because the individual receives wage income that is invested in the securities.

To derive the proper continuous-time dynamics for this investor's portfolio, we will first consider the analogous discrete-time dynamics where each discrete period is of length h. We will then take the limit as $h \to 0$. Therefore, let $w_i(t)$ be the number of shares held by the investor in asset i from date t to $t+h$. The value of the portfolio at the beginning of date t is denoted as $H(t)$ and equals the prior period's holdings at date t prices:

$$H(t) = \sum_{i=1}^{n} w_i(t-h) S_i(t) \tag{9.2}$$

Given these date t prices, the individual may choose to liquidate some of the portfolio or augment it with new funds. The net cash outflow over the period is $F(t)\ h$, which must equal the net sales of assets. Note that $F(t)$ should be interpreted as the average liquidation rate over the interval from t to $t+h$:

$$-F(t)\ h = \sum_{i=1}^{n} [w_i(t) - w_i(t-h)]\, S_i(t) \tag{9.3}$$

To properly derive the limits of equations (9.2) and (9.3) as of date t and as $h \to 0$, we need to convert backward differences, such as $w_i(t) - w_i(t-h)$, to forward differences. We do this by updating one period, so that at the start of the next period, $t+h$, we have

$$\begin{aligned} -F\,(t+h)\; h &= \sum_{i=1}^{n}[w_i(t+h)-w_i(t)]\,S_i(t+h) \\ &= \sum_{i=1}^{n}[w_i(t+h)-w_i(t)]\,[S_i(t+h)-S_i(t)] \\ &\quad + \sum_{i=1}^{n}[w_i(t+h)-w_i(t)]\,S_i(t) \end{aligned} \tag{9.4}$$

and

$$H\,(t+h) \;=\; \sum_{i=1}^{n} w_i(t)\,S_i(t+h) \tag{9.5}$$

Taking the limits of (9.4) and (9.5) as $h \to 0$ gives the results

$$-F\,(t)\;dt \;=\; \sum_{i=1}^{n} dw_i(t)\,dS_i(t) \;+\; \sum_{i=1}^{n} dw_i(t)\,S_i(t) \tag{9.6}$$

and

$$H\,(t) \;=\; \sum_{i=1}^{n} w_i(t)\,S_i(t) \tag{9.7}$$

Applying Itô's lemma to (9.7), we can derive the dynamics of the portfolio's value to be

$$dH\,(t) \;=\; \sum_{i=1}^{n} w_i(t)\,dS_i(t) \;+\; \sum_{i=1}^{n} dw_i(t)S_i(t) \;+\; \sum_{i=1}^{n} dw_i(t)\,dS_i(t) \tag{9.8}$$

Substituting (9.6) into (9.8), we obtain

$$dH\,(t) \;=\; \sum_{i=1}^{n} w_i(t)\,dS_i(t) \;-F\,(t)\;dt \tag{9.9}$$

Equation (9.9) says that the portfolio's value changes due to capital gains income less net cash outflows. Substituting (9.1) into (9.9), we arrive at

$$\begin{aligned} dH(t) &= \sum_{i=1}^{n} w_i(t)\, dS_i(t) \;- F(t)\; dt \\ &= \sum_{i=1}^{n} w_i(t)\; [\mu_i\, S_i dt \;+\; \sigma_i S_i\, dz_i] \;- F(t)\; dt \end{aligned} \tag{9.10}$$

Now, in some cases, rather than write a portfolio's dynamics in terms of the number of shares of each asset, $w_i(t)$, $i = 1, \ldots, n$, we may wish to write it in terms of each asset's proportion of the total portfolio value. If we define the proportion of $H(t)$ invested in asset i as $\omega_i(t) = w_i(t)S_i(t)/H(t)$, then (9.10) becomes

$$dH(t) \;= \sum_{i=1}^{n} \omega_i(t)\, H(t)\; [\mu_i dt \;+\; \sigma_i\, dz_i] \;- F(t)\; dt \tag{9.11}$$

or

$$dH(t) = \left[\sum_{i=1}^{n} \omega_i(t)\, H(t)\, \mu_i \;- F(t) \right] dt \;+\; \sum_{i=1}^{n} \omega_i(t)\, H(t)\, \sigma_i\, dz_i \tag{9.12}$$

Note from (9.7) that $\sum_{i=1}^{n} \omega_i(t) = 1$; that is, the portfolio proportions invested in the n risky assets must sum to 1. However, consider the introduction of a new risk-free asset. If, in addition to n risky assets, there is an asset that pays an instantaneously risk-free rate of return, then this would correspond to an asset having an instantaneous standard deviation, σ_i, of zero and an expected rate of return, μ_i, equal to the instantaneous risk-free rate, which we denote as $r(t)$. In this case, the portfolio proportion invested in the risk-free asset equals $1 - \sum_{i=1}^{n} \omega_i(t)$. With this extension, equation (9.12) becomes

$$dH(t) = \left[\sum_{i=1}^{n} \omega_i(t)\, (\mu_i - r)\, H(t) \;+ rH(t) - F(t) \right] dt + \sum_{i=1}^{n} \omega_i(t)\, H(t)\, \sigma_i\, dz_i \tag{9.13}$$

Having derived the continuous-time dynamics of an investment portfolio, we now turn to the Black-Scholes approach to valuing contingent claims.

9.1.2 Black-Scholes Model Assumptions

The Black-Scholes model assumes that there is a contingent claim whose underlying asset pays no dividends. We will refer to this underlying asset as a stock, and its date t price per share, $S(t)$, is assumed to follow the diffusion process

$$dS = \mu S\,dt + \sigma S\,dz \tag{9.14}$$

where the instantaneous expected rate of return on the stock, μ, may be a function of S and t, that is, $\mu(S,t)$. However, the standard deviation of the stock's rate of return, σ, is assumed to be constant. It is also assumed that there is a risk-free asset that earns a constant rate of return equal to r per unit time. Hence, if an amount $B(t)$ is invested in the risk-free asset, this value follows the process

$$dB = rBdt \tag{9.15}$$

Now consider a European call option on this stock that matures at date T and has an exercise price of X. Denote the option's date t value as $c(S,\,t)$. We assume it is a function of both calendar time, t, and the current stock price, $S(t)$, since at the maturity date $t = T$, the option's payoff depends on $S(T)$:

$$c(S(T),T) = \max[0,\,S(T) - X] \tag{9.16}$$

Given that the option's value depends on the stock price and calendar time, what process does it follow prior to maturity?[3] Let us assume that $c(S,t)$ is a twice-differentiable function of S and is differentiable in t. Later, we will

[3] The option's value also depends on the risk-free rate, r, but since r is assumed to be constant, it need not be an explicit argument of the option's value.

verify that the no-arbitrage value of $c(S,t)$ does indeed satisfy these conditions. Then we can apply Itô's lemma to state that the option's value must follow a process of the form

$$dc = \left[\frac{\partial c}{\partial S} \mu S + \frac{\partial c}{\partial t} + \frac{1}{2} \frac{\partial^2 c}{\partial S^2} \sigma^2 S^2 \right] dt + \frac{\partial c}{\partial S} \sigma S \, dz \tag{9.17}$$

Hence, the call option inherits the same source of risk as the underlying stock, reflected in the Wiener process dz.

9.1.3 The Hedge Portfolio

Now consider forming a portfolio that includes -1 unit of the option and a position in the underlying stock and the risk-free asset. Such a portfolio would reflect the wealth position of an option dealer who has just sold one call option to a customer and now attempts to hedge this liability by purchasing some of the underlying stock and investing or borrowing at the risk-free rate. We restrict this portfolio to require zero net investment; that is, after selling one unit of the call option and taking a hedge position in the underlying stock, the remaining surplus or deficit of funds is made up by borrowing or lending at the risk-free rate. Moreover, we require that the portfolio be self-financing, that is, $F(t) = 0$ $\forall t$, by which we mean that any surplus or deficit of funds from the option and stock positions are made up by investing or acquiring funds at the risk-free rate. Hence, if we let $w(t)$ be the number of shares invested in the stock, then this zero-net-investment, self-financing restriction implies that the amount invested in the risk-free asset for all dates t must be $B(t) = c(t) - w(t) S(t)$. Therefore, denoting the value of this hedge portfolio as $H(t)$ implies that its instantaneous return satisfies

$$dH(t) = -dc(t) + w(t) \, dS(t) + [c(t) - w(t) S(t)] \, r dt \tag{9.18}$$

Substituting (9.14) and (9.17) into (9.18), we obtain

$$\begin{aligned} dH(t) &= -\left[\frac{\partial c}{\partial S}\mu S + \frac{\partial c}{\partial t} + \frac{1}{2}\frac{\partial^2 c}{\partial S^2}\sigma^2 S^2\right] dt - \frac{\partial c}{\partial S}\sigma S\, dz \\ &\quad + w(t)(\mu S\, dt + \sigma S\, dz) + [c(t) - w(t) S(t)]\, rdt \end{aligned} \tag{9.19}$$

Now consider selecting the number of shares invested in the stock in such a way as to offset the risk of the return on the option. Specifically, suppose that the option dealer chooses $w(t) = \partial c/\partial S$ units (shares) of the stock, which is the local sensitivity of the option's value to the value of the underlying stock, also known as the "hedge ratio."[4] Hence, the hedging portfolio involves maintaining a unit short position in the option and a position of $\partial c/\partial S$ shares of stock, with any surplus or deficit of funds required to maintain this hedge being invested or acquired at the risk-free rate. As will be verified, since $c(S,t)$ is a nonlinear function of S and t, $w(t) = \partial c/\partial S$ varies continuously over time as S and t change: the hedge portfolio's number of shares invested in the stock is not constant, but is continuously rebalanced.[5] However, as long as a position of $\partial c/\partial S$ shares of stock are held, we can substitute $w(t) = \partial c/\partial S$ into (9.19) to obtain

$$\begin{aligned} dH(t) &= -\left[\frac{\partial c}{\partial S}\mu S + \frac{\partial c}{\partial t} + \frac{1}{2}\frac{\partial^2 c}{\partial S^2}\sigma^2 S^2\right] dt - \frac{\partial c}{\partial S}\sigma S\, dz \\ &\quad + \frac{\partial c}{\partial S}(\mu S\, dt + \sigma S\, dz) + \left[c(t) - \frac{\partial c}{\partial S}S(t)\right] rdt \\ &= \left[-\frac{\partial c}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 c}{\partial S^2} + rc(t) - rS(t)\frac{\partial c}{\partial S}\right] dt \end{aligned} \tag{9.20}$$

[4] $\partial c/\partial S$ is analogous to the hedge ratio Δ in the binomial option pricing model. Recall that the optimal choice of this hedge ratio was $\Delta^* = (c_u - c_d)/(uS - dS)$, which is essentially the same partial derivative.

[5] Since $c(S,t)$ is yet to be determined, the question arises as to how $w(t) = \partial c/\partial S$ would be known to create the hedge portfolio. We will verify that if such a position in the stock is maintained, then a no-arbitrage value for the option, $c(S,t)$, is determined, which, in turn, makes known the hedge ratio $w(t) = \partial c/\partial S$.

Note that, by design, the return on this portfolio is instantaneously riskless. Not only do the dz terms in the first line of (9.20) drop out but so do the terms that depend on the stock's drift, μ. By continually readjusting the number of shares held in the stock so that it always equals $\partial c/\partial S$, the risk of the option is perfectly hedged. Dynamic trading in the stock is able to replicate the risk of the option because both the option and stock depend on the same (continuous-time) Brownian motion process, dz. In this sense, when assets follow continuous-time stochastic processes, dynamic (continuous) trading can lead to a complete market and permit the pricing of contingent claims.

9.1.4 No-Arbitrage Implies a PDE

Since the rate of return on this "hedge" portfolio is riskless, to avoid arbitrage it must equal the competitive risk-free rate of return, r. But since we restricted the hedge portfolio to require zero net investment at the initial date, say, $t = 0$, then $H(0) = 0$ and

$$dH(0) = rH(0)\,dt = r0dt = 0 \tag{9.21}$$

This implies $H(t) = 0\ \forall t$ so that $dH(t) = 0\ \forall t$. This no-arbitrage condition along with (9.20) allows us to write

$$\frac{\partial c}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 c}{\partial S^2} + r\,S\,\frac{\partial c}{\partial S} - r\,c = 0 \tag{9.22}$$

which is the Black-Scholes partial differential equation. The call option's value must satisfy this partial differential equation subject to the boundary condition

$$c(S(T),\,T) = \max[\,0,\,S(T) - X] \tag{9.23}$$

The solution to (9.22) and (9.23) is[6]

$$c(S(t), t) = S(t)\, N(d_1) - X\, e^{-r\,(T-t)}\, N(d_2) \tag{9.24}$$

where

$$\begin{aligned} d_1 &= \frac{\ln\left(S(t)/X\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \\ d_2 &= d_1 - \sigma\sqrt{T-t} \end{aligned} \tag{9.25}$$

and $N(\cdot)$ is the standard normal distribution function. Similar to the binomial option pricing formula, the value of the call option does not depend on the stock's expected rate of return, μ, but on only its current price, $S(t)$, and volatility, σ. The value of a European put option follows immediately from put-call parity:[7]

$$\begin{aligned} p(S(t), t) &= c(S(t), t) + X\, e^{-r\,(T-t)} - S(t) \\ &= S(t)\, N(d_1) - X\, e^{-r\,(T-t)}\, N(d_2) + X\, e^{-r\,(T-t)} - S(t) \\ &= X\, e^{-r\,(T-t)} N(-d_2) - S(t) N(-d_1) \end{aligned} \tag{9.26}$$

By taking the partial derivatives of (9.24) and (9.26) with respect to $S(t)$, the call and put options' hedge ratios are shown to be[8]

---

[6] The solution can be derived using a separation of variables method (Churchill and Brown 1978) or a LaPlace transform method (Shimko 1992). Also, in Chapter 10, we will show how (9.24) can be derived using risk-neutral valuation.

[7] The last line uses the symmetry property of the normal distribution $1 - N(x) = N(-x)$.

[8] Deriving these partial derivatives is more tedious than it might first appear since d_1 and d_2 are both functions of $S(t)$. Note that $\partial c/\partial S = N(d_1) + Sn(d_1)\frac{\partial d_1}{\partial S} - Xe^{-r(T-t)}n(d_2)\frac{\partial d_2}{\partial S}$ where $n(d) = \frac{1}{\sqrt{2\pi}}\exp\left(\frac{1}{2}d^2\right)$ is the standard normal probability density function. This reduces to (9.27) because it can be shown that $Sn(d_1)\frac{\partial d_1}{\partial S} = Xe^{-r(T-t)}n(d_2)\frac{\partial d_2}{\partial S}$. Practitioners refer to the hedge ratios in (9.27) and (9.28) as the options' *deltas*.

$$\frac{\partial c}{\partial S} = N\left(d_1\right) \tag{9.27}$$

$$\frac{\partial p}{\partial S} = -N\left(-d_1\right) \tag{9.28}$$

which implies $0 < \partial c/\partial S < 1$ and $-1 < \partial p/\partial S < 0$. Hence, hedging a call option requires a long position in less than one share of the underlying stock, whereas hedging a put option requires a short position in less than one share of the underlying stock. Since d_1 is an increasing function of $S\left(t\right)$, the hedge portfolio for a call option increases the share amount in the stock as its price rises. A similar argument shows that the hedge portfolio for a put option increases the share amount sold short as the price of the stock falls. Thus, because $S\left(t\right)$ moves in a continuous fashion, so will the hedge portfolio's position in the stock. Finally, based on the solution in (9.24), we can verify that both $\partial^2 c/\partial S^2$ and $\partial c/\partial t$ exist, which justifies our use of Itô's lemma in deriving the process followed by the option's price.[9]

We now turn to another application of the Black-Scholes-Merton hedging argument for deriving security prices. However, rather than derive the price of a contingent security in terms of an underlying asset price, we next consider pricing securities that pay known (fixed) cashflows at different future dates. That is, we derive the relationship between the prices of different maturity bonds, also known as fixed-income securities. This provides an introduction into the literature on the term structure of interest rates (or bond yields).

[9] Using (9.27) and (9.28), it is easy to see that $\partial^2 c/\partial S^2 = \partial^2 p/\partial S^2 = n\left(d_1\right)/\left[S\sigma\sqrt{T-t}\right] > 0$ where $n\left(x\right) = \partial N\left(x\right)/\partial x = e^{-x^2/2}/\sqrt{2\pi}$ is the standard normal probability density function. Hence, both call and put options are convex functions of the underlying asset price. Practitioners refer to this second derivative as the option's *gamma*. The larger an option's gamma, the larger is the required change in the hedge ratio for a given change in the underlying asset's price. The option's *theta* or *time decay* is given by $\partial c/\partial\left(T-t\right) = -Sn\left(d_1\right)\sigma/\left[2\left(T-t\right)\right] - rXe^{-r(T-t)}N\left(d_2\right)$.

