CHAPTER SIX

The Informational Role of Asset Prices
The Case of Implied Volatility

ZVI BODIE AND ROBERT C. MERTON

An important function of the financial system is to serve as a key source of information that helps coordinate decentralized decision-making in various sectors of the economy. Households and investors use interest rates and asset prices observed in fixed-income and equity markets in making their consumption saving decisions and their portfolio allocation decisions. Interest rates and prices also provide important signals to managers of firms selecting investment projects and financing.

The informational role of asset prices has long been recognized. In his classic work on the theory of interest, Irving Fisher (1907, 1930) explains in detail how information extractable from competitive financial markets facilitates the efficient separation and decentralization of intertemporal consumption and production decisions, thereby improving social welfare.

Samuelson (1965) and Fama (1965) expanded Fisher’s analysis by developing the efficient markets hypothesis, which holds that in a well-functioning and informed capital market, the best estimate of an asset’s future price is the current price, adjusted for a “fair” expected rate of return. That is, an asset’s current price fully reflects all publicly available information about future economic fundamentals affecting the asset’s value.¹

¹ The reasoning behind the efficient markets hypothesis (EMH) is that if security prices are not informationally efficient, well-informed investors have opportunities to make

Information that is reflected in asset prices can, under certain circumstances, be extracted from those prices. That is the specific focus of this chapter: extracting information about the volatility of changes in stock, bond, currency, and commodity prices.

Volatility is a measure of the uncertainty about future changes in price or rate of return on assets. It is a fundamental measure used to quantify risk in modern finance theory, and a critical input for virtually all decisions relating to risk management and strategic financial planning.

Until 1973, volatility was generally estimated by using historical data. Since the appearance of exchange-traded options in 1973, the concurrent development of the theory of contingent claims pricing has made it possible to infer beliefs about the future volatility of an asset directly from the prices of options and other securities whose payoffs depend in a nonlinear way on the price of the underlying asset. The estimate extracted in this way is called implied volatility.

This chapter begins with a brief historical perspective on the informational role of bond prices. We then illustrate the general informational role of asset prices—both spot and futures—in contributing to efficient resource allocation. Our substantive focus is the use of security prices to derive estimates of asset return volatility. We show how information about asset price volatility can be extracted from option prices using models such as the one developed by Black and Scholes (1973). Extensions of this analysis indicate that we can use implied volatilities and prices of publicly traded securities to value nontraded firm liabilities (such as employee stock options).

The Yield Curve in Historical Perspective

Perhaps the most basic use of asset prices in informing financial decisions is the use of interest rates derived from the market prices of default-free bonds. The earliest recorded interest rates date from around 3000 B.C. in Mesopotamia, and the earliest prices of market-traded bonds date from Venice in the fourteenth century A.D. Today, knowledge of the term structure of interest rates (also known as the yield curve) is the starting point for virtually any valuation model used in finance.

The most useful form in which to have this information is as a "dis-

“excess profits." Over time, competition leads to elimination of the excess profit opportunities, and asset prices become more reflective of the best available information about underlying fundamentals. For a review of the literature on the EMH, see Bodie, Kane, and Marcus (1993, Ch. 12).

count function” giving the price of a unit of currency to be received at various dates in the future. Effectively, this means one would like to have the prices of zero-coupon bonds of all possible maturities. This form greatly facilitates valuation of all default-free fixed-income instruments promising any arbitrary temporal pattern of cash flows.

But until very recently, bond market price data rarely came in zero-coupon form. Even today the U.S. Treasury issues only two different security types: bills with maturities up to one year, which are issued in pure discount form, and notes and bonds with a semiannual coupon, which are usually issued at or close to par. For many years “implied” interest rates on zero-coupon bonds with maturities longer than one year could only be inferred from the observed prices of coupon bonds.

More recently, a secondary market has developed for U.S. Treasury strips created by investment banking firms. These firms buy U.S. Treasury bonds and sell off the component cash flows. By observing the market prices of the resulting U.S. Treasury strips, one can now have direct knowledge of the zero-coupon yield curve. Similarly, the development of a swap market in international bonds makes it possible to have direct knowledge of the yield curve in different currencies.

This improvement in the information provided by financial markets was not the intention of the firms that created the U.S. Treasury strips or the bond swaps, but rather is a by-product of their profit-making activities. As the markets for fixed-income securities and their derivatives continue to evolve, more and more information useful in valuing other fixed-income instruments will become available.

Financial institutions rarely have as their manifest function to provide information. The production of price information is more properly viewed as a latent function of the financial system.3 In recent years, however, firms have emerged that specialize in gathering and analyzing asset price information.4

Futures Prices and the Efficient Storage of Commodities

The economics literature on the role of interest rates and asset prices in guiding resource allocation decisions is vast. Rather than attempt a survey, we instead offer a single example—commodity futures—to illustrate how the information that prices in the financial markets provide can enhance allocational efficiency.

---

3. That social behavior and organizations have manifest and latent functions is developed by R.K. Merton (1957, Ch. 1).
4. Prominent examples are Bloomberg and Reuters.
Commodity futures contracts are among the oldest financial instruments in existence. Futures contracts for rice were likely traded in ancient China. While the manifest function of commodity futures markets is to facilitate the reallocation of exposure to commodity price risk among market participants, commodity futures prices also play an important informational role. By providing a means to hedge the price risk associated with storing a commodity, futures contracts make it possible to separate the decision on whether to store a commodity physically from the decision to expose oneself financially to its price changes.

An example shows how this works. Suppose it is one month before the next harvest, and a wheat distributor has a ton of wheat in storage from the last harvest. The spot price of wheat is \( S \), and the futures price for delivery a month from now (after the new crop has been harvested) is \( F \). The distributor can hedge its exposure to price changes by either (1) selling the wheat in the spot market for \( S \), and delivering it immediately, or (2) selling a futures contract short at a price of \( F \), and delivering the wheat a month from now. In either case, the hedger has complete certainty about the price to be received for the wheat (assuming no risk of contract default).

Suppose a distributor’s cost of physically storing the wheat, the “cost of carry,” which includes interest, warehousing, and spoilage costs, is \( C \). This distributor will choose alternative (2) and carry the ton of wheat for another month (i.e., past the next harvest) only if \( F > S + C \). To put it slightly differently, the distributor will choose to carry the wheat for another month only if the cost of carrying it is less than the difference between the futures and the spot prices of wheat, (i.e., if \( C < F - S \)). Thus, the spread between the futures and the spot price governs how much wheat will be stored in aggregate and by whom: Wheat will be stored only by those distributors whose cost of carry is less than \( F - S \).

Suppose now that the next wheat harvest is expected to be a bountiful one. In that case, the equilibrium futures price of wheat may well be lower than the current spot price (\( F < S \)), and it will not pay for anyone to store wheat from this growing season into the next, even if it costs nothing to do so (i.e., \( C = 0 \)). The futures price of wheat conveys this information to all producers, distributors, and consumers of wheat, including those not transacting in the futures market.

To see how the futures market enhances allocational efficiency, consider what would happen in its absence. Now, every producer, distributor, and

---

5. For a brief review of the history of futures markets, see Miller (1990).
6. The decision to eliminate exposure to price changes is called “hedging.”

200 THE GLOBAL FINANCIAL SYSTEM
consumer of wheat would have to rely on forecasts of the future spot price of wheat in deciding whether to store it for another month. The futures market eliminates the need for all market participants to gather and process information in order to forecast the future spot price. Producers, distributors, and consumers of wheat may be in the best position to forecast future wheat prices (perhaps because they have low costs of gathering the relevant information), but others can also use the market. Competition among active forecasters in the futures markets will encourage those who have a comparative advantage in forecasting wheat prices to specialize in it.

Extracting Information from Asset Prices

Sometimes the information needed to make a particular financial decision is directly observable in the financial markets. If you are considering whether to refinance a mortgage, you can find out interest rates and other terms of the various financing alternatives directly and at low cost by looking in the newspaper or by calling a mortgage lender. Often, however, the information required to make a financial decision is not directly observable. It must be inferred or extracted from the prices you directly observe.

To extract information from the observed market prices of assets, an analyst must use a model that connects the information sought to the price. The analyst may have high or low confidence in the validity of the model as well as in the accuracy of the recorded market prices. Table 6-1 summarizes the possibilities.

Let us explore the implications of each of the four possible situations in Table 6-1. First, consider the cell labeled 1. If analysts are absolutely sure that the model used for extraction is correct, and that the observed prices are also correct, they can be certain that the information they are extracting is valid.

Suppose we want to obtain information about the current rate of exchange between the U.S. dollar and the British pound. We can observe only the rates of exchange between the U.S. dollar and the Japanese yen and between the pound and the yen.

Here the relevant model is based on the Law of One Price. This law holds that, in a competitive market, if two assets are functionally perfect

7. While the futures price "reflects" information about the future spot price, it is not necessarily an unbiased estimate of the future spot price, even in theory. For a discussion of the relation between futures prices and expected future spot prices of commodities, see Siegel and Siegel (1990).
Table 6-1  Matrix of Quality of Information Extraction

<table>
<thead>
<tr>
<th></th>
<th>Model Is Correct</th>
<th>Model Is Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prices Are Correct</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Prices Are Incorrect</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

substitutes, then they will have the same price. The Law of One Price is enforced by a trading process called *arbitrage*—the purchase and immediate sale of equivalent assets in order to earn a sure profit from a difference in their prices. When it is applied to currency exchange rates, the Law of One Price implies that, for any three currencies that are freely convertible in competitive markets, it is enough to observe the exchange rates between any two in order to infer a reliable value for the third.