9.2 An Equilibrium Term Structure Model

The previous section showed that in a continuous-time environment, the absence of arbitrage restricts a derivative's price in terms of its underlying asset's price. We now consider a second example of how the absence of arbitrage links security prices. When the prices of default-free bonds are assumed to be driven by continuous-time stochastic processes, continuous trading and the no-arbitrage condition can lead to equilibrium relationships between the prices of different maturity bonds. The simplest equilibrium bond pricing models assume that a single source of uncertainty affects bonds of all maturities. For these "one-factor" bond pricing models, it is often convenient to think of this uncertainty as being summarized by the yield on the shortest (instantaneous) maturity bond, $r(t)$.[10] This is the assumption we make in presenting the Oldrich Vasicek (Vasicek 1977) model of the term structure of interest rates.

Define $P(t,\tau)$ as the date t price of a bond that makes a single payment of \$1 in τ periods, at date $T = t + \tau$. Hence, τ denotes this "zero-coupon" or "pure discount" bond's time until maturity. The instantaneous rate of return on the bond is given by $\frac{dP(t,\tau)}{P(t,\tau)}$. Also note that, by definition, $P(t,0) = \$1$. The instantaneous yield, $r(t)$, is defined as

$$\lim_{\tau \to 0} \frac{dP(t,\tau)}{P(t,\tau)} \equiv r(t)\, dt \tag{9.29}$$

The Vasicek model assumes $r(t)$ follows an Ornstein-Uhlenbeck process:

$$dr(t) = \alpha \left[\overline{r} - r(t)\right] dt + \sigma_r dz_r \tag{9.30}$$

where α, $\overline{r}$, and σ_r are positive constants. The parameter σ_r measures the

[10] Other approaches to modeling the term structure of interest rates are considered in Chapter 17. For example, we will discuss research by David Heath, Robert Jarrow, and Andrew Morton (Heath, Jarrow, and Morton 1992) that assumes *forward* interest rates of all maturities are affected by one or more sources of risk.

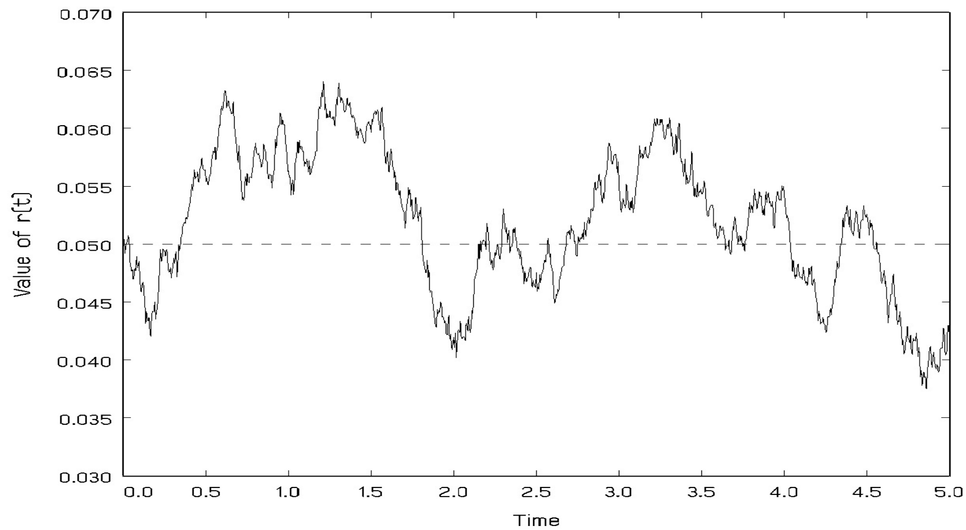

Figure 9.1: Ornstein-Uhlenbeck Interest Rate Process

instantaneous volatility of $r(t)$, while α measures the strength of the process's mean reversion to $\overline{r}$, the unconditional mean value of the process. In discrete time, (9.30) is equivalent to a normally distributed, autoregressive (1) process.[11] Figure 9.1 illustrates a typical sample path for $r(t)$ that assumes the annualized parameter values of $r(0) = \overline{r} = 0.05$, $\alpha = 0.3$, and $\sigma_r = 0.02$.

Now assume that bond prices of all maturities depend on only a single source of uncertainty and that this single "factor" is summarized by the current level of $r(t)$.[12] Then we can write a τ-maturity bond's price as $P(r(t), \tau)$, and Itô's lemma implies that it follows the process

[11] The discrete-time expected value and variance implied by the continuous-time process in (9.30) are $E_t[r(t+\tau)] = \overline{r} + e^{-\alpha\tau}(r(t) - \overline{r})$ and $Var_t[r(t+\tau)] = \frac{\sigma_r^2}{2\alpha}\left(1 - e^{-2\alpha\tau}\right)$, respectively. See exercise 4 at the end of Chapter 8.

[12] For example, a central bank may implement monetary policy by changing the level of the short-term interest rate. Other macroeconomic effects on bond prices might be summarized in the level of the short rate.

$$
\begin{aligned}
dP(r,\tau) &= \frac{\partial P}{\partial r}dr + \frac{\partial P}{\partial t}dt + \tfrac{1}{2}\frac{\partial^2 P}{\partial r^2}(dr)^2 \\
&= \left[P_r \alpha(\bar{r}-r) + P_t + \tfrac{1}{2}P_{rr}\sigma_r^2\right]dt + P_r\sigma_r dz_r \\
&= \mu_p(r,\tau)\,P(r,\tau)\,dt - \sigma_p(\tau)\,P(r,\tau)\,dz_r
\end{aligned} \tag{9.31}
$$

where the subscripts on P denote partial derivatives and where $\mu_p(r,\tau) \equiv \frac{\left[P_r\alpha(\bar{r}-r)+P_t+\frac{1}{2}P_{rr}\sigma_r^2\right]}{P(r,\tau)}$ and $\sigma_p(\tau) \equiv -\frac{P_r\sigma_r}{P(r,\tau)}$ are the mean and standard deviation, respectively, of the bond's instantaneous rate of return.[13]

Consider forming a portfolio containing one bond of maturity τ_1 and $-\frac{\sigma_p(\tau_1)P(r,\tau_1)}{\sigma_p(\tau_2)P(r,\tau_2)}$ units of a bond with maturity τ_2. In other words, we have a unit long position in a bond of maturity τ_1 and a short position in a bond with maturity τ_2 in an amount that reflects the ratio of bond 1's return standard deviation to that of bond 2's. Since both bonds are driven by the same Wiener process, dz_r, this portfolio is a hedged position. If we continually readjust the amount of the τ_2-maturity bonds to equal $-\frac{\sigma_p(\tau_1)P(r,\tau_1)}{\sigma_p(\tau_2)P(r,\tau_2)}$ as $r(t)$ changes, the value of this hedge portfolio, $H(t)$, is

$$
\begin{aligned}
H(t) &= P(r,\tau_1) - \frac{\sigma_p(\tau_1)\,P(r,\tau_1)}{\sigma_p(\tau_2)\,P(r,\tau_2)}P(r,\tau_2) \\
&= P(r,\tau_1)\left[1 - \frac{\sigma_p(\tau_1)}{\sigma_p(\tau_2)}\right]
\end{aligned} \tag{9.32}
$$

Furthermore, the hedge portfolio's instantaneous return is

[13] We define $\sigma_p(\tau) \equiv -P_r\sigma_r/P(r,\tau)$ rather than $\sigma_p(\tau) \equiv P_r\sigma_r/P(r,\tau)$ because it will turn out that $P_r < 0$. Hence, if we want both σ_r and σ_p to denote standard deviations, we need them to be positive. This choice of definition makes no material difference since the instantaneous variance of the change in the interest rate and the bond's rate of return will always be σ_r^2 and σ_p^2, respectively.

$$\begin{aligned} dH(t) &= dP(r,\tau_1) - \frac{\sigma_p(\tau_1) P(r,\tau_1)}{\sigma_p(\tau_2) P(r,\tau_2)} dP(r,\tau_2) \\ &= \mu_p(r,\tau_1) P(r,\tau_1) dt - \sigma_p(\tau_1) P(r,\tau_1) dz_r \\ &\quad - \frac{\sigma_p(\tau_1)}{\sigma_p(\tau_2)} P(r,\tau_1) \mu_p(r,\tau_2) dt + \sigma_p(\tau_1) P(r,\tau_1) dz_r \\ &= \mu_p(r,\tau_1) P(r,\tau_1) dt - \frac{\sigma_p(\tau_1)}{\sigma_p(\tau_2)} P(r,\tau_1) \mu_p(r,\tau_2) dt \end{aligned} \tag{9.33}$$

where the second equality in (9.33) reflects substitution of (9.31). Since the portfolio return is riskless at each instant of time, the absence of arbitrage implies that its rate of return must equal the instantaneous riskless interest rate, $r(t)$:

$$\begin{aligned} dH(t) &= \left[\mu_p(r,\tau_1) - \frac{\sigma_p(\tau_1)}{\sigma_p(\tau_2)} \mu_p(r,\tau_2)\right] P(r,\tau_1) dt \\ &= r(t) H(t) dt = r(t) \left[1 - \frac{\sigma_p(\tau_1)}{\sigma_p(\tau_2)}\right] P(r,\tau_1) dt \end{aligned} \tag{9.34}$$

Equating the terms that precede $P(r,\tau_1)$ on the first and second lines of (9.34), we see that an implication of this equation is

$$\frac{\mu_p(r,\tau_1) - r(t)}{\sigma_p(\tau_1)} = \frac{\mu_p(r,\tau_2) - r(t)}{\sigma_p(\tau_2)} \tag{9.35}$$

which relates the risk premiums or Sharpe ratios on the different maturity bonds.

9.2.1 A Bond Risk Premium

Equation (9.35) says that bonds' expected rates of return in excess of the instantaneous maturity rate, divided by their standard deviations, must be equal at all points in time. This equality of Sharpe ratios must hold for any set of bonds τ_1, τ_2, τ_3, and so on. Each of the different bonds' reward-to-risk ratios (Sharpe ratios) derives from the single source of risk represented by the dz_r process driving the short-term interest rate, $r(t)$. Hence, condition (9.35) can be interpreted as a law of one price that requires all bonds to have a uniform *market price of interest rate risk*.

To derive the equilibrium prices for bonds, we must specify the form of this market price of bond risk. Chapter 13 outlines a general equilibrium model by John Cox, Jonathan Ingersoll, and Stephen Ross, (Cox, Ingersoll, and Ross 1985a) and (Cox, Ingersoll, and Ross 1985b), that shows how this bond risk premium can be derived from individuals' preferences (utilities) and the economy's technologies. For now, however, we simply assume that the market price of bond risk is constant over time and equal to q. Thus, we have for any bond maturity, τ,

$$\frac{\mu_p(r,\tau) - r(t)}{\sigma_p(\tau)} = q \tag{9.36}$$

or

$$\mu_p(r,\tau) = r(t) + q\sigma_p(\tau) \tag{9.37}$$

which says that the expected rate of return on a bond with maturity τ equals the instantaneous risk-free rate plus a risk premium proportional to the bond's standard deviation. Substituting $\mu_p(r,\tau)$ and $\sigma_p(\tau)$ from Itô's lemma into (9.37) and simplifying, we obtain

$$P_r\alpha(\overline{r} - r) + P_t + \tfrac{1}{2}P_{rr}\sigma_r^2 = rP - q\sigma_r P_r \tag{9.38}$$

This can be rewritten as

$$\tfrac{\sigma_r^2}{2} P_{rr} + (\alpha \overline{r} + q\sigma_r - \alpha r) P_r - rP + P_t = 0 \tag{9.39}$$

Equation (9.39) is the equilibrium partial differential equation that all bonds must satisfy. Since $\tau \equiv T - t$, so that $P_t \equiv \frac{\partial P}{\partial t} = -\frac{\partial P}{\partial \tau} \equiv -P_\tau$, equation (9.39) can be rewritten as

$$\tfrac{\sigma_r^2}{2} P_{rr} + [\alpha (\overline{r} - r) + q\sigma_r] P_r - rP - P_\tau = 0 \tag{9.40}$$

and, solved subject to the boundary condition that at $\tau = 0$, the bond price equals \$1; that is, $P(r, 0) = 1$. Doing so, gives the following solution:[14]

$$P(r(t), \tau) = A(\tau) e^{-B(\tau) r(t)} \tag{9.41}$$

where

$$B(\tau) \equiv \frac{1 - e^{-\alpha\tau}}{\alpha} \tag{9.42}$$

$$A(\tau) \equiv \exp\left[(B(\tau) - \tau) \left(\overline{r} + q\frac{\sigma_r}{\alpha} - \tfrac{1}{2}\frac{\sigma_r^2}{\alpha^2} \right) - \frac{\sigma_r^2 B(\tau)^2}{4\alpha} \right] \tag{9.43}$$

9.2.2 Characteristics of Bond Prices

Using equation (9.41), we see that

$$\sigma_p(\tau) = -\sigma_r \frac{P_r}{P} = \sigma_r B(\tau) = \frac{\sigma_r}{\alpha}\left(1 - e^{-\alpha\tau}\right) \tag{9.44}$$

[14] The solution can be derived by "guessing" a solution of the form in (9.41) and substituting it into (9.40). Noting that the terms multiplied by $r(t)$ and those terms not multiplied by $r(t)$ must each be zero for all $r(t)$ leads to simple ordinary differential equations for $A(\tau)$ and $B(\tau)$. These equations are solved subject to the boundary condition $P(r, \tau = 0) = 1$, which implies $A(\tau = 0) = 1$ and $B(\tau = 0) = 0$. See Chapter 17 for details and a generalization to bond prices that are influenced by multiple factors.

which implies that a bond's rate of return standard deviation (volatility) is an increasing but concave function of its maturity, τ. Moreover, (9.44) confirms that as the bond approaches its maturity date, its price volatility shrinks to zero, $\sigma_p(\tau = 0) = 0$, since the instantaneous maturity bond's return is riskless. The tendency for price volatility to decrease over time is a fundamental property of finitely lived, fixed-income securities that distinguishes them from potentially infinitely lived securities such as common or preferred stocks. While it may be reasonable to assume as in (9.14) that the volatility of a stock's price need not be a function of calendar time, this cannot be the case for a zero-coupon bond.

Given that $\sigma_p(\tau)$ is an increasing function maturity, equation (9.37) says that a bond's expected rate of return increases (*decreases*) with its time until maturity if the market price of risk, q, is positive (*negative*). Since historical returns on longer-maturity bonds have exceeded those of shorter-maturity ones in most (though not all) countries, this suggests that q is likely to be positive.[15] Additional evidence on the value of q can be gleaned by observing the yields to maturity on different maturity bonds. A τ-maturity bond's continuously compounded yield to maturity, denoted $Y(r(t), \tau)$, can be derived from its price in (9.41):

$$\begin{aligned} Y(r(t), \tau) &\equiv -\frac{1}{\tau}\ln\left[P(r(t), \tau)\right] \\ &= -\frac{1}{\tau}\ln\left[A(\tau)\right] + \frac{B(\tau)}{\tau}r(t) \\ &= Y_\infty + \left[r(t) - Y_\infty\right]\frac{B(\tau)}{\tau} + \frac{\sigma_r^2 B(\tau)^2}{4\alpha\tau} \end{aligned} \tag{9.45}$$

where $Y_\infty \equiv \overline{r} + q\frac{\sigma_r}{\alpha} - \frac{1}{2}\frac{\sigma_r^2}{\alpha^2}$. Note that $\lim_{\tau\to\infty} Y(r(t), \tau) = Y_\infty$, so that the yield to maturity on a very long maturity bond approaches Y_∞. Hence, the yield curve,

[15] See (Dimson, Marsh, and Staunton 2002) for an account of the historical evidence.

which is the graph of $Y(r(t), \tau)$ as a function of τ, equals $r(t)$ at $\tau = 0$ and asymptotes to Y_∞ for τ large. When $r(t) \leq Y_\infty - \frac{\sigma_r^2}{4\alpha^2} = \overline{r} + q\frac{\sigma_r}{\alpha} - \frac{3\sigma_r^2}{4\alpha^2}$, the yield curve is monotonically increasing. When $Y_\infty - \frac{\sigma_r^2}{4\alpha^2} < r(t) < Y_\infty + \frac{\sigma_r^2}{2\alpha^2} = \overline{r} + q\frac{\sigma_r}{\alpha}$, the yield curve has a humped shape. A monotonically downward sloping, or "inverted," yield curve occurs when $\overline{r} + q\frac{\sigma_r}{\alpha} \leq r(t)$. Since the unconditional mean of the short rate is $\overline{r}$ and, empirically, the yield curve is normally upward sloping, this suggests that $\overline{r} < \overline{r} + q\frac{\sigma_r}{\alpha} - \frac{3\sigma_r^2}{4\alpha^2}$, or $q > \frac{3\sigma_r}{4\alpha}$. Therefore, a yield curve that typically is upward sloping is also evidence of a positive market price of bond risk.