In our case, suppose the dollar price of the yen is 1 cent per yen (or equivalently 100¥ to the dollar), and the price of the yen in terms of pounds is a half pence (£0.005) to the yen (or equivalently, ¥200 to the pound). Then, by the Law of One Price, one can infer that the dollar-pound exchange rate is $2 per pound (in the absence of taxes and transactions costs).

The reasoning behind this inference is as follows. There are two ways to buy pounds for dollars. One is *indirectly* through the yen market—by first buying yen for dollars, and then using the yen to buy pounds. Because one pound costs ¥200, and ¥200 costs $2.00, this indirect way costs $2.00 per pound. Another way is to purchase pounds *directly* with dollars. The direct purchase of pounds with dollars must cost the same as the indirect purchase of pounds with dollars because of the Law of One Price. If this equivalence is violated, there will be an arbitrage opportunity that cannot persist for very long.8

Suppose you believe the markets for all three currencies satisfy the conditions for being considered competitive, i.e., there are no impediments to the purchase or sale of any quantities of any of the three cur-

---

8. To see how the force of arbitrage works to uphold the Law of One Price in this example, let's look at what would happen if the price of the pound were $2.10 rather than $2.00. Suppose you walk into a bank in New York City, and you observe three exchange rates: 1 cent per yen, ¥200 per pound, and $2.10 per pound. Suppose there is one window for exchanging dollars and yen, another for exchanging yen and pounds, and a third window for exchanging dollars and pounds. Here is how you can make a $10 profit without leaving the confines of the bank:

1. At the dollar/yen window, convert $200 into ¥20,000.
2. At the yen/pound window, convert the ¥20,000 into £100.
3. At the dollar/pound window, convert the £100 into $210.

202  THE GLOBAL FINANCIAL SYSTEM
rencies. You are therefore very confident that the no-arbitrage opportunities model is correct and that the Law of One Price holds.

Now suppose you can also directly observe the dollar/pound market, and the exchange rate is $2.10 per pound. Then you must conclude either that the prices you observe are incorrect (i.e., cell 2 in Table 6-1), or that the model is incorrect (i.e., cell 3). There are two senses in which observed exchange rates can be incorrect. The first is that they are reported with error. That is, they all may not have been observed at the same time of day, or they are an average of the bid and asked prices rather than actual transaction prices. The second sense is that the data are genuinely synchronous transaction prices, but that the market is mispricing the three currencies.

The model may be incorrect because of some barrier to the operation of the force of arbitrage. For example, there may be government exchange controls or some other form of government intervention in one of the three markets. Yet, if you are truly confident that the model is correct and that the observed prices are genuine transaction prices, then you must conclude that an arbitrage opportunity exists. If you and others similarly informed proceed to exploit that opportunity, your trading activity will over time eliminate the arbitrage opportunity. Thus, such a situation is unstable and transitory, and eventually we will find ourselves in cell 1 of Table 6-1.

Of course, an analyst is almost never absolutely sure about both the validity of the model and of the observed prices. Each is believed with some degree of confidence that is less than 100%. Nevertheless, the underlying principle of inference illustrated here remains applicable. No information extraction is possible without some model.

A broad set of examples will clarify the process of information extraction and its uses.

**Pure Default-Free Cash Flows**

There is a form of "hierarchy" in the extraction of information from asset prices—a progression from the simple to the complex. We have already given an example of the simplest kind of extraction of information: computing the implied exchange rate between two currencies when the exchange rate cannot be observed directly. An example that raises the order of complexity is extracting information about the prices of pure default-free future cash flows, or, equivalently, extraction of the pure default-free term structure of interest rates.

To illustrate, suppose we have a pattern of cash flows denominated in U.S. dollars. The data used for valuing these flows are the prices of U.S.
Treasury bonds. As in the case of currencies, the basic model follows from the Law of One Price. Thus, we assume that inferences drawn from the prices of U.S. Treasury bonds are valid for all dollar-denominated cash flows that are free of default risk.

Let us first restrict the data to the prices of Treasury strips, which represent the value of a single cash payment on a specified maturity date. The objective is to use these Treasury strip prices to extract information about the prices for all possible maturity dates. Because there is a continuum of possible maturity dates, and strips are available for only a discrete number of maturity dates, additional assumptions are needed in order to infer the missing prices. One common approach is to assume that pure discount prices can be represented as a polynomial function of time to maturity. The model is calibrated by fitting its parameters to the observed prices. The estimated parameters are then used either to interpolate between observed maturity dates or to extrapolate beyond their range.

Even this relatively simple procedure of information extraction can give rise to both a difficulty and, perhaps, an opportunity: There may be more than one reported price for strips with the same maturity date. Different market prices for the same promised future cash flows can be the result of differences among bonds in liquidity, tax treatment, or other features such as call provisions. There are three ways to reconcile these differences:

- Explicitly incorporate the features that account for pricing differences by developing a more complex pricing model.
- Leave the features unspecified, assume they give rise to an “error term” that is uncorrelated with time to maturity, and use econometric methods to infer the hypothetical “pure” prices. This approach takes advantage of a powerful set of tools of statistical inference. Chief among these is the technique of multiple regression analysis.
- Conclude that there is a genuine arbitrage opportunity due to mispricing in the market for U.S. Treasury bonds and seek to exploit it.

These avenues to reconciliation of price differences are not mutually exclusive. The analyst typically pursues all three.

Nontraded Financial Assets

In many countries including the United States, regulatory bodies such as the Federal Reserve Board, rule-making bodies such as the Financial Accounting Standards Board, and guaranty agencies such as the Federal Deposit Insurance Corporation need information about the market values of the assets and liabilities of financial intermediaries. Such information
is essential in monitoring the capital adequacy of banks, insurance companies, and pension funds.

The market prices of securities that are similar to the financial assets held by the intermediary can provide such information. Thus the prices of low-credit-quality "junk" bonds provide useful information for valuing the commercial loan portfolios of banks or the privately placed bond portfolios of insurance companies. This is sometimes called "pricing by analogy" or "matrix pricing."

As in the valuation of pure default-free cash flows, rarely does one find a traded security that is an exact analogue of the nontraded one. It is therefore necessary to resort to formal methods of inference such as pricing models to translate observed market prices into the desired valuations.

**Residential Real Estate**

It is a common practice in the United States that local governments levy taxes on the current market value of residential property. Because the vast majority of properties do not trade frequently, the typical practice is to infer their market values using a model calibrated to the observed prices of properties recently sold in arm's length transactions.

Models of real estate appraisal for tax purposes vary widely in their details. Generally, they invoke the Law of One Price and use multiple regression techniques. Observed transaction prices are regressed against a set of observable features of each property that should matter in valuation, such as the acreage and building size. A set of coefficients relating value to these features are estimated. The market values of all properties are then inferred from their observable features by applying the estimated coefficients from the regression model.

**Credit Guarantees**

A next-level increase of complexity in the hierarchy of information extraction provides that the output of one stage of extraction becomes the input in a subsequent stage. For example, suppose a firm is considering guaranteeing bonds issued by a certain corporation against default, and wants an estimate of the market value of such a guarantee.

The two-stage procedure first derives the prices of pure default-free cash flows from the prices of U.S. Treasury strips using the approach discussed above. Stage two then uses those prices to compute what corporate bonds would sell for if they were free of default risk. The difference
between that computed default-free value and the actual market price of the corporate bonds is the *implied value* of the guarantee.⁹

**Implied Tax Brackets**

In many countries including the United States, income tax rates are progressive or "graduated." Private investors and tax authorities alike are interested in knowing the marginal tax bracket at which it becomes advantageous for investors to switch from holding taxable to tax-exempt securities. Market interest rates (computed from bond prices) provide one means of computing this implied "switch-over" tax bracket.

Implied tax brackets are useful in identifying and separating the effects of taxes from other factors that influence bond prices and yields. For example, the yield spread between corporate bonds and U.S. Treasury bonds can be attributed to three sources: liquidity differences, default risk, and the partially tax-exempt status of the government bonds.¹⁰ The implied effect of this tax exemption can be extracted from the spread between high-grade corporates or U.S. government agency bonds and Treasury securities.

**A Caveat**

Even if information is perfectly and accurately reflected in asset prices, it may not be possible to extract that information from those prices. (In the language of mathematics, the "inverse function" may not exist.) To illustrate this point, suppose that the price of XYZ stock always accurately reflects investor expectations of XYZ’s expected future earnings per share, E, discounted at a market capitalization rate, k. Thus \( P = E/k \). Yet, neither E nor k is directly observable. One cannot extract the separate values of E and k from the observed price, \( P \), without additional information. Establishing the conditions under which specific information can or cannot be extracted from prices is a particular instance of the general *identification problem* of statistics and econometrics.¹¹

As a further demonstration of this central issue in information extraction, consider two typical examples.

*Example 1: Extracting estimates of the expected future value of a stock index*

---

⁹. Merton (1990) develops this procedure in detail and derives a set of guarantee values for low-credit-quality corporate bonds.

¹⁰. The interest earned on U.S. Treasury bonds is exempt from state and local income taxes. This exemption does not apply to debt issues of U.S. government agencies.