9.3 Option Pricing with Random Interest Rates

This last example of the Black-Scholes hedging argument combines aspects of the first two in that we now consider option pricing in an environment where interest rates can be random. We follow Robert Merton (Merton 1973b) in valuing a European call option when the risk-free interest rate is stochastic and bond prices satisfy the Vasicek model. The main alteration to the Black-Scholes derivation is to realize that the call option's payoff, $\max[S(T) - X, 0]$, depends not only on the maturity date, T, and the stock price at that date, $S(T)$, but on the present value of the exercise price, X, which can be interpreted as the value of a default-free bond that pays X at its maturity date of T. Given the randomness of interest rates, even the value of this exercise price is stochastic prior to the option's maturity. This motivates us to consider the process of a bond maturing in $\tau \equiv T - t$ periods to be another underlying asset, in addition to the stock, affecting the option's value. Writing this bond price as $P(t, \tau)$, the option's value can now be expressed as $c(S(t), P(t, \tau), t)$. Consistent with the Vasicek model, we write this bond's process as

$$dP(t,\tau) = \mu_p(t,\tau)\,P(t,\tau)\,dt + \sigma_p(\tau)\,P(t,\tau)\,dz_p \tag{9.46}$$

where from equation (9.31) we define $dz_p \equiv -dz_r$. In general, the bond's return will be correlated with that of the stock, and we allow for this possibility by assuming $dz_p dz = \rho dt$. Given the option's dependence on both the stock and the bond, Itô's lemma says that the option price satisfies

$$\begin{aligned} dc &= \left[\frac{\partial c}{\partial S}\,\mu S + \frac{\partial c}{\partial P}\,\mu_p P + \frac{\partial c}{\partial t} + \frac{1}{2}\frac{\partial^2 c}{\partial S^2}\,\sigma^2 S^2 + \frac{1}{2}\frac{\partial^2 c}{\partial P^2}\,\sigma_p^2 P^2 \right. \\ &\quad \left. + \frac{\partial^2 c}{\partial S \partial P}\,\rho\sigma\sigma_p SP\right] dt + \frac{\partial c}{\partial S}\,\sigma S\,dz + \frac{\partial c}{\partial P}\,\sigma_p P\,dz_p \qquad (9.47) \\ &\equiv \mu_c c dt + \frac{\partial c}{\partial S}\,\sigma S\,dz + \frac{\partial c}{\partial P}\,\sigma_p P\,dz_p \end{aligned}$$

where $\mu_c c$ is defined as those bracketed terms in the first two lines of (9.47). Similar to our first example in which a dealer wishes to hedge the sale of an option, let us form a hedge portfolio consisting of a unit short position in the option, and a purchase of $w_s(t)$ units of the underlying stock, and a purchase of $w_p(t)$ units of the τ-maturity bond, where we also restrict the portfolio to require a zero net investment. The zero-net-investment restriction implies

$$c(t) - w_s(t)\,S(t) - w_p(t)\,P(t,\tau) = 0 \tag{9.48}$$

The hedge portfolio's return can then be written as

$$\begin{aligned} dH(t) &= -dc(t) + w_s(t)\,dS(t) + w_p(t)\,dP(t,\tau) \qquad (9.49)\\ &= \left[-\mu_c c + w_s(t)\,\mu S + w_p(t)\,\mu_p P\right] dt \\ &\quad + \left[-\frac{\partial c}{\partial S}\,\sigma S + w_s(t)\,\sigma S\right] dz \\ &\quad + \left[-\frac{\partial c}{\partial P}\,\sigma_p P + w_p(t)\,\sigma_p P\right] dz_p \\ &= \left[w_s(t)\,(\mu - \mu_c)\,S + w_p(t)\left(\mu_p - \mu_c\right) P\right] dt \\ &\quad + \left[w_s(t) - \frac{\partial c}{\partial S}\right] \sigma S dz \\ &\quad + \left[w_p(t) - \frac{\partial c}{\partial P}\right] \sigma_p P\, dz_p \end{aligned}$$

where, in the last equality of (9.49), we have substituted in for c using the zero-net-investment condition (9.48). If $w_s(t)$ and $w_p(t)$ can be chosen to make the hedge portfolio's return riskless, then it must be the case that the terms in brackets in the last line of (9.49) can be made to equal zero. In other words, the following two conditions must hold:

$$w_s(t) = \frac{\partial c}{\partial S} \qquad (9.50)$$

$$w_p(t) = \frac{\partial c}{\partial P} \qquad (9.51)$$

but from the zero-net-investment condition (9.48), this can only be possible if it happens to be the case that

$$\begin{aligned} c &= w_s(t)\,S + w_p(t)\,P \\ &= S\frac{\partial c}{\partial S} + P\frac{\partial c}{\partial P} \qquad (9.52) \end{aligned}$$

By Euler's theorem, condition (9.52) holds if the option price is a homogeneous of degree 1 function of S and P.[16] What this means is that if the stock's price and the bond's price happened to increase by the same proportion, then the option's price would increase by that same proportion. That is, for $k > 0$, $c\left(kS\left(t\right), kP\left(t,\tau\right), t\right) = kc\left(S\left(t\right), P\left(t,\tau\right), t\right)$.[17] We assume this to be so and later verify that the solution indeed satisfies this homogeneity condition.

Given that condition (9.52) does hold, so that we can choose $w_s\left(t\right) = \partial c/\partial S$ and $w_p\left(t\right) = \partial c/\partial P$ to make the hedge portfolio's return riskless, then as in the first example the zero-net-investment portfolio's riskless return must equal zero in the absence of arbitrage:

$$w_s\left(t\right)\left(\mu - \mu_c\right)S + w_p\left(t\right)\left(\mu_p - \mu_c\right)P = 0 \tag{9.53}$$

or

$$\frac{\partial c}{\partial S}\left(\mu - \mu_c\right)S + \frac{\partial c}{\partial P}\left(\mu_p - \mu_c\right)P = 0 \tag{9.54}$$

which, using (9.52), can be rewritten as

$$\frac{\partial c}{\partial S}\mu S + \frac{\partial c}{\partial P}\mu_p P - \mu_c c = 0 \tag{9.55}$$

Substituting for $\mu_c c$ from (9.47), we obtain

$$-\frac{\partial c}{\partial t} - \frac{1}{2}\frac{\partial^2 c}{\partial S^2}\sigma^2 S^2 - \frac{1}{2}\frac{\partial^2 c}{\partial P^2}\sigma_p^2 P^2 - \frac{\partial^2 c}{\partial S \partial P}\rho\sigma\sigma_p SP = 0 \tag{9.56}$$

which, since $\tau \equiv T - t$, can also be written as

$$\frac{1}{2}\left[\frac{\partial^2 c}{\partial S^2}\sigma^2 S^2 + \frac{\partial^2 c}{\partial P^2}\sigma_p^2 P^2 + 2\frac{\partial^2 c}{\partial S \partial P}\rho\sigma\sigma_p SP\right] - \frac{\partial c}{\partial \tau} = 0 \tag{9.57}$$

[16] A function $f\left(x_1, ..., x_n\right)$ is defined to be homogeneous of degree r (where r is an integer) if for every $k > 0$, then $f\left(kx_1, ..., kx_n\right) = k^r f\left(x_1, ..., x_n\right)$. Euler's theorem states that if $f\left(x_1, ..., x_n\right)$ is homogeneous of degree r and differentiable, then $\sum_{i=1}^n x_i \frac{\partial f}{\partial x_i} = rf$.

[17] For example, suppose there was a general rise in inflation that increased the stock's and bond's prices but did not change their *relative* price, S/P. Then the homogeneity condition implies that the option's price would rise by the same increase in inflation.

Equation (9.57) is the equilibrium partial differential equation that the option's value must satisfy. Importantly, it does not depend on either the expected rate of return on the stock, μ, or the expected rate of return on the bond, μ_p. The appropriate boundary condition for a European call option is similar to before, with $c(S(T), P(T,0), T) = c(S(T), 1, T) = \max[S(T) - X, 0]$, where we impose the condition $P(t = T, \tau = 0) = 1$. Robert Merton (Merton 1973b) shows that the solution to this equation is

$$c(S(t), P(t,\tau), \tau) = S(t)\, N(h_1) - P(t,\tau)\, XN(h_2) \tag{9.58}$$

where

$$\begin{aligned} h_1 &= \frac{\ln\left(\frac{S(t)}{P(t,\tau)X}\right) + \frac{1}{2}v^2}{v} \\ h_2 &= h_1 - v \end{aligned} \tag{9.59}$$

where

$$v^2 = \int_0^{\tau} \left(\sigma^2 + \sigma_p(y)^2 - 2\rho\sigma\sigma_p(y)\right) dy \tag{9.60}$$

The solution is essentially the same as the Black-Scholes constant interest rate formula (9.24) but where the parameter v^2 replaces $\sigma^2\tau$. v^2 is the total variance of the ratio of the stock price to the discounted exercise price over the life of the option.[18] In other words, it is the variance of the ratio $\frac{S(t)}{P(t,\tau)X}$ from date t to date T, an interval of τ periods. Because the instantaneous variance of the bond, and hence the variance of the discounted exercise price, shrinks as the option approaches maturity, this changing variance is accounted for by making $\sigma_p(y)$ a function of the time until maturity in (9.60). If we assume that the

[18] As one would expect, when interest rates are nonstochastic so that the volatility of bond prices is zero, that is, $\sigma_p(y) = 0$, then $v^2 = \sigma^2\tau$, and we obtain the standard Black-Scholes formula.

bond's volatility is that of the Vasicek model, $\sigma_p(y) = \frac{\sigma_r}{\alpha}\left(1 - e^{-\alpha y}\right)$, then (9.60) becomes

$$\begin{aligned} v^2 &= \int_0^\tau \left(\sigma^2 + \frac{\sigma_r^2}{\alpha^2}\left(1 - 2e^{-\alpha y} + e^{-2\alpha y}\right) - 2\rho\sigma\frac{\sigma_r}{\alpha}\left(1 - e^{-\alpha y}\right)\right) dt \qquad (9.61) \\ &= \sigma^2\tau + \frac{\sigma_r^2}{\alpha^3}\left(\alpha\tau + \frac{1 - e^{-2\alpha\tau}}{2} - 2\left(1 - e^{-\alpha\tau}\right)\right) - 2\rho\sigma\frac{\sigma_r}{\alpha^2}\left[\alpha\tau - \left(1 - e^{-\alpha\tau}\right)\right] \end{aligned}$$

Finally, note that the solution is homogeneous of degree 1 in $S(t)$ and $P(t,\tau)$, which verifies condition (9.52).

9.4 Summary

Fischer Black, Myron Scholes, and Robert Merton made a fundamental discovery that profoundly changed the pricing of contingent securities. They showed that when an underlying asset follows a diffusion, and trade is allowed to occur continuously, a portfolio can be created that fully hedges the risk of the contingent claim. Therefore, in the absence of arbitrage, the hedge portfolio's return must be riskless, and this implies that the contingent claim's price must satisfy a particular partial differential equation subject to a boundary condition that its value must equal its terminal payoff. Solving this equation led to a surprising result: the contingent claim's value did not depend directly on the underlying security's expected rate of return, but only on its volatility. This was an attractive feature because estimating a risky asset's expected rate of return is much more difficult than estimating its volatility.[19]

As our second example illustrated, the Black-Scholes-Merton hedging argument can be used to derive models of the default-free term structure of interest

[19] The accuracy of estimates for a risky asset's expected rate of return is proportional to the time interval over which its average return is computed. In contrast, the accuracy of a risky asset's standard deviation of return is proportional to the number of times the return is sampled over any fixed time interval. See Merton (Merton 1980) and Chapter 9.3.2 of Campbell, Lo, and MacKinlay (Campbell, Lo, and MacKinlay 1997).

rates. The pricing of different maturity bonds and of fixed-income derivatives is a large and ever-growing field of asset pricing. Chapter 17 is devoted solely to this subject. A related topic is the pricing of default-risky bonds. As the title of Black and Scholes's seminal paper suggests, it was readily recognized that a satisfactory model of option pricing could be applied to valuing the liabilities of corporations that were subject to default. This link between option pricing and credit risk also will be explored in Chapter 18.

9.5 Exercises

1. Suppose that the price of a non-dividend-paying stock follows the process

$$dS = \alpha S dt + \beta S^{\gamma} dz$$

where α, β, and γ are constants. The risk-free interest rate equals a constant, r. Denote $p(S(t), t)$ as the current price of a European put option on this stock having an exercise price of X and a maturity date of T. Derive the equilibrium partial differential equation and boundary condition for the price of this put option using the Black-Scholes hedging argument.

2. Define $P(r(t), \tau)$ as the date t price of a pure discount bond that pays \$1 in τ periods. The bond price depends on the instantaneous maturity yield, $r(t)$, which follows the process

$$dr(t) = \alpha [\overline{r} - r(t)] dt + \sigma \sqrt{r} dz$$

where α, $\overline{r}$, and σ are positive constants. If the process followed by the price of a bond having τ periods until maturity is

$$dP\left(r,\tau\right)/P\left(r,\tau\right)=\mu\left(r,\tau\right)dt-\sigma_{p}\left(r,\tau\right)dz$$

and the market price of bond risk is

$$\frac{\mu\left(r,\tau\right)-r\left(t\right)}{\sigma_{p}\left(r,\tau\right)}=\lambda\sqrt{r}$$

then write down the equilibrium partial differential equation and boundary condition that this bond price satisfies.

3. The date t price of stock A, $A\left(t\right)$, follows the process

$$dA/A=\mu_{A}dt+\sigma_{A}dz$$

and the date t price of stock B, $B\left(t\right)$, follows the process

$$dB/B=\mu_{B}dt+\sigma_{B}dq$$

where σ_A and σ_B are constants and dz and dq are Brownian motion processes for which $dzdq=\rho dt$. Let $c\left(t\right)$ be the date t price of a European option written on the difference between these two stocks' prices. Specifically, at this option's maturity date, T, the value of the option equals

$$c\left(T\right)\;=\;\max\left[0,A\left(T\right)-B\left(T\right)\right]$$

a. Using Itô's lemma, derive the process followed by this option.

b. Suppose that you are an option dealer who has just sold (written) one of these options for a customer. You now wish to form a hedge portfolio composed of your unit short position in the option and positions in the two stocks. Let $H(t)$ denote the date t value of this hedge portfolio. Write down an equation for $H(t)$ that indicates the amount of shares of stocks A and B that should be held.

c. Write down the dynamics for $dH(t)$, showing that its return is riskless.

d. Assuming the absence of arbitrage, derive the equilibrium partial differential equation that this option must satisfy.

4. Let $S(t)$ be the date t price of an asset that continuously pays a dividend that is a fixed proportion of its price. Specifically, the asset pays a dividend of $\delta S(t)\, dt$ over the time interval dt. The process followed by this asset's price can be written as

$$dS = (\mu - \delta)\, Sdt + \sigma S dz$$

where σ is the standard deviation of the asset's rate of return and μ is the asset's total expected rate of return, which includes its dividend payment and price appreciation. Note that the total rate of return earned by the owner of one share of this asset is $dS/S + \delta dt = \mu dt + \sigma dz$. Consider a European call option written on this asset that has an exercise price of X and a maturity date of $T > t$. Assuming a constant interest rate equal to r, use a Black-Scholes hedging argument to derive the equilibrium partial differential equation that this option's price, $c(t)$, must satisfy.

Chapter 10

Arbitrage, Martingales, and Pricing Kernels

In Chapters 4 and 7, we examined the asset pricing implications of market completeness in a discrete-time model. It was shown that when the number of nonredundant assets equaled the number of states of nature, markets were complete and the absence of arbitrage ensured that state prices and a state price deflator would exist. Pricing could be performed using risk-neutral valuation. The current chapter extends these results in a continuous-time environment. We formally show that when asset prices follow diffusion processes and trading is continuous, then the absence of arbitrage may allow us to value assets using a *martingale pricing* technique, a generalization of risk-neutral pricing. Under these conditions, a continuous-time stochastic discount factor, or pricing kernel, also exists.

These results were developed by John Cox and Stephen Ross (Cox and Ross 1976), John Harrison and David Kreps (Harrison and Kreps 1979), and John Harrison and Stanley Pliska (Harrison and Pliska 1981) and have proved to be very popular approaches to valuing a wide variety of contingent claims. Valuing contingent claims using risk-neutral pricing, or a pricing kernel method, can be an alternative to the previous chapter's partial differential equation approach.

The first section of this chapter reviews the derivation of the Black-Scholes partial differential equation and points out that this equation also implies that the market price of risk must be uniform for a contingent claim and its underlying asset. It also shows how the contingent claim's price process can be transformed into a driftless process by adjusting its Brownian motion process by the market price of risk and then deflating the contingent claim's price by that of a riskless asset. This driftless (zero expected change) process is known as a *martingale*. The contingent claim's value then can be computed as the expectation of its terminal value under this transformed process.

The second section derives the form of a continuous-time state price deflator that can also be used to price contingent claims. It also demonstrates how the continuous-time state price deflator transforms actual probabilities into risk-neutral probabilities. The third section shows how problems of valuing a contingent claim sometimes can be simplified by deflating the contingent claim's price by that of another risky asset. An example is given by valuing an option written on the difference between the prices of two risky assets. The final section of the chapter examines applications of the martingale approach. It is used to value an option written on an asset that pays a continuous dividend, examples of which include an option written on a foreign currency and an option written on a futures price. The martingale pricing technique is also applied to rederiving a model of the term structure of interest rates.