¹¹. For a survey article on the identification problem in econometrics, see Hsiao (1983).
from the current futures price on that index. It can be tempting to believe that the futures price of a stock index is an indicator of its expected future spot price, and therefore that the creation of a financial futures market can provide information about investor expectations that is not extractable from spot market prices. However, the futures price of the index is related to the spot price through the arbitrage relation:

$$F = (1 + r - d)S$$

where \(F\) is the futures price, \(r\) is the risk-free interest rate, \(d\) is the dividend yield on the index portfolio, and \(S\) is the spot price of the index.

This relation does not produce a "perfect" arbitrage opportunity unless we know the dividend yield, \(d\), with certainty. Moreover, the expected future spot price of the index does not appear at all in this relation, and therefore no information about it can be extracted from the futures price that was not already extractable from the spot price.  

Example 2: Inferring the expected direction of change in a stock's price from the relative prices of puts and calls on the stock. It can be tempting to believe that the ratio of the price of a call to the price of a put on a particular stock can serve as an indicator of investor expectations about future stock price changes. Since a call pays off only if the stock price rises above its exercise price and a put only pays off if the stock price falls below its exercise, it would seem that the call–put ratio should be higher when investors are relatively "bullish" and lower when they are "bearish." Yet, no such inference is warranted.

The prices of puts and calls are related to the price of the underlying stock through an arbitrage relation that does not depend on the expected future price of the stock. In fact, the only incremental information reliably extractable from option prices is information about the volatility of the stock price, not about its expected return.

12. The relation between the spot and futures prices and the risk-free interest rate can be used to estimate the dividend yield as:

$$d = r - \left( \frac{F}{S} - 1 \right)$$

13. Indeed, the futures price is almost certainly a downward-biased estimate of the future spot value of the index, because the stock market offers a risk premium to investors.  

14. Sprengle (1961) provides an early attempt at a model to extract information about expected returns and investor risk preferences from option prices. Later research on option pricing shows that such inferences cannot be reliably identified [Merton (1992, p. 282)].
What Is Volatility, and Why Is It Important?

Volatility is a measure of the uncertainty about future changes in an asset’s price or its rate of return. Volatility is related to the range of possible rates of return from holding the security and to their likelihood of occurring. A security’s volatility is greater, the wider the range of possible outcomes, and the higher the probabilities of the returns at the extremes of the range. The most common statistic used in finance to measure the volatility of a security’s probability distribution of returns is standard deviation—σ. It is an indicator of deviations of per period rates of return from the expected rate of return weighted by their probabilities of occurring.\(^{15}\)

Volatility is a fundamental variable in financial decision-making. Estimation of risk is essential for virtually every type of financial decision: whether capital investment and financing, asset allocation, buying or selling insurance and guarantees, or valuing risky debt and other corporate liabilities. In the absence of uncertainty, investment and financing decisions would simplify to a comparison of the present values of future cash flows discounted at observed interest rates. Portfolio selection rules would reduce to the simple imperative: Maximize the investor’s return. There would be no need for insurance, and derivative security valuation would be a trivial matter.

To document the wide range of uses for volatility estimates in financial decision-making, we briefly describe applications in three broad categories: (1) financial services, (2) corporate finance, and (3) regulation.

Uses of Volatility Estimates in the Financial Services Industry

Portfolio selection. For more than forty years, starting with the pioneering work of Markowitz (1952), the portfolio optimization models used by professional investment managers have focused on estimating the efficient portfolio frontier.\(^{16}\) The central idea is to find, for any given level of risk, the portfolio that maximizes the investor’s expected rate of return.

\(^{15}\) Standard deviation is computed as follows:

\[
\sigma = \sqrt{\sum_{i=1}^{n} P_i (r_i - E(r))^2}
\]

16. Markowitz received the 1990 Nobel Prize in Economics for this work.

208  THE GLOBAL FINANCIAL SYSTEM
These optimization models require return distribution information for all of the risky assets to be included in the investment portfolio. Specifically:

- Expected rates of return, which are derived from statistical analysis of historical data on asset returns and analyst assessments of future scenarios.\(^\text{17}\)
- Estimates of rate of return volatilities including both variances and covariances, which until recently have been derived almost exclusively from historical analysis of past returns data.

Valuation. Valuation is the lifeblood of firms engaged in investment banking activities. Firms that underwrite and sell corporate securities or arrange mergers and divestitures must have reliable valuation models if they are to remain competitive. Often the securities and other corporate liabilities that have to be evaluated are not traded in markets. Examples are executive stock options and unfunded pension liabilities. The modern technology for pricing these securities and contracts is contingent claims analysis, which requires estimates of the volatilities of the underlying economic variables.\(^\text{18}\)

Risk management. Firms that provide risk management services create customized contracts tailored to the unique circumstances of their clients. Examples are swaps, caps, collars, and floors contingent on interest rates, exchange rates, and commodity prices.\(^\text{19}\) The models that financial intermediaries use in the production and pricing of these customized derivative securities require volatility estimates as inputs.