10.1 Arbitrage and Martingales

We begin by reviewing the Black-Scholes derivation of contingent claims prices. Let S be the value of a risky asset that follows a general scalar diffusion process

$$dS = \mu S dt + \sigma S dz \tag{10.1}$$

where both $\mu = \mu(S,t)$ and $\sigma = \sigma(S,t)$ may be functions of S and t and dz is a standard, pure Brownian motion (or Wiener) process. For ease of presentation, we assume that $S(t)$ is a scalar process. Later we discuss how multivariate processes can be handled by the theory, such that μ and σ can depend on other variables that follow diffusion processes (driven by additional Brownian motions) in addition to $S(t)$. In this way, asset values can depend on multiple sources of uncertainty.

Next let $c(S,t)$ denote the value of a contingent claim whose payoff depends solely on S and t. From Itô's lemma, we know that this value satisfies

$$dc = \mu_c c dt + \sigma_c c dz \tag{10.2}$$

where $\mu_c c = c_t + \mu S c_S + \frac{1}{2}\sigma^2 S^2 c_{SS}$ and $\sigma_c c = \sigma S c_S$, and the subscripts on c denote partial derivatives.

Similar to our earlier analysis, we employ a form of the Black-Scholes hedging argument by considering a portfolio of -1 units of the contingent claim and c_S units of the risky asset. The value of this portfolio, H, satisfies[1]

$$H = -c + c_S S \tag{10.3}$$

and the change in value of this portfolio over the next instant is

[1] Unlike last chapter's derivation, we do not restrict this portfolio to be a zero-net-investment portfolio. As will be clear, the lack of this restriction does not change the nature of our results.

$$
\begin{aligned}
dH &= -dc + c_S dS \\
&= -\mu_c c dt - \sigma_c c dz + c_S \mu S dt + c_S \sigma S dz \\
&= \left[c_S \mu S - \mu_c c\right] dt
\end{aligned}
\tag{10.4}
$$

Since the portfolio is riskless, the absence of arbitrage implies that it must earn the risk-free rate. Denoting the (possibly stochastic) instantaneous risk-free rate as $r(t)$, we have[2]

$$
dH = \left[c_S \mu S - \mu_c c\right] dt = rHdt = r[-c + c_S S]dt \tag{10.5}
$$

which implies

$$
c_S \mu S - \mu_c c = r[-c + c_S S] \tag{10.6}
$$

If we substitute $\mu_c c = c_t + \mu S c_S + \frac{1}{2}\sigma^2 S^2 c_{SS}$ into (10.6), we obtain the Black-Scholes equilibrium partial differential equation (PDE):

$$
\frac{1}{2}\sigma^2 S^2 c_{SS} + rSc_S - rc + c_t = 0 \tag{10.7}
$$

However, consider a different interpretation of equation (10.6). From Itô's lemma, we can substitute $c_S = \frac{\sigma_c c}{\sigma S}$ into (10.6) and rearrange to obtain

[2] For simplicity, we have assumed that the contingent claim's value depends only on a single risky asset price, $S(t)$. However, when the interest rate is stochastic, the contingent claim's value also might be a function of $r(t)$, that is, $c(S, r, t)$. If, for example, the interest rate followed the process $dr = \mu_r(r)\, dt + \sigma_r(r)\, dz_r$ where dz_r is an additional Wiener process affecting interest rate movements, then the contingent claim's process would be given by a bivariate version of Itô's lemma. Also, to create a portfolio that earns an instantaneous risk-free rate, the portfolio would need to include a bond whose price is driven by dz_r. Later, we discuss how our results generalize to multiple sources of uncertainty. However, the current univariate setting can be fully consistent with stochastic interest rates if the risky asset is, itself, a bond so that $S(r, t)$ and $dz = dz_r$. The contingent claim could then be interpreted as a fixed-income (bond) derivative security.

$$\frac{\mu - r}{\sigma} = \frac{\mu_c - r}{\sigma_c} \equiv \theta(t) \tag{10.8}$$

Condition (10.8) is the familiar no-arbitrage condition that requires a unique market price of risk, which we denote as $\theta(t)$. Then the stochastic process for the contingent claim can be written as

$$dc = \mu_c c dt + \sigma_c c dz = [rc + \theta\sigma_c c] dt + \sigma_c c dz \tag{10.9}$$

Note that the drift of this process depends on the market price of risk, $\theta(t)$, which may not be directly observable or easily estimated. We now consider an approach to valuing contingent claims that is an alternative to solving the PDE in (10.7) but that shares with it the benefit of not having to know $\theta(t)$. The next topic discusses how a contingent claim's risk premium can be eliminated by reinterpreting the probability distribution generating asset returns.

10.1.1 A Change in Probability: Girsanov's Theorem

Girsanov's theorem says that by shifting the Brownian motion process, one can change the drift of a diffusion process when this process is interpreted under a new probability distribution. Moreover, this shift in Brownian motion changes the future probability distribution for asset prices in a particular way. To see how this works, consider a new process $\widehat{z}_t = z_t + \int_0^t \theta(s) ds$, so that $d\widehat{z}_t = dz_t + \theta(t) dt$. Then substituting $dz_t = d\widehat{z}_t - \theta(t) dt$ in equation (10.9), it can be rewritten:

$$\begin{aligned} dc &= [rc + \theta\sigma_c c] dt + \sigma_c c [d\widehat{z} - \theta dt] \\ &= rc dt + \sigma_c c d\widehat{z} \end{aligned} \tag{10.10}$$

Hence, converting from the Brownian motion process dz to $d\widehat{z}$, which removes the risk premium $\theta\sigma_c c$ from the drift term on the right-hand side of (10.9), results in the expected rate of return of c being equal to the risk-free rate *if* we were now to view $d\widehat{z}$, rather than dz, as a Brownian motion process. The probability distribution of future values of c that are generated by $d\widehat{z}$, a probability distribution that we define as the Q probability measure, is referred to as the risk-neutral probability measure.[3] This is in contrast to the actual probability distribution for c generated by the dz Brownian motion in (10.9), the original "physical," or "statistical," probability distribution that is denoted as the P measure.

Girsanov's theorem states that as long as $\theta(t)$ is well behaved in the sense that it follows a process that does not vary too much over time, then the probability density function for a random variable at some future date T, such as $c(T)$, under the risk-neutral Q distribution bears a particular relationship to that of the physical P distribution.[4] Specifically, denote dP_T as the instantaneous change in the physical distribution function at date T generated by dz_t, which makes it the physical probability density function at date T.[5] Similarly, let dQ_T be the risk-neutral probability density function generated by $d\widehat{z}_t$. Then

---

[3] The idea of a probability measure (or distribution), P, is as follows. Define a set function, f, which assigns a real number to a set E, where E could be a set of real numbers, such as an interval on the real line. Formally, $f(E) \in R$. This function is *countably additive* if $f\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} f(E_i)$ where $\langle E_i \rangle$ is a finite or countably infinite sequence of disjoint sets. A *measure* is defined as a nonnegative set function that is countably additive. Note that probabilities are measures since they assign a nonnegative probability to a particular set. For example, let the domain of a continuous probability distribution for a random variable, x, be the entire real line; that is, $\int_{-\infty}^{\infty} dP(x) = 1$ where P is the probability measure (probability distribution function). Now let a set $E_1 = [a,b]$ be an interval on this line. The probability of $x \in E_1$ is $f(E_1) = \int_a^b dP(x) \geq 0$. Similarly, if $E_2 = [c,d]$, which is assumed to be an interval that does not overlap with E_1, then $f(E_1 \bigcup E_2) = \int_a^b dP(x) + \int_c^d dP(x) = f(E_1) + f(E_2)$. Hence, probabilities are nonnegative and countably additive.

[4] The restriction on $\theta(t)$ is that $E_t\left[\exp\left(\int_t^T \theta(u)^2 \, du\right)\right] < \infty$, which is known as the Novikov condition. Ioannis Karatzas and Steven Shreve (Karatzas and Shreve 1991) give a formal statement and proof of Girsanov's theorem.

[5] Recall that since a probability distribution function, P, is an integral over the probability density function, $\int dP$, the density function can be interpreted as the derivative of the probability distribution function.

Girsanov's theorem says that at some date t where $0 < t < T$, the relationship between the two probability densities at date T is

$$\begin{aligned} dQ_T &= \exp\left[-\int_t^T \theta(u)\, dz - \frac{1}{2}\int_t^T \theta(u)^2\, du\right] dP_T \\ &= (\xi_T/\xi_t)\, dP_T \end{aligned} \tag{10.11}$$

where ξ_t is a positive random process that depends on $\theta(t)$ and z_t and is given by

$$\xi_\tau = \exp\left[-\int_0^\tau \theta(u)\, dz - \frac{1}{2}\int_0^\tau \theta(u)^2\, du\right] \tag{10.12}$$

In other words, by multiplying the physical probability density at date T by the factor ξ_T/ξ_t, we can determine the risk-neutral probability density at date T. Since from (10.12) we see that $\xi_T/\xi_t > 0$, equation (10.11) implies that whenever dP_T has positive probability, so does dQ_T. Because they share this characteristic, the physical P measure and the risk-neutral Q measure are called *equivalent* probability measures in that any future value of c that has positive probability (density) under the physical measure also has positive probability (density) under the risk-neutral measure.[6] We can rearrange (10.11) to obtain

$$\frac{dQ_T}{dP_T} = \xi_T/\xi_t \tag{10.13}$$

which clarifies that ξ_T/ξ_t can be interpreted as the derivative of the risk-neutral measure Q_T with respect to the physical measure P_T. Indeed, ξ_T/ξ_t is known as the Radon-Nikodym derivative of Q with respect to P. Later in this chapter we will return to an interpretation of this derivative ξ_T/ξ_t following a discussion

[6] An example illustrates this equivalency. Suppose in (10.1) that μ and σ are constant and the risk-free interest rate, r, is constant. Then the process $dS/S = \mu dt + \sigma dz$ has a discrete time lognormal distribution under the P measure. Under the Q measure the process is $dS/S = rdt + \sigma d\widehat{z}$, which is also lognormally distributed but with r replacing μ. Since these lognormal distributions both have positive probability density over the domain from 0 to ∞, they are referred to as *equivalent*.

of the continuous-time pricing kernel approach to valuing contingent securities.

In summary, we have seen that a transformation of a Brownian motion by the market price of risk transforms a security's expected rate of return to equal the risk-free rate. This transformation from the physical Brownian motion to a risk-neutral one also transforms the probability density functions for random variables at future dates.

10.1.2 Money Market Deflator

As a final step in deriving a new valuation formula for contingent claims, we now show that the contingent claim's appropriately deflated price process can be made driftless (a martingale) under the probability measure Q. Let $B(t)$ be the value of an investment in a "money market fund," that is, an investment in the instantaneous maturity risk-free asset.[7] Then

$$dB/B = r(t)dt \tag{10.14}$$

Note that $B(T) = B(t)\, e^{\int_t^T r(u)du}$ for any date $T \geq t$. Now define $C(t) \equiv c(t)/B(t)$ as the deflated price process for the contingent claim. Essentially, $C(t)$ is the value of the contingent claim measured in terms of the value of the riskless safe investment that grows at rate $r(t)$. A trivial application of Itô's lemma gives

$$\begin{aligned} dC &= \frac{1}{B}dc - \frac{c}{B^2}dB \\ &= \frac{rc}{B}dt + \frac{\sigma_c c}{B}d\widehat{z} - r\frac{c}{B}dt \\ &= \sigma_c C d\widehat{z} \end{aligned} \tag{10.15}$$

[7] An investment that earns the instantaneous maturity risk-free rate is sometimes referred to as a money market fund because money market mutual funds invest in short-maturity, high-credit quality (nearly risk-free) debt instruments.

Thus, the deflated price process under the equivalent probability measure generated by $d\widehat{z}$ is a driftless process: its expected change is zero. An implication of (10.15) is that the expectation under the risk-neutral, or Q, measure of any future value of C is the current value of C. This can be stated as

$$C(t) = \widehat{E}_t [C(T)] \qquad \forall T \geq t \tag{10.16}$$

where $\widehat{E}_t [\cdot]$ denotes the expectation operator under the probability measure generated by $d\widehat{z}$.[8] The mathematical name for a process such as (10.16) is a martingale, which is essentially a random walk in discrete time.[9]

To summarize, we showed that the absence of arbitrage implies the existence of an equivalent probability measure such that the deflated price process is a martingale. Note that (10.16) holds for any deflated contingent claim, including the deflated underlying risky asset, S/B, since we could define the contingent claim as $c = S$.

10.1.3 Feynman-Kac Solution

Now if we rewrite (10.16) in terms of the undeflated contingent claims price, we obtain

$$\begin{aligned} c(t) &= B(t)\widehat{E}_t \left[c(T) \frac{1}{B(T)} \right] \\ &= \widehat{E}_t \left[e^{-\int_t^T r(u)du} c(T) \right] \end{aligned} \tag{10.17}$$

---

[8] Another common notation for this risk-neutral, or Q, measure expectation is $E_t^Q [\cdot]$.

[9] More formally, define a family of information sets, I_t, that start at date $t = 0$ and continue for all future dates, $\{I_t, t \in [0, \infty]\}$. Also, assume that information at date t includes all information from previous dates, so that for $t_0 < t_1 < t_2$, $I_{t_0} \subseteq I_{t_1} \subseteq I_{t_2}$. Such a family of information sets is referred to as a *filtration*. A process is a martingale with respect to I_t if it satisfies $E[C(T)|I_t] = C(t)$ $\forall t < T$ where I_t includes the value of $C(t)$, and $E[|C(T)|] < \infty$; that is, the unconditional expectation of the process is finite.

Equation (10.17) can be interpreted as a solution to the Black-Scholes partial differential equation (10.7) and, indeed, is referred to as the Feynman-Kac solution.[10] From a computational point of view, equation (10.17) says that we can price (value) a contingent security by taking the expected value of its discounted payoff, where we discount at the risk-free rate but also assume that when taking the expectation of $c(T)$ the rate of return on c (and all other asset prices, such as S) equals the risk-free rate, a rate that may be changing over time. As when the contingent security's value is found directly from the partial differential equation (10.7), no assumption regarding the market price of risk, $\theta(t)$, is required, because it was eliminated from all assets' return processes when converting to the Q measure. Equivalently, one can use equation (10.16) to value $c(t)/B(t)$ by taking expectations of the deflated price process, where this deflated process has zero drift. Both of these procedures are continuous-time extensions of the discrete-time, risk-neutral valuation technique that we examined in Chapters 4 and 7.

10.2 Arbitrage and Pricing Kernels

This is not the first time that we have computed an expectation to value a security. Recall from the single- or multiperiod consumption-portfolio choice problem with time-separable utility that we obtained an Euler condition of the form[11]

[10] To solve (10.7), a boundary condition for the derivative is needed. For example, in the case of a European call option, it would be $c(T) = \max[0, S(T) - X]$. The solution given by (10.17) incorporates this boundary condition, $c(T)$.

[11] In equation (10.18) we are assuming that the contingent claim pays no dividends between dates t and T.

$$\begin{aligned} c(t) &= E_t\left[m_{t,T}c(T)\right] \\ &= E_t\left[\frac{M_T}{M_t}c(T)\right] \end{aligned} \tag{10.18}$$

where date $T \geq t$, $m_{t,T} \equiv M_T/M_t$ and $M_t = U_c(C_t, t)$ was the marginal utility of consumption at date t. In Chapter 4, we also showed in a discrete time-discrete state model that the absence of arbitrage implies that a stochastic discount factor, $m_{t,T}$, exists whenever markets are complete. We now show that this same result applies in a continuous-time environment whenever markets are dynamically complete. The absence of arbitrage opportunities, which earlier guaranteed the existence of an equivalent martingale measure, also determines a pricing kernel, or state price deflator, M_t. In fact, the concepts of an equivalent martingale measure and state pricing kernel are one and the same.

Note that we can rewrite (10.18) as

$$c(t)\,M_t = E_t\left[c(T)\,M_T\right] \tag{10.19}$$

which says that the deflated price process, $c(t)\,M_t$, is a martingale. But note the difference here versus our earlier analysis: the expectation in (10.19) is taken under the physical probability measure, P, while in (10.16) and (10.17) the expectation is taken under the risk-neutral measure, Q.

Since in the standard, time-separable utility portfolio choice model M_t is the marginal utility of consumption, this suggests that M_t should be a positive process even when we consider more general environments where a stochastic discount factor pricing relationship would hold. Hence, we assume that the state price deflator, M_t, follows a strictly positive diffusion process of the general form

$$dM = \mu_m dt + \sigma_m dz \tag{10.20}$$

Now consider the restrictions that the Black-Scholes no-arbitrage conditions place on μ_m and σ_m if (10.19) and (10.20) hold. For any arbitrary security or contingent claim, c, define $c^m = cM$ and apply Itô's lemma:

$$\begin{aligned} dc^m &= cdM + Mdc + (dc)(dM) \\ &= [c\mu_m + M\mu_c c + \sigma_c c\sigma_m] dt + [c\sigma_m + M\sigma_c c] dz \end{aligned} \tag{10.21}$$

If $c^m = cM$ satisfies (10.19), that is, c^m is a martingale, then its drift in (10.21) must be zero, implying

$$\mu_c = -\frac{\mu_m}{M} - \frac{\sigma_c \sigma_m}{M} \tag{10.22}$$

Now consider the case in which c is the instantaneously riskless asset; that is, $c(t) = B(t)$ is the money market investment following the process in equation (10.14). This implies that $\sigma_c = 0$ and $\mu_c = r(t)$. Using (10.22) requires

$$r(t) = -\frac{\mu_m}{M} \tag{10.23}$$

In other words, the expected rate of change of the pricing kernel must equal minus the instantaneous risk-free interest rate.

Next, consider the general case where the asset c is risky, so that $\sigma_c \neq 0$. Using (10.22) and (10.23) together, we obtain

$$\mu_c = r(t) - \frac{\sigma_c \sigma_m}{M} \tag{10.24}$$

or

$$\frac{\mu_c - r}{\sigma_c} = -\frac{\sigma_m}{M} \tag{10.25}$$

Comparing (10.25) to (10.8), we see that

$$-\frac{\sigma_m}{M} = \theta(t) \tag{10.26}$$

Thus, the no-arbitrage condition implies that the form of the pricing kernel must be

$$dM/M = -r(t)\,dt - \theta(t)\,dz \tag{10.27}$$

Note that if we define $\mathrm{m}_t \equiv \ln M_t$, then $d\mathrm{m}= -\left[r + \frac{1}{2}\theta^2\right] dt - \theta dz$. Hence, in using the pricing kernel to value any contingent claim, we can rewrite (10.18) as

$$\begin{aligned} c(t) &= E_t\left[c(T)\, M_T/M_t\right] = E_t\left[c(T)\, e^{\mathrm{m}_T - \mathrm{m}_t}\right] \\ &= E_t\left[c(T)\, e^{-\int_t^T \left[r(u)+\frac{1}{2}\theta^2(u)\right]du - \int_t^T \theta(u)dz}\right] \end{aligned} \tag{10.28}$$

Given processes for $r(t)$, $\theta(t)$, and the contingent claim's payoff, $c(T)$, in some instances it may be easier to compute (10.28) rather than, say, (10.16) or (10.17). Of course, in computing (10.28), we need to use the actual drift for c; that is, we compute expectations under the P measure, not the Q measure.

10.2.1 Linking the Valuation Methods

To better understand the connection between the pricing kernel (stochastic discount factor) approach and the martingale (risk-neutral) valuation approach, we now show how M_t is related to the change in probability distribution accomplished using Girsanov's theorem. Equating (10.17) to (10.28), we have

$$\begin{aligned}\widehat{E}_t\left[e^{-\int_t^T r(u)du}c(T)\right] &= E_t\left[c(T)\,M_T/M_t\right] \\ &= E_t\left[e^{-\int_t^T r(u)du}c(T)\,e^{-\int_t^T \frac{1}{2}\theta^2(u)du-\int_t^T \theta(u)dz}\right]\end{aligned} \tag{10.29}$$

and then if we substitute using the definition of ξ_τ from (10.12), we have

$$\begin{aligned}\widehat{E}_t\left[e^{-\int_t^T r(u)du}c(T)\right] &= E_t\left[e^{-\int_t^T r(u)du}c(T)\left(\xi_T/\xi_t\right)\right] \\ \widehat{E}_t\left[C(T)\right] &= E_t\left[C(T)\left(\xi_T/\xi_t\right)\right] \\ \int C(T)\,dQ_T &= \int C(T)\left(\xi_T/\xi_t\right)dP_T\end{aligned} \tag{10.30}$$

where, you may recall, $C(t) = c(t)/B(t)$. From the first two lines of (10.30), we see that on both sides of the equation, the terms in brackets are exactly the same except that the expectation under P includes the Radon-Nikodym derivative ξ_T/ξ_t. As predicted by Girsanov's theorem, this factor transforms the physical probability density at date T to the risk-neutral probability density at date T. Furthermore, relating (10.29) to (10.30) implies

$$M_T/M_t = e^{-\int_t^T r(u)du}\left(\xi_T/\xi_t\right) \tag{10.31}$$

so that the continuous-time pricing kernel (stochastic discount factor) is the product of a risk-free rate discount factor and the Radon-Nikodym derivative. Hence, M_T/M_t can be interpreted as providing both discounting at the risk-free rate and transforming the probability distribution to the risk-neutral one. Indeed, if contingent security prices are deflated by the money market investment, thereby removing the risk-free discount factor, the second line of (10.30) shows that the pricing kernel, M_T/M_t, and the Radon-Nikodym derivative, ξ_T/ξ_t, are exactly the same.

Similar to the discrete-time case discussed in Chapter 4, the role of this derivative (ξ_T/ξ_t or M_T/M_t) is to adjust the risk-neutral probability, Q, to give it greater probability density for "bad" outcomes and less probability density for "good" outcomes relative to the physical probability, P. In continuous time, the extent to which an outcome, as reflected by a realization of dz, is bad or good depends on the sign and magnitude of its market price of risk, $\theta(t)$. This explains why in equation (10.27) the stochastic component of the pricing kernel is of the form $-\theta(t)\,dz$.

10.2.2 The Multivariate Case

The previous analysis has assumed that contingent claims prices depend on only a single source of uncertainty, dz. In a straightforward manner, the results can be generalized to permit multiple independent sources of risk. Suppose we had asset returns depending on an $n \times 1$ vector of independent Brownian motion processes, $\mathbf{dZ} = (dz_1 ... dz_n)'$ where $dz_i dz_j = 0$ for $i \neq j$.[12] A contingent claim whose payoff depended on these asset returns then would have a price that followed the process

$$dc/c = \mu_c dt + \mathbf{\Sigma}_c \mathbf{dZ} \tag{10.32}$$

where $\mathbf{\Sigma}_c$ is a $1 \times n$ vector $\mathbf{\Sigma}_c = (\sigma_{c1} ... \sigma_{cn})$.[13] Let the corresponding $n \times 1$ vector of market prices of risks associated with each of the Brownian motions be $\mathbf{\Theta} = (\theta_1 ... \theta_n)'$. Then, it is straightforward to show that we would have the no-arbitrage condition

$$\mu_c - r = \mathbf{\Sigma}_c \mathbf{\Theta} \tag{10.33}$$

[12] The independence assumption is not important. If there are correlated sources of risk (Brownian motions), they can be redefined by a linear transformation to be represented by n orthogonal risk sources.

[13] Both μ_c and the elements of $\mathbf{\Sigma}_c$ may be functions of state variables driven by the Brownian motion components of $\mathbf{dZ}$.

Equations (10.16) and (10.17) would still hold, and now the pricing kernel's process would be given by

$$dM/M = -r(t)\,dt - \mathbf{\Theta}(t)'\,\mathbf{dZ} \tag{10.34}$$

10.3 Alternative Price Deflators

In previous sections, we found it convenient to deflate a contingent claim price by the money market fund's price, $B(t)$. Sometimes, however, it may be convenient to deflate or "normalize" a contingent claims price by the price of a different type of security. Such a situation can occur when a contingent claim's payoff depends on multiple risky assets. Let's now consider an example of this, in particular, where the contingent claim is an option written on the difference between two securities' (stocks') prices. The date t price of stock 1, $S_1(t)$, follows the process

$$dS_1/S_1 = \mu_1 dt + \sigma_1 dz_1 \tag{10.35}$$

and the date t price of stock 2, $S_2(t)$, follows the process

$$dS_2/S_2 = \mu_2 dt + \sigma_2 dz_2 \tag{10.36}$$

where σ_1 and σ_2 are assumed to be constants and dz_1 and dz_2 are Brownian motion processes for which $dz_1 dz_2 = \rho dt$. Let $C(t)$ be the date t price of a European option written on the difference between these two stocks' prices. Specifically, at this option's maturity date, T, the value of the option equals

$$C(T) = \max[0, S_1(T) - S_2(T)] \tag{10.37}$$

Now define $c(t) = C(t)/S_2(t)$, $s(t) \equiv S_1(t)/S_2(t)$, and $B(t) = S_2(t)/S_2(t) = 1$ as the deflated price processes, where the prices of the option, stock 1, and stock 2 are all normalized by the price of stock 2. With this normalized price system, the terminal payoff corresponding to (10.37) is now

$$c(T) = \max[0, s(T) - 1] \tag{10.38}$$

Applying Itô's lemma, the process for $s(t)$ is given by

$$ds/s = \mu_s dt + \sigma_s dz_3 \tag{10.39}$$

where $\mu_s \equiv \mu_1 - \mu_2 + \sigma_2^2 - \rho\sigma_1\sigma_2$, $\sigma_s dz_3 \equiv \sigma_1 dz_1 - \sigma_2 dz_2$, and $\sigma_s^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$. Further, when prices are measured in terms of stock 2, the deflated price of stock 2 becomes the riskless asset, with the riskless rate of return given by $dB/B = 0dt$. That is, because the deflated price of stock 2 never changes, it returns a riskless rate of zero. Using Itô's lemma once again, the deflated option price, $c(s(t), t)$, follows the process

$$dc = \left[c_s \mu_s s + c_t + \frac{1}{2} c_{ss} \sigma_s^2 s^2\right] dt + c_s \sigma_s s \, dz_3 \tag{10.40}$$

With this normalized price system, the usual Black-Scholes hedge portfolio can be created from the option and stock 1. The hedge portfolio's value is given by

$$H = -c + c_s s \tag{10.41}$$

and the instantaneous change in value of the portfolio is

$$\begin{aligned}
dH &= -\,dc + c_s ds \\
&= -\left[c_s\mu_s s + c_t + \frac{1}{2}c_{ss}\,\sigma_s^2 s^2\right] dt - c_s\,\sigma_s s\,dz_3 + c_s\mu_s s\,dt + c_s\sigma_s s\,dz_3 \\
&= -\left[c_t + \frac{1}{2}c_{ss}\,\sigma_s^2 s^2\right] dt
\end{aligned} \tag{10.42}$$

When measured in terms of stock 2's price, the return on this portfolio is instantaneously riskless. In the absence of arbitrage, it must earn the riskless return, which as noted previously, equals zero under this deflated price system. Thus we can write

$$dH = -\left[c_t + \frac{1}{2}c_{ss}\,\sigma_s^2 s^2\right] dt = 0 \tag{10.43}$$

which implies

$$c_t + \frac{1}{2}c_{ss}\,\sigma_s^2 s^2 = 0 \tag{10.44}$$

which is the Black-Scholes partial differential equation but with the risk-free rate, r, set to zero. Solving it subject to the boundary condition (10.38), which implies a unit exercise price, gives the usual Black-Scholes formula

$$c(s,\,t) = s\,N(d_1) - N(d_2) \tag{10.45}$$

where

$$\begin{aligned}
d_1 &= \frac{\ln\left(s(t)\right) + \frac{1}{2}\sigma_s^2\,(T-t)}{\sigma_s\,\sqrt{T-t}} \\
d_2 &= d_1 - \sigma_s\sqrt{T-t}
\end{aligned} \tag{10.46}$$

To convert back to the undeflated price system, we simply multiply (10.45) by

$S_2(t)$ and obtain

$$C(t) = S_1 N(d_1) - S_2 N(d_2) \tag{10.47}$$

Note that the option price does not depend on the nondeflated price system's risk-free rate, $r(t)$. Hence, the formula holds even for stochastic interest rates.

10.4 Applications

This section illustrates the usefulness of the martingale pricing technique. The first set of applications deals with options written on assets that continuously pay dividends. Examples include an option written on a foreign currency and an option written on a futures price. The second application is to value bonds of different maturities, which determines the term structure of interest rates.

10.4.1 Continuous Dividends

Many types of contingent claims depend on an underlying asset that can be interpreted as paying a continuous dividend that is proportional to the asset's price. Let us apply the risk-neutral pricing method to value an option on such an asset. Denote as $S(t)$ the date t price of an asset that continuously pays a dividend that is a fixed proportion of its price. Specifically, the asset pays a dividend of $\delta S(t)\,dt$ over the time interval dt. The process followed by this asset's price can be written as

$$dS = (\mu - \delta)\,Sdt + \sigma Sdz \tag{10.48}$$

where σ is the standard deviation of the asset's rate of return and μ is the asset's total expected rate of return, which includes its dividend payment and price appreciation. Similar to the assumptions of Black and Scholes, σ and δ are assumed to be constant, but μ may be a function of S and t. Now note that the total rate of return earned by the owner of one share of this asset is $dS/S + \delta dt = \mu dt + \sigma dz$. Consider a European call option written on this asset that has an exercise price of X and a maturity date of $T > t$, where we define $\tau \equiv T - t$. Assuming a constant interest rate equal to r, we use equation (10.17) to write the date t price of this option as

$$\begin{aligned} c(t) &= \widehat{E}_t \left[e^{-r\tau} c(T) \right] \\ &= e^{-r\tau} \widehat{E}_t \left[\max \left[S(T) - X, 0 \right] \right] \end{aligned} \tag{10.49}$$

To calculate the expectation in (10.49), we need to consider the distribution of $S(T)$. Note that because μ could be a function of S and t, the distribution of $S(T)$ under the physical P measure cannot be determined until this functional relationship $\mu(S,t)$ is specified. However, (10.49) requires the distribution of $S(T)$ under the risk-neutral Q measure, and given the assumption of a constant risk-free rate, this distribution already is determined. As in (10.10), converting from the physical measure generated by dz to the risk-neutral measure generated by $d\widehat{z}$ removes the risk premium from the asset's expected rate of return. Hence, the risk-neutral process for the stock price becomes

$$dS = (r - \delta) S dt + \sigma S d\widehat{z} \tag{10.50}$$

Since $r - \delta$ and σ are constants, we know that S follows geometric Brownian motion, and hence is lognormally distributed, under Q. From our previous results, we also know that the risk-neutral distribution of $\ln[S(T)]$ is normal:

$$\ln[S(T)] \sim N\left(\ln[S(t)] + (r - \delta - \frac{1}{2}\sigma^2)\tau,\ \sigma^2\tau \right) \tag{10.51}$$

Equation (10.49) can now be computed as

$$\begin{aligned} c(t) &= e^{-r\tau}\widehat{E}_t\left[\max\left[S(T)-X,0\right]\right] \\ &= e^{-r\tau}\int_X^{\infty}\left(S(T)-X\right)g(S(T))\,dS(T) \end{aligned} \tag{10.52}$$

where $g(S_T)$ is the lognormal probability density function. This integral can be evaluated by making the change in variable

$$Y = \frac{\ln\left[S(T)/S(t)\right]-\left(r-\delta-\frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}} \tag{10.53}$$

which from (10.51) transforms the lognormally distributed $S(T)$ into the variable Y distributed $N(0,1)$. The result is the modified Black-Scholes formula

$$c = Se^{-\delta\tau}N(d_1) - Xe^{-r\tau}N(d_2) \tag{10.54}$$

where

$$\begin{aligned} d_1 &= \frac{\ln(S/X)+\left(r-\delta+\frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}} \\ d_2 &= d_1-\sigma\sqrt{\tau} \end{aligned} \tag{10.55}$$

In this case of a European call option where the payoff is $c(T) = c(S(T)) = \max\left[S(T)-X,0\right]$ and the risk-neutral process for $S(t)$ is geometric Brownian motion, it is possible to derive a closed-form solution for the expectation in (10.49). However, it should be noted that for many applications where contingent claims have more complex payoffs and/or the underlying asset follows a more complicated risk-neutral process, a closed-form solution may not be possible. Still, it is often the case that a contingent claim's value of the form $c(t) = \widehat{E}_t\left[e^{-r\tau}c(S(T))\right]$ can be computed numerically. One important approach, pioneered by Phelim Boyle (Boyle 1977), uses Monte Carlo simulation. A random number generator is used to simulate a large number of risk-neutral paths for $S(t)$ in order to generate a risk-neutral frequency distribution of the underlying asset's value at date T.[14] By taking the discounted average of the random outcomes $c(S(T))$, the risk-neutral expectation $\widehat{E}_t\left[e^{-r\tau}c(S(T))\right]$ is then computed.