Uses of Volatility Estimates in Corporate Financial Management

Contingent claims analysis has been applied in capital budgeting decisions to explicitly recognize the value of flexibility in investment decisions.\(^\text{20}\) For example, suppose a public utility is going to build a new power plant. It can choose between a less expensive, single-fuel plant that uses either only oil or only natural gas, or a more expensive, multi-fuel plant that can switch back and forth between oil and gas. In general, the value of fuel input type flexibility increases with increases in the uncertainty about

\(^{17}\) For a textbook exposition of how this is typically done, see, for example, Bodie, Kane, and Marcus (1995, pp. 894–900).

\(^{18}\) For development of contingent claims analysis and examples of its applications in security pricing, see Merton (1992, Chapters 11–13).

\(^{19}\) For a textbook introduction to these derivatives, see, for example, Hull (1993).

\(^{20}\) For a survey of the uses of contingent claims analysis in capital budgeting, see Merton (1992, pp. 425–427) and Dixit and Pindyck (1994).
the future relative prices of those fuels. Thus a critical consideration in
the evaluation of these alternatives for a power plant is the volatility of
the prices of gas and oil.

Financial contracting is an alternative that can substitute for such physi-
cal investments in plant and equipment that increase the firm’s flexibility.
If the utility wants the economic benefit of the ability to switch between
two different types of fuel, it can enter into option-like contractual agree-
ments that provide for cash compensation to the firm whenever the price
of fuel it actually uses in its specialized plant exceeds the price of the
alternative fuel. Similarly, a crude oil producer can create a “synthetic
refinery,” by entering into a cracking-spread contract for the firm to
deliver crude oil and receive in return refined products such as gasoline
and heating oil.

Uses of Volatility Estimates by Government Regulators

Financial guarantees. Governments all over the world provide finan-
cial guarantees either through specific guarantee programs such as de-
posit insurance or by providing credit directly. In the United States, for
example, the federal government is the nation’s largest underwriter of
default risk. The U.S. Office of Management and Budget (OMB) estimates
that three-fifths of all nonfederal credit outstanding in 1992 was assisted
by federal credit programs, government-sponsored enterprises (such as
the Federal National Mortgage Corporation), or deposit insurance.21 OMB
estimates the total present value of the costs of federal guarantees in force
in 1992 at between $203 and $294 billion. Measurement and management
of the government’s exposure through these various guarantee programs
requires estimates of volatilities of the underlying economic variables.
[See Merton and Bodie (1992).]

Monetary and fiscal policy. The volatility of certain economic indica-
tors such as the consumer price level, the growth rate of gross domestic
product, or the unemployment rate is increasingly a target as well as an
indicator of macroeconomic policy. For example, the long-standing issue
of the trade-off between inflation and unemployment has been recast in
terms of a trade-off between the variabilities of these two economic indi-
cators.22 In the United States, the Federal Reserve takes into account the
volatility of stocks, bonds, currencies, and commodities in establishing
monetary policy [See Nasar (1992)].

22. For example, see Taylor (1994).
Implied Volatility

Implied volatility is an estimate of $\sigma$ that is extracted from the prices of options and other derivative securities with nonlinear payoff structures. The "watershed" date for implied volatility is 1973. In that year, two events occurred:

- Formal publication of the Black and Scholes option pricing model.
- Creation of an organized options exchange in the United States—the Chicago Board Options Exchange (CBOE).23

The basic insight underlying the Black–Scholes model is that a dynamic portfolio trading strategy in the stock can be found that will replicate the returns from an option on that stock. Hence, to avoid arbitrage opportunities, the option price must always equal the value of this replicating portfolio.

The Black–Scholes model characterizes a family of theoretical models of option pricing in which the only unobservable is $\sigma$, and the option pricing formula is monotonic in $\sigma$. Hence a unique solution for $\sigma$ can be found that equates the market price of the option to the model's price.

The Black–Scholes formula for the price of a European call option on a nondividend-paying stock is:

\[
C = N(d_1)S - N(d_2)Ke^{-rT}
\]

\[
d_1 = \frac{\ln(S/E) + (r - d + \sigma^2/2)T}{\sigma\sqrt{T}}
\]

\[
d_2 = d_1 - \sigma\sqrt{T}
\]

where:

- $C$ = price of the call
- $N(d)$ = the probability that a random draw from a standard normal distribution will be less than $d$
- $S$ = price of the stock
- $E$ = exercise price

23. Prior to establishment of the CBOE, options had been traded over the counter. The CBOE began trading the first listed options in the United States in April 1973, a month before official publication of the Black–Scholes model. By 1975, traders on the CBOE were using the model both to price and hedge their options positions. For a description of the story behind the development of the model and its impact on finance practice, see Bernstein (1992, Chapter 11).
\( e \) = the base of the natural log function (approximately 2.71828)
\( r \) = risk-free interest rate (the annualized continuously compounded rate on a safe asset with the same maturity as the option)
\( T \) = time to maturity of the option in years
\( \ln \) = natural logarithm
\( \sigma \) = standard deviation of the annualized continuously compounded rate of return on the stock