Also, note that if one compares the formula in (10.54) and (10.55) to Chapter 9's equations (9.24) and (9.25), the value of an option written on an asset that pays no dividends, the only difference is that the non-dividend-paying asset's price, $S(t)$, is replaced with the dividend-discounted price of the dividend-paying asset, $S(t)e^{-\delta\tau}$. The intuition behind this can be seen by realizing that if no dividends are paid, then $\hat{E}_t[S(T)] = S(t)e^{r\tau}$. However, with dividends, the risk-neutral expected asset price appreciates at rate $r-\delta$, rather than r. This is because with dividends paid out at rate δ, expected price appreciation must be at rate $r-\delta$ to keep the total expected rate of return equal to $\delta + r - \delta = r$. Thus, the risk-neutral expectation of $S(T)$ is

$$\begin{aligned} \hat{E}_t[S(T)] &= S(t)e^{(r-\delta)\tau} \\ &= S(t)e^{-\delta\tau}e^{r\tau} = \overline{S}(t)e^{r\tau} \end{aligned} \tag{10.56}$$

where we define $\overline{S}(t) \equiv S(t)e^{-\delta\tau}$. This shows that the value of an option on

[14] This simulation is similar to that which is illustrated in Figure 8.2. However, in Figure 8.2 the underlying asset's physical process with $\mu = 0.10$ and $\sigma = 0.30$ was simulated. To simulate its risk-neutral process, one would replace μ with the risk-free interest rate, say $r = 0.05$. The exact discrete-time distribution to simulate with a random number generator may be found from the underlying asset's continuous-time distribution using the Kolmogorov equation (8.31).

a dividend-paying asset with current price S equals the value of an option on a non-dividend-paying asset having current price $\overline{S} = Se^{-\delta\tau}$.

Formula (10.54) can be applied to an option on a foreign currency. If $S(t)$ is defined as the domestic currency value of a unit of foreign currency, that is, the spot exchange rate, then assuming this rate has a constant volatility gives it a process satisfying (10.48). Since purchase of a foreign currency allows the owner to invest it in an interest-earning asset yielding the foreign currency interest rate, r_f, the dividend yield will equal this foreign currency rate, $\delta = r_f$. Hence, $\hat{E}_t[S(T)] = S(t)\, e^{(r-r_f)\tau}$, where the domestic and foreign currency interest rates are those for a risk-free investment having a maturity equal to that of the option. Note that this expression is the no-arbitrage value of the date t forward exchange rate having a time until maturity of τ, that is, $F_{t,\tau} = Se^{(r-r_f)\tau}$.[15] Therefore, equation (10.54) can be written as

$$c(t) = e^{-r\tau}\left[F_{t,\tau} N(d_1) - XN(d_2)\right] \tag{10.57}$$

where $d_1 = \frac{\ln[F_{t,\tau}/X] + \frac{\sigma^2}{2}\tau}{\sigma\sqrt{\tau}}$, and $d_2 = d_1 - \sigma\sqrt{\tau}$.

A final example is an option written on a futures price. Options are written on the futures prices of commodities, equities, bonds, and currencies. Futures prices are similar to forward prices.[16] Like a forward contract, futures contracts involve long and short parties, and if both parties maintain their positions until the maturity of the contract, their total profits equal the difference between the underlying asset's maturity value and the initial future price. The main difference between futures contracts and forward contracts is that a futures

[15] This is the same formula as (3.19) or (7.2) but with continuously compounded yields.

[16] See (Cox, Ingersoll, and Ross 1981) and (Jarrow and Oldfield 1981) for a comparison of forward and futures contracts. If markets are frictionless, there are no arbitrage opportunities, and default-free interest rates are nonstochastic, then it can be shown that forward and futures prices are equivalent for contracts written on the same underlying asset and having the same maturity date. When interest rates are stochastic, then futures prices will be greater (*less*) than equivalent contract forward prices if the underlying asset is positively (*negatively*) correlated with short-term interest rates.

contract is "marked-to-market" daily; that is, the futures price for a particular maturity contract is recomputed daily and profits equal to the difference between today's and yesterday's future price are transferred (settled) from the short party to the long party on a daily basis. Thus, if F_{t,t^*} is the date t futures price for a contract maturing at date t^*, then the undiscounted profit (*loss*) earned by the long (*short*) party over the period from date t to date $T \leq t^*$ is simply $F_{T,t^*} - F_{t,t^*}$. Like forward contracts, there is no initial cost for the parties who enter into a futures contract. Hence, in a risk-neutral world, their expected profits must be zero. This implies that

$$\hat{E}_t \left[F_{T,t^*} - F_{t,t^*}\right] = 0 \tag{10.58}$$

or that under the Q measure, the futures price is a martingale:

$$\hat{E}_t \left[F_{T,t^*}\right] = F_{t,t^*} \tag{10.59}$$

Thus, while under the Q measure a non-dividend-paying asset price would be expected to grow at rate r, a futures price would be expected to grow at rate 0. Hence, futures are like assets with a dividend yield $\delta = r$. From this, one can derive the value of a futures call option that matures in τ periods where $\tau \leq (t^* - t)$ as

$$c\left(t\right) = e^{-r\tau}\left[F_{t,t^*} N\left(d_1\right) - XN\left(d_2\right)\right] \tag{10.60}$$

where $d_1 = \frac{\ln\left[F_{t,t^*}/X\right] + \frac{\sigma^2}{2}\tau}{\sigma\sqrt{\tau}}$, and $d_2 = d_1 - \sigma\sqrt{\tau}$. Note that this is similar in form to an option on a foreign currency written in terms of the forward exchange rate.

10.4.2 The Term Structure Revisited

The martingale pricing equation (10.17) can be applied to deriving the date t price of a default-free bond that matures in τ periods and pays \$1 at the maturity date $T = t + \tau$. This allows us to value default-free bonds in a manner that is an alternative to the partial differential equation approach of the previous chapter. Using the same notation as in Chapter 9, let $P(t, \tau)$ denote this bond's current price. Then, since $c(T) = P(T, 0) = 1$, equation (10.17) becomes

$$P(t, \tau) = \widehat{E}_t \left[e^{-\int_t^T r(u)du} 1 \right] \tag{10.61}$$

We now rederive the Vasicek model using this equation. To apply equation (10.61), we need to find the risk-neutral (Q measure) process for the instantaneous maturity interest rate, $r(t)$. Recall that the physical (P measure) process for the interest rate was assumed to be the Ornstein-Uhlenbeck process

$$dr(t) = \alpha \left[\overline{r} - r(t) \right] dt + \sigma_r dz_r \tag{10.62}$$

and that the market price of bond risk, q, was assumed to be a constant. This implied that the expected rate of return on all bonds satisfied

$$\mu_p(r, \tau) = r(t) + q\sigma_p(\tau) \tag{10.63}$$

where $\sigma_p(\tau) = -P_r \sigma_r / P$. Thus, the physical process for a bond's price, given by equation (9.31), can be rewritten as

$$\begin{aligned} dP(r, \tau) / P(r, \tau) &= \mu_p(r, \tau) dt - \sigma_p(\tau) dz_r \\ &= \left[r(t) + q\sigma_p(\tau) \right] dt - \sigma_p(\tau) dz_r \end{aligned} \tag{10.64}$$

Now note that if we define the transformed Brownian motion process $d\widehat{z}_r = dz_r - qdt$, then equation (10.64) becomes

$$\begin{aligned} dP(t,\tau)/P(t,\tau) &= [r(t) + q\sigma_p(\tau)]\,dt - \sigma_p(\tau)[d\widehat{z}_r + qdt] \\ &= r(t)\,dt - \sigma_p(\tau)\,d\widehat{z}_r \end{aligned} \tag{10.65}$$

which is the risk-neutral, Q measure process for the bond price. This is so because under this transformation all bond prices now have an expected rate of return equal to the instantaneously risk-free rate, $r(t)$. Therefore, applying this same Brownian motion transformation to equation (10.62), we find that the instantaneous maturity interest rate process under the Q measure is

$$\begin{aligned} dr(t) &= \alpha[\overline{r} - r(t)]\,dt + \sigma_r[d\widehat{z}_r + qdt] \\ &= \alpha\left[\left(\overline{r} + \frac{q\sigma_r}{\alpha}\right) - r(t)\right]dt + \sigma_r d\widehat{z}_r \end{aligned} \tag{10.66}$$

Hence, we see that the risk-neutral process for $r(t)$ continues to be an Ornstein-Uhlenbeck process but with a different unconditional mean, $\overline{r} + q\sigma_r/\alpha$. Thus, we can use the valuation equation (10.61) to compute the discounted value of the bond's \$1 payoff, $P(t,\tau) = \widehat{E}_t\left[\exp\left(-\int_t^T r(u)\,du\right)\right]$, assuming $r(t)$ follows the process in (10.66). Doing so leads to the same solution given in the previous chapter, equation (9.41).[17]

The intuition for why (10.66) is the appropriate risk-neutral process for $r(t)$ is as follows. Note that if the market price of risk, q, is positive, then the risk-neutral mean, $\overline{r} + q\sigma_r/\alpha$, exceeds the physical process's mean, $\overline{r}$. In this case, when we use the valuation equation $P(t,\tau) = \widehat{E}_t\left[\exp\left(-\int_t^T r(u)\,du\right)\right]$,

[17]Since the Ornstein-Uhlenbeck process in (10.66) is normally distributed, the integral $\int_t^T r(u)\,du$ is also normally distributed based on the idea that sums (an integral) of normals are normal. Hence, $\exp\left[-\int_t^T r(u)\,du\right]$ is lognormally distributed.

the expected risk-neutral discount rate is greater than the physical expectation of $r(t)$. Therefore, ceteris paribus, the greater is q, the lower will be the bond's price, $P(t,\tau)$, and the greater will be its yield to maturity, $Y(t,\tau)$. Thus, the greater the market price of interest rate risk, the lower are bond prices and the greater are bond yields.

10.5 Summary

This chapter has covered much ground. Yet, many of its results are similar to discrete-time counterparts derived in Chapter 4. The martingale pricing method essentially is a generalization of risk-neutral pricing and is applicable in complete market economies when arbitrage opportunities are not present. A continuous-time state price deflator can also be derived when asset markets are dynamically complete. We demonstrated that this pricing kernel is expected to grow at minus the short-term interest rate and that the standard deviation of its growth is equal to the market price of risk. We also saw that contingent claims valuation often can be simplified by an appropriate normalization of asset prices. In some cases, this is done by deflating by the price of a riskless asset, and in others by deflating by a risky-asset price. A final set of results included showing how the martingale approach can be applied to valuing a contingent claim written on an asset that pays a continuous, proportional dividend. Important examples of this included options on foreign exchange and on futures prices. Also included was an illustration of how the martingale method can be applied to deriving the term structure of interest rates.

10.6 Exercises

1. In this problem, you are asked to derive the equivalent martingale measure and the pricing kernel for the case to two sources of risk. Let S_1 and S_2 be the values of two risky assets that follow the processes

$$dS_i/S_i = \mu_i dt + \sigma_i dz_i, i = 1, 2$$

where both μ_i and σ_i may be functions of S_1, S_2, and t, and dz_1 and dz_2 are two independent Brownian motion processes, implying $dz_1 dz_2 = 0$. Let $f(S_1, S_2, t)$ denote the value of a contingent claim whose payoff depends solely on S_1, S_2, and t. Also let $r(t)$ be the instantaneous, risk-free interest rate. From Itô's lemma, we know that the derivative's value satisfies

$$df = \mu_f f dt + \sigma_{f1} f dz_1 + \sigma_{f2} f dz_2$$

where $\mu_f f = f_3 + \mu_1 S_1 f_1 + \mu_2 S_2 f_2 + \frac{1}{2}\sigma_1^2 S_1^2 f_{11} + \frac{1}{2}\sigma_2^2 S_2^2 f_{22}$, $\sigma_{f1} f = \sigma_1 S_1 f_1$, $\sigma_{f2} f = \sigma_2 S_2 f_2$ and where the subscripts on f denote the partial derivatives with respect to its three arguments, S_1, S_2, and t.

a. By forming a riskless portfolio composed of the contingent claim and the two risky assets, show that in the absence of arbitrage an expression for μ_f can be derived in terms of r, $\theta_1 \equiv \frac{\mu_1 - r}{\sigma_1}$, and $\theta_2 \equiv \frac{\mu_2 - r}{\sigma_2}$.

b. Define the risk-neutral processes $d\widehat{z}_1$ and $d\widehat{z}_2$ in terms of the original Brownian motion processes, and then give the risk-neutral process for df in terms of $d\widehat{z}_1$ and $d\widehat{z}_2$.

c. Let $B(t)$ be the value of a "money market fund" that invests in the instantaneous maturity, risk-free asset. Show that $F(t) \equiv f(t)/B(t)$ is a martingale under the risk-neutral probability measure.

d. Let $M(t)$ be the state price deflator such that $f(t) M(t)$ is a martingale under the physical probability measure. If

$$dM = \mu_m dt + \sigma_{m1} dz_1 + \sigma_{m2} dz_2$$

what must be the values of μ_m, σ_{m1}, and σ_{m2} that preclude arbitrage? Show how you solve for these values.

2. The Cox, Ingersoll, and Ross (Cox, Ingersoll, and Ross 1985b) model of the term structure of interest rates assumes that the process followed by the instantaneous maturity, risk-free interest rate is

$$dr = \alpha(\gamma - r) dt + \sigma\sqrt{r} dz$$

where α, γ, and σ are constants. Let $P(t, \tau)$ be the date t price of a zero-coupon bond paying \$1 at date $t + \tau$. It is assumed that $r(t)$ is the only source of uncertainty affecting $P(t, \tau)$. Also, let $\mu_p(t, \tau)$ and $\sigma_p(t, \tau)$ be the instantaneous mean and standard deviation of the rate of return on this bond and assume

$$\frac{\mu_p(t, \tau) - r(t)}{\sigma_p(t, \tau)} = \beta\sqrt{r}$$

where β is a constant.

a. Write down the stochastic process followed by the pricing kernel (state price deflator), $M(t)$, for this problem, that is, the process dM/M. Also, apply Itô's lemma to derive the process for $\mathrm{m}(t) \equiv \ln(M)$, that is, the process $d\mathrm{m}$.

b. Let the current date be 0 and write down the formula for the bond price, $P(0, \tau)$, in terms of an expectation of $\mathrm{m}_\tau - \mathrm{m}_0$. Show how this can be written in terms of an expectation of functions of integrals of $r(t)$ and β.

3. If the price of a non-dividend-paying stock follows the process $dS/S = \mu dt + \sigma dz$ where σ is constant, and there is a constant risk-free interest rate equal to r, then the Black-Scholes derivation showed that the no-arbitrage value of a standard call option having τ periods to maturity and an exercise price of X is given by $c = SN(d_1) - Xe^{-r\tau}N(d_2)$ where $d_1 = \left[\ln(S/X) + \left(r + \frac{1}{2}\sigma^2\right)\tau\right] / (\sigma\sqrt{\tau})$ and $d_2 = d_1 - \sigma\sqrt{\tau}$.

   A *forward start* call option is similar to this standard option but with the difference that the option exercise price, X, is initially a random variable. The exercise price is set equal to the contemporaneous stock price at a future date prior to the maturity of the option. Specifically, let the current date be 0 and the option maturity date be τ. Then at date t where $0 < t < \tau$, the option's exercise price, X, is set equal to the date t value of the stock, denoted as $S(t)$. Hence, $X = S(t)$ is a random variable as of the current date 0.

   For a given date t, derive the date 0 value of this forward start call option. Hint: note the value of a standard call option when $S = X$, and then use a simple application of risk-neutral pricing to derive the value of the forward start option.

4. If the price of a non-dividend-paying stock follows the process $dS/S = \mu dt + \sigma dz$ where σ is constant, and there is a constant risk-free interest rate equal to r, then the Black-Scholes showed that the no-arbitrage value of a standard call option having τ periods to maturity and an exercise price of X is given by $c = SN(d_1) - Xe^{-r\tau}N(d_2)$ where $d_1 =$

$\left[\ln\left(S/X\right)+\left(r+\frac{1}{2}\sigma^2\right)\tau\right]/\left(\sigma\sqrt{\tau}\right)$ and $d_2 = d_1 - \sigma\sqrt{\tau}$. Based on this result and a simple application of risk-neutral pricing, derive the value of the following *binary* options. Continue to assume that the underlying stock price follows the process $dS/S = \mu dt + \sigma dz$, the risk-free interest rate equals r, and the option's time until maturity equals τ.

a. Consider the value of a *cash-or-nothing call*, *cnc*. If $S(T)$ is the stock's price at the option's maturity date of T, the payoff of this option is

$$cnc_T = \begin{cases} F & \text{if } S(T) > X \\ 0 & \text{if } S(T) \leq X \end{cases}$$

where F is a fixed amount. Derive the value of this option when its time until maturity is τ and the current stock price is S. Explain your reasoning.

b. Consider the value of an *asset-or-nothing call*, *anc*. If $S(T)$ is the stock's price at the option's maturity date of T, the payoff of this option is

$$anc_T = \begin{cases} S(T) & \text{if } S(T) > X \\ 0 & \text{if } S(T) \leq X \end{cases}$$

Derive the value of this option when its time until maturity is τ and the current stock price is S. Explain your reasoning.

5. Outline a derivation of the form of the multivariate state price deflator given in equations (10.33) and (10.34).

6. Consider a continuous-time version of a Lucas endowment economy (Lucas 1978). It is assumed that there is a single risky asset (e.g., fruit tree) that produces a perishable consumption good that is paid out as a continuous

dividend, g_t. This dividend satisfies the process

$$dg_t/g_t = \alpha dt + \sigma dz$$

where α and σ are constants. There is a representative agent who at date 0 maximizes lifetime consumption given by

$$E_0 \int_0^\infty U(C_t, t)\, dt$$

where $U(C_t, t) = e^{-\phi t} C_t^\gamma / \gamma$, $\gamma < 1$. Under the Lucas endowment economy assumption, we know that in equilibrium $C_t = g_t$.

a. Let $P_t(\tau)$ denote the date t price of a riskless discount (zero-coupon) bond that pays one unit of the consumption good in τ periods. Derive an (Euler equation) expression for $P_t(\tau)$ in terms of an expectation of a function of future dividends.

b. Let $m_{t,t+\tau} \equiv M_{t+\tau}/M_t$ be the stochastic discount factor (pricing kernel) for this economy. Based on your answer in part (a), write down the stochastic process for M_t. Hint: find an expression for M_t and then use Itô's lemma.

c. Based on your previous answers, write down the instantaneous, risk-free real interest rate. Is it constant or time varying?

Chapter 11

Mixing Diffusion and Jump Processes

We have studied the nature and application of diffusion processes, which are continuous-time stochastic processes whose uncertainty derives from Brownian motions. While these processes have proved useful in modeling many different types of economic and financial time series, they may be unrealistic for modeling random variables whose values can change very significantly over a short period of time. This is because diffusion processes have continuous sample paths and cannot model discontinuities, or "jumps," in their values. In some situations, it may be more accurate to allow for large, sudden changes in value. For example, when the release of significant new information results in an immediate, substantial change in the market value of an asset, then we need to augment the diffusion process with another type of uncertainty to capture this discontinuity in the asset's price. This is where Poisson jump processes can be useful. In particular, we can model an economic or financial time series as the sum of diffusion (Brownian motion-based) processes and jump processes.

The first section of this chapter introduces the mathematics of a process that is a mixture of a jump process and a diffusion process. Section 11.2 shows how Itô's lemma can be extended to derive the process of a variable that is a function of a mixed jump-diffusion process. It comes as no surprise that this function inherits the risk of both the Brownian motion component as well as the jump component of the underlying process. Section 11.3 revisits the problem of valuing a contingent claim, but now assumes that the underlying asset's price follows a mixed jump-diffusion process. Our analysis follows that of Robert Merton (Merton 1976), who first analyzed this subject. In general, the inclusion of a jump process means that a contingent claim's risk cannot be perfectly hedged by trading in the underlying asset. In this situation of market incompleteness, additional assumptions regarding the price of jump risk need to be made in order to value derivative securities. We show how an option can be valued when the underlying asset's jump risk is perfectly diversifiable. The problem of option valuation when the underlying asset is the market portfolio of all assets is also discussed.

11.1 Modeling Jumps in Continuous Time

Consider the following continuous-time process:

$$dS/S = (\mu - \lambda k)\, dt + \sigma\, dz + \gamma(Y)\, dq \tag{11.1}$$

where dz is a standard Wiener (Brownian motion) process and $q(t)$ is a Poisson counting process that increases by 1 whenever a Poisson-distributed event occurs. Specifically, $dq(t)$ satisfies

$$dq = \begin{cases} 1 & \text{if a jump occurs} \\ 0 & \text{otherwise} \end{cases} \tag{11.2}$$

During each time interval, dt, the probability that $q(t)$ will augment by 1 is $\lambda(t)\,dt$, where $\lambda(t)$ is referred to as the Poisson intensity. When a Poisson event does occur, say, at date $\widehat{t}$, then there is a discontinuous change in S equal to $dS = \gamma(Y)\,S$ where γ is a function of $Y\left(\widehat{t}\right)$, which may be a random variable realized at date $\widehat{t}$.[1] In other words, if a Poisson event occurs at date $\widehat{t}$, then $dS\left(\widehat{t}\right) = S\left(\widehat{t}^{+}\right) - S\left(\widehat{t}^{-}\right) = \gamma(Y)\,S\left(\widehat{t}^{-}\right)$, or

$$S\left(\widehat{t}^{+}\right) = [1 + \gamma(Y)]\,S\left(\widehat{t}^{-}\right) \tag{11.3}$$

Thus, if $\gamma(Y) > 0$, there is an upward jump in S; whereas if $\gamma(Y) < 0$, there is a downward jump in S. Now we can define $k \equiv E[\gamma(Y)]$ as the expected proportional jump given that a Poisson event occurs, so that the expected change in S from the jump component $\gamma(Y)\,dq$ over the time interval dt is $\lambda\,k\,dt$. Therefore,

if we wish to let the parameter μ denote the instantaneous total expected rate of return (rate of change) on S, we need to subtract off $\lambda\,k\,dt$ from the drift term of S:

$$\begin{aligned} E[dS/S] &= E[(\mu - \lambda k)\,dt] + E[\sigma\,dz] + E[\gamma(Y)\,dq] \\ &= (\mu - \lambda k)\,dt + 0 + \lambda\,k\,dt = \mu\,dt \end{aligned} \tag{11.4}$$

The sample path of $S(t)$ for a process described by equation (11.1) will be

[1] The date, or "point," of a jump, $\widehat{t}$, is associated with the attribute or "mark" $Y\left(\widehat{t}\right)$. Hence, $\left(\widehat{t}, Y\left(\widehat{t}\right)\right)$ is referred to as a *marked point process*, or space-time point process.

continuous most of the time, but can have finite jumps of differing signs and amplitudes at discrete points in time, where the timing of the jumps depends on the Poisson random variable $q(t)$ and the jump sizes depend on the random variable $Y(t)$. If $S(t)$ is an asset price, these jump events can be thought of as times when important information affecting the value of the asset is released.

Jump-diffusion processes can be generalized to a multivariate setting where the process for $S(t)$ can depend on multiple Brownian motion and Poisson jump components. Moreover, the functions μ, σ, λ, and γ may be time varying and depend on other variables that follow diffusion or jump-diffusion processes. In particular, if $\lambda(t)$ depends on a random state variable $x(t)$, where for example, $dx(t)$ follows a diffusion process, then $\lambda(t, x(t))$ is called a doubly stochastic Poisson process or Cox process. Wolfgang Runggaldier (Runggaldier 2003) gives an excellent review of univariate and multivariate specifications for jump-diffusion models. For simplicity, in this chapter we restrict our attention to univariate models.[2] Let us next consider an extension of Itô's lemma that covers univariate jump-diffusion processes.

11.2 Itô's Lemma for Jump-Diffusion Processes

Let $c(S, t)$ be the value of a variable that is a twice-differentiable function of $S(t)$, where $S(t)$ follows the jump-diffusion process in equation (11.1). For example, $c(S, t)$ might be the value of a derivative security whose payoff depends on an underlying asset having the current price $S(t)$. Itô's lemma can be extended to the case of mixed jump-diffusion processes, and this generalization implies that

[2] In Chapter 18, we consider examples of default risk models where λ and γ are permitted to be functions of other state variables that follow diffusion processes.

the value $c(S,t)$ follows the process

$$\begin{aligned} dc &= c_s\left[(\mu-\lambda k)S\,dt + \sigma S\,dz\right] + \frac{1}{2}c_{ss}\sigma^2 S^2\,dt + c_t\,dt \\ &\quad + \left\{c\left([1+\gamma(Y)]\,S,\,t\right) - c(S,\,t)\right\}dq \end{aligned} \tag{11.5}$$

where subscripts on c denote its partial derivatives. Note that the first line on the right-hand side of equation (11.5) is the standard form for Itô's lemma when $S(t)$ is restricted to following a diffusion process. The second line is what is new. It states that when S jumps, the contingent claim's value has a corresponding jump and moves from $c(S,\,t)$ to $c\left([1+\gamma(Y)]\,S,\,t\right)$. Now define μ_c as the instantaneous expected rate of return on c, that is, $E[dc/c] = \mu_c dt$. Also, define σ_c as the standard deviation of the instantaneous rate of return on c, conditional on a jump not occurring. Then we can rewrite equation (11.5) as

$$dc/c = \left[\mu_c - \lambda k_c(t)\right]dt + \sigma_c dz + \gamma_c(Y)\,dq \tag{11.6}$$

where

$$\mu_c \equiv \frac{1}{c}\left[c_s(\mu-\lambda k)\,S + \frac{1}{2}c_{ss}\sigma^2 S^2 + c_t\right] + \lambda k_c(t) \tag{11.7}$$

$$\sigma_c \equiv \frac{c_s}{c}\sigma S \tag{11.8}$$

$$\gamma_c = \left[c\left([1+\gamma(Y)]\,S,\,t\right) - c(S,t)\right]/c(S,t) \tag{11.9}$$

$$k_c(t) \equiv E_t\left[c\left([1+\gamma(Y)]\,S,\,t\right) - c(S,t)\right]/c(S,t) \tag{11.10}$$

Here, $k_c(t)$ is the expected proportional jump of the variable $c(S,t)$ given that a Poisson event occurs. In general, $k_c(t)$ is time varying. Let us now apply these results to valuing a contingent claim that depends on an asset whose price follows a jump-diffusion process.

11.3 Valuing Contingent Claims

This section follows work by Robert Merton (Merton 1976). For simplicity, the analysis that follows assumes that λ is constant over time and that $\gamma(Y) = (Y-1)$. Thus, if a jump occurs, the discontinuous change in S is $dS = (Y-1)S$. In other words, $S\left(\widehat{t}^-\right)$ goes to $S\left(\widehat{t}^+\right) = YS\left(\widehat{t}^-\right)$, where $\widehat{t}$ is the date of the jump. It is also assumed that successive random jump sizes, $(\widetilde{Y}-1)$, are independently and identically distributed.

Note that if μ and σ are constants, so that the continuous component of $S(t)$ is lognormally distributed, then conditional upon there being n jumps in the interval $(0, t)$,

$$\tilde{S}(t) = S(0)\, e^{(\mu-\frac{1}{2}\sigma^2-\lambda k)\, t + \sigma(\tilde{z}_t - z_0)}\, \tilde{y}(n) \tag{11.11}$$

where $\tilde{z}_t - z_0 \sim N(0, t)$ is the change in the Brownian motion process from date 0 to date t. Jump uncertainty is reflected in the random variable $\widetilde{y}(n)$, where $\tilde{y}(0) = 1$ and $\tilde{y}(n) = \prod_{i=1}^{n} \tilde{Y}_i$ for $n \geq 1$ where $\{\tilde{Y}_i\}_{i=1}^{n}$ is a set of independent identically distributed jumps. A verification of (11.11) is left as an exercise.

Similar to a Black-Scholes hedge portfolio, let us now consider an investment that includes a contingent claim (for example, a call option), its underlying asset, and the riskless asset.[3] Let the contingent claim's price be c and assume

[3] Our analysis regarding the return on a portfolio containing the underlying asset, the contingent claim, and the risk-free asset differs somewhat from our orginal Black-Scholes presentation, because here we write the portfolio's return in terms of the assets' portfolio proportions instead of units of their shares. To do this, we do not impose the requirement that the portfolio require zero net investment ($H(t) = 0$), since then portfolio proportions would be undefined. However, as before, we do require that the portfolio be self-financing.

the underlying asset's price follows the jump-diffusion process given in equation (11.1) with $\gamma(Y) = (Y-1)$. Furthermore, assume that the risk-free interest rate is a constant equal to r per unit time. Denote the proportions of the portfolio invested in the underlying asset, contingent claim, and risk-free asset as ω_1, ω_2, and $\omega_3 = 1-\omega_1-\omega_2$, respectively. The instantaneous rate of return on this portfolio, denoted dH/H, is given by

$$\begin{aligned} dH/H &= \omega_1\, dS/S + \omega_2\, dc/c + (1-\omega_1-\omega_2)r\, dt \qquad (11.12) \\ &= \left[\omega_1(\mu - r) + \omega_2(\mu_c - r) + r - \lambda(\omega_1 k + \omega_2 k_c)\right] dt \\ &\quad + (\omega_1\sigma + \omega_2\sigma_c)\, dz + \left[\omega_1\gamma(Y) + \omega_2\gamma_c(Y)\right] dq \end{aligned}$$

11.3.1 An Imperfect Hedge

Consider the possibility of choosing ω_1 and ω_2 in order to eliminate the risk from jumps. Note that while jumps occur simultaneously in the asset and the contingent claim, that is, jump risk is perfectly dependent for these two securities, these risks are not necessarily linearly dependent. This is because the contingent claim price, $c(S, t)$, is generally a nonlinear function of the asset price. Unlike Brownian motion-generated movements, jumps result in nonlocal changes in S and $c(S, t)$. When the underlying asset's jump size $(\tilde{Y} - 1)$ is random, the ratio between the size of the jump in S and the size of the jump in c, which is $\gamma\left(\widetilde{Y}\right)/\gamma_c\left(\widetilde{Y}\right)$, is unpredictable. Hence, a predetermined hedge ratio, ω_1/ω_2, that would eliminate all portfolio risk does not exist.[4] The implication is that one cannot perfectly replicate the contingent claim's payoff by a portfolio composed of the underlying asset and the risk-free asset. In this sense, the

[4] If the size of the jump is deterministic, a hedge that eliminates jump risk is possible. Alternatively, Phillip Jones (Jones 1984) shows that if the underlying asset's jump size has a discrete (finite state) distribution and a sufficient number of different contingent claims are written on this asset, a hedge portfolio that combines the underlying asset and these multiple contingent claims could also eliminate jump risk.

market for the contingent claim is incomplete.

Instead, suppose we pick ω_1 and ω_2 to eliminate only the risk from the continuous Brownian motion movements. This Black-Scholes hedge implies setting $\omega_1^*/\omega_2^* = -\sigma_c/\sigma = -c_s S/c$ from our definition of σ_c. This leads to the process for the value of the portfolio:

$$\begin{aligned} dH/H &= \left[\omega_1^* \left(\mu - r\right) + \omega_2^* \left(\mu_c - r\right) + r - \lambda \left(\omega_1^* \, k + \omega_2^* \, k_c\right)\right] dt \\ &\quad + \left[\omega_1^* \gamma \left(Y\right) + \omega_2^* \gamma_c \left(Y\right)\right] dq \end{aligned} \tag{11.13}$$

The return on this portfolio is a pure jump process. The return is deterministic, except when jumps occur. Using the definitions of γ, γ_c, and $\omega_1^* = -\omega_2^* c_s S/c$, we see that the portfolio jump term, $[\omega_1^* \gamma (Y) + \omega_2^* \gamma_c (Y)]\, dq$, equals

$$\begin{cases} \omega_2^* \left[\frac{c(S\tilde{Y},\, t) - c(S,\, t)}{c(S,\, t)} - c_s(S,\, t) \frac{S\tilde{Y} - S}{c(S,\, t)} \right] & \text{if a jump occurs} \\ 0 & \text{otherwise} \end{cases} \tag{11.14}$$

Now consider the case when the contingent claim is a European option on a stock with a time until expiration of τ and a strike price X. What would be the pattern of profits and losses on the (quasi-) hedge portfolio? We can answer this question by noting that if the rate of return on the underlying asset is independent of its price level, as is the case in equation (11.1), then the absence of arbitrage restricts the option price to a convex function of the asset price.[5] The option's convexity implies that $c(SY,\, t) - c(S,\, t) - c_s(S,\, t)[SY - S] \geq 0$ for all Y and t. This is illustrated in Figure 11.1 where the convex solid line gives the value of a call option as a function of its underlying asset's price.

From this fact and (11.14), we see that the unanticipated return on the hedge portfolio has the same sign as ω_2^*. This means that $\omega_1^* k + \omega_2^* k_c$, the

[5] For a proof, see Theorem 8.10 in Chapter 8 of (Merton 1992), which reproduces (Merton 1973b).

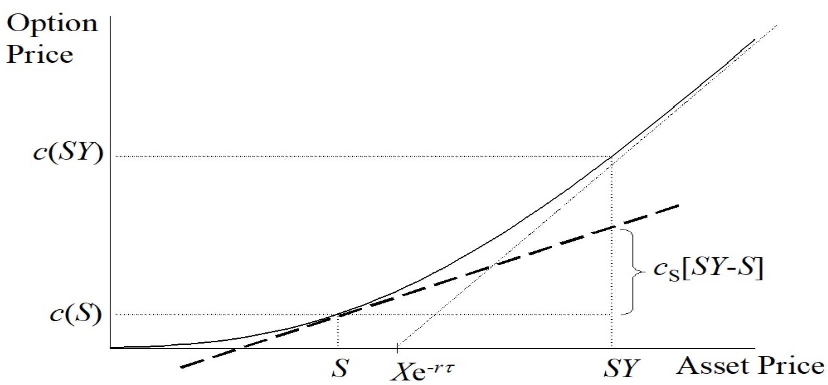

Figure 11.1: Hedge Portfolio Return with Jump

expected portfolio value jump size, also has the same sign as ω_2^*. Therefore, an option writer who follows this Black-Scholes hedge by being short the option ($\omega_2^* < 0$) and long the underlying asset earns, most of the time, more than the portfolio's expected rate of return. However, on those rare occasions when the underlying asset price jumps, a relatively large loss is incurred. Thus in "quiet" times, option writers appear to make positive excess returns. However, during infrequent "active" times, option writers suffer large losses.