The Black–Scholes formula has five parameters, four of them are directly observable: \( S \), the price of the stock; \( E \), the exercise price; \( r \), the risk-free interest rate (the annualized continuously compounded rate on a safe asset with the same maturity as the option); and \( T \), the time to maturity of the option. If we know the market price of an option, then the Black–Scholes formula can be used to derive a value for the unobservable \( \sigma \) that equates the price computed according to the Black–Scholes formula to the actual observed market price. This value of \( \sigma \) is called the *option's implied volatility* for the stock.\(^{24}\)

The Black–Scholes model assumes that \( \sigma \) is a constant. Merton (1973) generalizes the model to allow for a time-varying, nonstochastic \( \sigma \). Under this generalization, the implied volatility that is extracted from the Black–Scholes formula is the *average* \( \sigma \) over the life of the option.

Black and Scholes also assume that no dividends are paid during the life of the option. Merton (1973) generalizes the model to allow for a constant continuous dividend yield, \( d \).\(^{25}\) Merton’s dividend-adjusted formula is:

\[
C = N(d_1)Se^{-dT} - N(d_2)Ke^{-rT}
\]

\[
d_1 = \frac{\ln(S/E) + (r - d + \sigma^2/2)T}{\sigma \sqrt{T}}
\]

\[
d_2 = d_1 - \sigma \sqrt{T}
\]


25. Merton’s formula reduces to the Black–Scholes formula when \( d = 0 \). Whaley (1993) uses the cash-dividend-adjusted, Cox-Ross-Rubinstein (1979) binomial method to derive implied volatilities for the CBOE’s index of implied volatility.
In the real world, both $\sigma$ and $d$ are rarely known with certainty and empirical evidence suggests that both vary stochastically over time. Models that incorporate these stochastic variations have been developed and are used in practice. Nonetheless, the implied volatility that is extracted from the dividend-adjusted Black-Scholes model can still serve as a reliable indicator (even if a biased one) of investor expectations for volatility over the life of the option.

As is evident by inspecting the Black-Scholes formula, investors who disagree about the expected return on an underlying security, but have the same estimate of $\sigma$, will agree on the "fair value" of the option. Therefore, option prices can reflect consensus views about $\sigma$ even if there is no consensus about the expected return of the underlying stock.

Of course, there will not be a consensus among all investors about $\sigma$. Rather, the $\sigma$ that is embodied in option prices is a weighted average of the $\sigma$ used either explicitly or implicitly by option traders. Estimates of $\sigma$ therefore form a distribution of values around the "true" value. If the observed market price of the option reflects an unbiased estimate of the Black-Scholes model price with stochastic volatility, then, because the Black-Scholes formula is nonlinear in $\sigma$, the implied value of $\sigma$ derived from it will not generally be unbiased.

This bias can be minimized by selecting options with exercise prices equal to the forward price of the stock (that is, $E = S e^{rT}$). For such options, a linear approximation can be used for the values of both a European put and a European call on a nondividend-paying stock:26

$$C = P \approx \frac{1}{\sqrt{2\pi}} S \sigma \sqrt{T}$$

By inverting this equation, one can derive the implied volatility (i.e., $\sigma$) as a linear function of the option price and the underlying stock price:

$$\sigma = \frac{C \sqrt{2\pi}}{S \sqrt{T}}$$

The significance of this approximate linear relation is that implied volatility estimates derived from options whose exercise price equals the forward price will be nearly unbiased in a statistical sense.

The CBOE has constructed an implied volatility index (VIX) for the S&P 100 stock price index to be used as the basis for creating new futures and

---

options contracts on implied volatility itself [see Whaley (1993, pp. 80-82)]. This implied volatility index is designed to minimize statistical bias while using information from the prices of eight S&P 100 index options with a maturity of approximately 30 days.

Figure 6-1 shows the value of the VIX over the period 1986-1993. It is evident that the implied volatility of the stock price index has fluctuated quite a bit. There was a huge "spike" at the time of the October 1987 stock market crash. Thereafter, implied volatility trended downward slowly as investors regained confidence in the future prospects for stocks. By the end of 1993, the level of the VIX was below 15%.

It is also evident from Figure 6-1 that implied volatility is affected by current political and economic developments.

Implied Volatility Versus Estimates from Historical Data

Instead of extracting volatility estimates from option prices, we can estimate volatility using the time series of past price changes. It has been shown [Merton (1980)] that, in estimating \( \sigma \) from past price data, the accuracy of the estimate increases with the number of observations for a given overall observation period. Therefore for a given calendar period (e.g., the last ten years), the estimates of volatility can be improved by moving from monthly to weekly or from weekly to daily data. In sharp contrast, the accuracy of the estimate of expected return increases only with the length of the observation period and not with the number of observations over a given period.