11.3.2 Diversifiable Jump Risk

Since the hedge portfolio is not riskless but is exposed to jump risk, we cannot use the previous no-arbitrage argument to equate the hedge portfolio's rate of return to the risk-free rate. The hedge portfolio is exposed to jump risk and, in general, there may be a "market price" to such risk. One assumption might be that this jump risk is the result of purely firm specific information and, hence, the jump risk is perfectly diversifiable. This would imply that the market price of jump risk is zero. In this case, all of the risk of the hedge portfolio is diversifiable, so that its expected rate of return must equal the risk-free rate, r. Making this assumption implies

$$\omega_1^*(\mu - r) + \omega_2^*(\mu_c - r) + r = r \tag{11.15}$$

or

$$\omega_1^*/\omega_2^* = -\sigma_c/\sigma = -(\mu_c - r)/(\mu - r) \tag{11.16}$$

Now denote T as the maturity date of the contingent claim, and let us use the time until maturity $\tau \equiv T - t$ as the second argument for $c(S, \cdot)$ rather than calendar time, t. Hence, $c(S, \tau)$ is the price of the contingent claim when the current asset price is S and the time until maturity of the contingent claim is τ. With this redefinition, note that $c_\tau = -c_t$. Using (11.16) and substituting in for μ_c and σ_c from the definitions (11.7) and (11.8), we obtain the equilibrium partial differential equation

$$\frac{1}{2}\sigma^2 S^2 c_{ss} + (r - \lambda k)Sc_s - c_\tau - rc + \lambda E_t\left[c(S\tilde{Y}, \tau) - c(S, \tau)\right] = 0 \tag{11.17}$$

For a call option, this is solved subject to the boundary conditions $c(0, \tau) = 0$ and $c(S(T), 0) = \max[S(T) - X, 0]$. Note that when $\lambda = 0$, equation (11.17) is the standard Black-Scholes equation, which we know has the solution

$$b(S, \tau, X, \sigma^2, r) \equiv S\,N(d_1) - Xe^{-r\tau}N(d_2) \tag{11.18}$$

where $d_1 = [\ln(S/X) + (r + \frac{1}{2}\sigma^2)\tau]/(\sigma\sqrt{\tau})$ and $d_2 = d_1 - \sigma\sqrt{\tau}$. Robert Merton (Merton 1976) shows that the general solution to (11.17) is

$$c(S, \tau) = \sum_{n=0}^{\infty} \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} E_t\left[b(S\,\tilde{y}(n)\,e^{-\lambda k\tau}, \tau, X, \sigma^2, r)\right] \tag{11.19}$$

where, you may recall, $\tilde{y}(0) = 1$ and $\tilde{y}(n) = \prod_{i=1}^{n} \tilde{Y}_i$ for $n \geq 1$. The intuition behind the formula in (11.19) is that the option is a probability-weighted average of expected Black-Scholes option prices. Note that if the underlying asset price followed (11.1), then conditional on no jumps occurring over the life of the option, risk-neutral valuation would imply that the Black-Scholes option price would be $b(Se^{-\lambda k \tau}, \tau, X, \sigma^2, r)$.[6] Similarly, conditional on one jump occurring, risk-neutral valuation would imply that the option price would be $b(Sy(1)e^{-\lambda k \tau}, \tau, X, \sigma^2, r)$. Conditional on two jumps, it would be $b(Sy(2)e^{-\lambda k \tau}, \tau, X, \sigma^2, r)$, and thus for n jumps, it would be $b(Sy(n)e^{-\lambda k \tau}, \tau, X, \sigma^2, r)$.

Since $\frac{e^{-\lambda \tau}(\lambda \tau)^n}{n!}$ is the probability of n jumps occurring, we see that (11.19) is the jump-probability-weighted average of expected option values conditioned over all possible numbers of jumps.

11.3.3 Lognormal Jump Proportions

Under particular assumptions regarding the distribution of $\tilde{Y}$, solutions to (11.19) can be calculated numerically or, in some cases, in closed form. Here, we consider a case that leads to a closed-form solution, namely, the case in which $\tilde{Y}$ is lognormally distributed. Thus, if $E[\ln \tilde{Y}] \equiv \alpha - \frac{1}{2}\delta^2$ where $var[\ln \tilde{Y}] \equiv \delta^2$, then $E[\tilde{Y}] = e^{\alpha} = 1 + k$. Hence, $\alpha \equiv \ln(1+k)$. Given this assumption, if μ is assumed to be constant, the probability density for $\ln[S(t+\tau)]$, conditional on the value of $S(t)$, is

$$\sum_{n=0}^{\infty} g(\ln[S(t+\tau)/S(t)] \,|\, n) h(n) \tag{11.20}$$

[6] Recall that since the drift is $\mu - \lambda k$, and risk-neutral valuation sets $\mu = r$, then λk is like a dividend yield. Hence, $b(Se^{-\lambda k \tau}, \tau, X, \sigma^2, r)$ is the Black-Scholes formula for an asset with a dividend yield of λk.

where $g(\cdot\,|\,n)$ is the conditional density function given that n jumps occur during the interval between t and $t+\tau$, and $h(n)$ is the probability that n jumps occur between t and $t+\tau$. The values of these expressions are

$$g\left(\ln\left[\frac{S(t+\tau)}{S(t)}\right]|\,n\right) \equiv \frac{\exp\left[-\frac{\left(\ln\left[\frac{S(t+\tau)}{S(t)}\right]-\left(\mu-\lambda k+\frac{n\alpha}{\tau}-\frac{\nu_n^2}{2}\right)\tau\right)^2}{2\nu_n^2\tau}\right]}{\sqrt{2\pi\nu_n^2\tau}} \tag{11.21}$$

$$h(n) \equiv \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!} \tag{11.22}$$

where $\nu_n^2 \equiv \sigma^2+n\delta^2/\tau$ is the "average" variance per unit time. From (11.21), we see that conditional on n jumps occurring, $\ln[S(t+\tau)/S(t)]$ is normally distributed. Using the Cox-Ross risk-neutral (equivalent martingale) transformation, which allows us to set $\mu = r$, we can compute the date t risk-neutral expectation of $\max[S(T)-X,0]$, discounted by the risk-free rate, and conditional on n jumps occurring. This is given by

$$\begin{aligned} E_t[\,b(S\tilde{y}(n)e^{-\lambda k\tau},\tau,X,\sigma^2,r)\,] &= e^{-\lambda k\tau}(1+k)^n\,b(S,\tau,X,\nu_n^2,r_n) \\ &= e^{-\lambda k\tau}(1+k)^n\,b_n(S,\,\tau) \end{aligned} \tag{11.23}$$

where $b_n(S,\,\tau) \equiv b(S,\tau,X,\nu_n^2,r_n)$ and where $r_n \equiv r-\lambda k+n\alpha/\tau$. The actual value of the option is then the weighted average of these conditional values, where each weight equals the probability that a Poisson random variable with characteristic parameter $\lambda\tau$ will take on the value n. Defining $\lambda' \equiv \lambda(1+k)$, this equals

$$\begin{aligned} c(S,\,\tau) &= \sum_{n=0}^{\infty}\frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!}\,e^{-\lambda k\tau}\,(1+k)^n\,b_n(S,\,\tau) \\ &= \sum_{n=0}^{\infty}\frac{e^{-\lambda'\tau}(\lambda'\tau)^n}{n!}\,b_n(S,\,\tau) \end{aligned} \tag{11.24}$$

11.3.4 Nondiversifiable Jump Risk

In some circumstances, it is unrealistic to assume that jump risk is nonpriced risk. For example, David Bates (Bates 1991) investigated the U.S. stock market crash of 1987, an event that certainly was not firm specific but affected the entire market for equities. Similar work by Vasanttilak Naik and Moon Lee (Naik and Lee 1990) considered nondiversifiable jump risk. The models in these articles assume that aggregate wealth in the economy follows a mixed jump-diffusion process. This could result from a representative agent, Cox, Ingersoll, and Ross-type production economy in which technologies follow a jump-diffusion process and individuals select investments in these technologies such that their optimally invested aggregate wealth follows a mixed jump-diffusion process (Bates 1991). Or it can simply be assumed that the economy is a Lucas-type endowment economy and there is an exogenous firm dividend process that follows a mixed jump-diffusion process, and these dividends cannot be invested but must be consumed (Naik and Lee 1990).

In both articles, jumps in aggregate wealth or consumption (endowment) are assumed to be of the lognormal type that we assumed earlier. Further, representative individuals are assumed to have constant relative-risk-aversion utility. These assumptions allow the authors to solve for the general equilibrium price of jump risk. Given this setup, contingent claims, which are assumed to be in zero net supply, can be priced. For example, the formula for a call option derived by Bates has a series solution that is similar in form to equation (11.24).

11.3.5 Black-Scholes versus Jump-Diffusion Model

Having derived a model for pricing options written on an underlying asset whose price follows a jump-diffusion process, the natural question to ask is whether this makes any difference vis-à-vis the Black-Scholes option pricing model, which does not permit the underlying's price to jump. The answer is yes, and the jump-diffusion model appears to better fit the actual prices of many options written on stocks, stock indices, and foreign exchange. In most types of options, the Black-Scholes model underprices out-of-the-money and in-the-money options relative to at-the-money-options. What this means is that the prices of actual options whose exercise price is substantially different from the current price of the underlying are priced higher than the theoretical Black-Scholes price, while the prices of actual options whose exercise price is close to the current price of the underlying are priced lower than the theoretical Black-Scholes price. This phenomenon has been described as a *volatility smile* or *volatility smirk*.[7]

This empirical deficiency can be traced to the Black-Scholes model's assumption that the underlying's terminal price has a risk-neutral distribution that is lognormal. Apparently, investors price actual options under the belief that the risk-neutral distribution has much fatter "tails" than those of the lognormal distribution. In other words, investors price securities as if they believe that extreme asset prices are more likely than what would be predicted by a lognormal distribution, because actual in- and out-of-the-money options are priced relatively high versus the Black-Scholes theoretical prices. A model that permits the underlying asset's price to jump, with jumps possibly being both positive and negative, can generate a distribution for the asset's price that has

[7] Note that if the Black-Scholes model correctly priced all options having the same maturity date and the same underlying asset but different exercise prices, there would be one volatility parameter, σ, consistent with all of these options. However, the implied volatilities, σ, needed to fit in- and, especially, out-of-the-money call options are greater than the volatility parameter needed to fit at-the-money options. Hence, when implied volatility is graphed against call options' exercise prices, it forms an inverted hump, or "smile," or in the case of equity index options, a downward sloping curve, or "smirk." These characteristics of option prices are equivalent to the Black-Scholes model giving relatively low prices for in- and out-of-the-money options because options prices are increasing functions of the underlying's volatility, σ. The Black-Scholes model needs relatively high estimated volatility for in- and out-of-the-money options versus at-the-money options. If a (theoretically correct) single volatility parameter were used for all options, in- and out-of-the-money options would be relatively underpriced by the model. See (Hull 2000) for a review of this issue.

fatter tails than the lognormal. The possibility of jumps makes extreme price changes more likely and, indeed, the jump-diffusion option pricing model can better match the market prices of many types of options.

However, there are other aspects of actual option prices for which even the standard jump-diffusion model cannot account. The volatility parameters implied by actual option prices change over time and appear to follow a mean-reverting stochastic process. To account for this empirical time variation, *stochastic volatility* option pricing models have been developed. These models start by assuming that the underlying asset price follows a diffusion process such as $dS/S = \mu dt + \sigma dz$, but where the volatility, σ, is stochastic. The volatility follows a mean-reverting process of the form $d\sigma = \alpha(\sigma)\,dt + \beta(\sigma)\,dz_\sigma$, where dz_σ is another Brownian motion process possibly correlated with dz. Similar to the jump-diffusion model, one must assign a market price of risk associated with the volatility uncertainty reflected in the dz_σ term.[8]

While stochastic volatility option pricing models also produce fatter-tailed distributions relative to the lognormal, empirically these distributions do not tend to be fat enough to explain volatility smiles and smirks. To capture both time variation in volatilities and cross-sectional differences in volatility due to different degrees of "moneyness" (volatility smiles or smirks), it appears that an option pricing model that allows for both stochastic volatility and jumps is required.[9] For recent reviews of the empirical option pricing literature, see (Bates 2002) and (Bakshi, Cao, and Chen 1997).

[8] Steven Heston (Heston 1993) developed a popular stochastic volatility model.

[9] David Bates (Bates 1996) derived an option pricing model that combines both jumps and stochastic volatility and estimated its parameters using options on foreign exchange.

11.4 Summary

Allowing for the possibility of discontinuous movements can add realism to the modeling of asset prices. For example, a firm's stock price might experience a sudden, large change upon the public announcement that it is involved in a corporate merger. While the mixed jump-diffusion process captures such asset price dynamics, it complicates the valuation of contingent claims written on such an asset. In general, we showed that the contingent claim's payoff cannot be perfectly replicated by a dynamic trading strategy involving the underlying asset and risk-free asset. In this situation of market incompleteness, additional theory that assigns a market risk premium to jump risk is required to determine the contingent claim's value.

The additional complications in deriving jump-diffusion models of option pricing appear worthwhile. Because jumps increase the likelihood of extreme price movements, they generate a risk-neutral distribution of asset prices whose tails are fatter than the Black-Scholes model's lognormal distribution. Since the actual prices of many types of options appear to reflect significant probabilities of extreme movements in the underlying's price, the jump-diffusion model has better empirical performance.

Having seen that pricing contingent claims sometimes requires specifying market prices of risk, the following chapters turn to the subject of deriving equilibrium risk premia for assets in continuous-time economies. As a preliminary, we revisit the individual's consumption and portfolio choice problem when asset prices, and the individual's consumption and portfolio choices, can change continuously. Based on this structure of consumption and asset demands, we then derive assets' risk premia in a general equilibrium, continuous-time economy.

11.5 Exercises

1. Verify that (11.11) holds by using Itô's lemma to find the process followed by $\ln(S(t))$.

2. Let $S(t)$ be the U.S. dollar price of a stock. It is assumed to follow the process

$$dS/S = [\mu_s - \lambda k]\, dt + \sigma_s dz_s + \gamma\left(\widetilde{Y}\right) dq \qquad (*)$$

where dz_s is a standard Wiener process, $q(t)$ is a Poisson counting process, and $\gamma\left(\widetilde{Y}\right) = \left(\widetilde{Y} - 1\right)$. The probability that q will jump during the time interval dt is λdt. $k \equiv E\left[\tilde{Y} - 1\right]$ is the expected jump size. Let F be the foreign exchange rate between U.S. dollars and Japanese yen, denominated as U.S. dollars per yen. F follows the process

$$dF/F = \mu_f dt + \sigma_f dz_f$$

where $dz_s dz_f = \rho dt$. Define $x(t)$ as the Japanese yen price of the stock whose U.S. dollar price follows the process in (*). Derive the stochastic process followed by $x(t)$.

3. Suppose that the instantaneous-maturity, default-free interest rate follows the jump-diffusion process

$$dr(t) = \kappa\, [\theta - r(t)]\, dt + \sigma dz + r\gamma\,(Y)\, dq$$

where dz is a standard Wiener process and $q(t)$ is a Poisson counting process having the arrival rate of λdt. The arrival of jumps is assumed to be independent of the Wiener process, dz. $\gamma(Y) = (Y - 1)$ where $Y > 1$ is a known positive constant.

a. Define $P(r,\tau)$ as the price of a default-free discount bond that pays \$1 in τ periods. Using Itô's lemma for the case of jump-diffusion processes, write down the process followed by $dP(r,\tau)$.

b. Assume that the market price of jump risk is zero, but that the market price of Brownian motion (dz) risk is given by ϕ, so that $\phi = \left[\alpha_p - r(t)\right]/\sigma_p$, where $\alpha_p\left(r,\tau\right)$ is the expected rate of return on the bond and $\sigma_p\left(\tau\right)$ is the standard deviation of the bond's rate of return from Brownian motion risk (not including the risk from jumps). Derive the equilibrium partial differential equation that the value $P(r,\tau)$ must satisfy.

4. Suppose that a security's price follows a jump-diffusion process and yields a continuous dividend at a constant rate of δdt. For example, its price, $S\left(t\right)$, follows the process

$$dS/S = \left[\mu\left(S,t\right) - \lambda k - \delta\right]dt + \sigma\left(S,t\right)dz + \gamma\left(\widetilde{Y}\right)dq$$

where $q\left(t\right)$ is a Poisson counting process and $\gamma\left(\widetilde{Y}\right) = \left(\widetilde{Y}-1\right)$. Also let $k \equiv E\left[\widetilde{Y}-1\right]$; let the probability of a jump be λdt; and denote $\mu(S,t)$ as the asset's total expected rate of return. Consider a forward contract written on this security that is negotiated at date t and matures at date T where $\tau = T - t > 0$. Let $r(t,\tau)$ be the date t continuously compounded, risk-free interest rate for borrowing or lending between dates t and T. Assuming that one can trade continuously in the security, derive the equilibrium date t forward price using an argument that rules out arbitrage. Hint: some information in this problem is extraneous. The solution is relatively simple.