Estimates of \( \sigma \) derived exclusively from historical price data are by their nature unconditional forecasts. Such historical estimates thus do not exploit other publicly available information that would help forecast future volatility. For example, consider the period just before the outbreak of the Persian Gulf War in 1991. The United Nations on November 29, 1990, gave Iraq a deadline for removing Iraqi forces from Kuwait—January 15, 1991. Suppose at that time you wanted to estimate the volatility of stock or bond prices in the period between November 29, 1990, and January 15, 1991. An estimate of \( \sigma \) derived exclusively from past data clearly could not reflect the additional degree of uncertainty regarding future security prices engendered by uncertainty about the impending war.

Note that information about the possibility of war in the Persian Gulf was available to anyone who watched television news or read a newspaper. It was not insider information, and hence, was reflected in the prices of stocks and options. But an estimate of \( \sigma \) based exclusively on past data would not have picked up this information. One based on option prices (i.e., implied volatility) would have.
Jan. 15, 1991: War breaks out

Aug. 1990: Iraq invades Kuwait

Nov. 29, 1990: Iraq is given a deadline to remove its forces from Kuwait

CBOE MARKET VOLATILITY INDEX

Fig. 6-1  Implied Volatility: 1986-1993
Another advantage of implied volatility over historical volatility stems from the nonstationarity of volatility. There is substantial empirical evidence that the volatilities of security prices, interest rates, exchange rates, and commodity prices are not constant over time.\textsuperscript{27} Theoretical research on speculative prices also supports the hypothesis that volatility generally changes over time. There have been, and are sure to be in the future, periods of greater or lesser degrees of uncertainty regarding economic conditions, and those different degrees of uncertainty will be reflected in the volatility of financial markets. Volatility should not in general be modeled as a constant.

Nonstationarity complicates estimation of volatility from past data using standard statistical techniques. If volatility is "slowly" changing, one way to reduce the problem of nonstationarity of the volatility parameter is to use shorter time intervals (more frequent sampling) to achieve the same level of accuracy in estimating $\sigma$ for a shorter overall observation period. Sophisticated econometric techniques, which allow for changing volatility (such as GARCH models), have had some success. As Black and Scholes (1972) conclude in the first empirical evaluation of their model, the best approach is to combine the information provided by implied volatility and historical volatility.\textsuperscript{28}

\textbf{The Term Structure of Implied Volatilities}

The availability of government bond prices of different maturities makes it possible to derive an accurate term structure of interest rates. Just so, the availability of prices on multiple options with different maturities allows one to infer the term structure of volatility. In the exchange-traded options markets, there are different maturities for every strike price of an option. Implied volatilities can be computed from the prices of these different maturity options to yield a term structure of volatilities. The implied-volatility term structure is generally not flat, nor even monotonic.\textsuperscript{29}

For example, Table 6-2 shows the term structure of volatility derived from options on the S&P 100 stock index (OEX) on December 3, 1990. Note that options maturing before the January 15, 1991, deadline for Iraq to leave Kuwait have an implied annualized volatility of 18.68\%.\textsuperscript{30} The

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
Maturity (days) & Implied Volatility & Reference \hline
10 & 19.25 & \cite{28} \hline
20 & 18.68 & \cite{29} \hline
30 & 18.68 & \cite{30} \hline
\end{tabular}
\caption{Implied Volatility Term Structure}
\end{table}

27. See the substantial literature on the ARCH and GARCH models for estimating dynamic changes in volatility. Engle, Kane, and Noh (1993) and Nelson (1991) are examples.
28. Brenner and Galai (1989) concur when they propose a standardized volatility index that is a weighted average of implied and historical volatility measures.
29. See, for example, Heynan, Kenna, and Vorst (1992).
30. The prospect of a "jump" distribution for nonlocal changes in stock prices when the
Table 6-2  Term Structure of Volatility on December 3, 1990

<table>
<thead>
<tr>
<th>Contract and Maturity Date</th>
<th>Trading Days to Maturity</th>
<th>Average Implied Volatility</th>
<th>Incremental Implied Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec 1990 Call</td>
<td>15</td>
<td>$ 6.625</td>
<td>18.68%</td>
</tr>
<tr>
<td>January 1991 Call</td>
<td>35</td>
<td>$11.250</td>
<td>20.45%</td>
</tr>
<tr>
<td>February 1991 Put</td>
<td>55</td>
<td>$12.000</td>
<td>25.66%</td>
</tr>
</tbody>
</table>

Note: Exercise price of all options is 305; index closing value is 305.73; T-bill rate is 7.02% per year; number of assumed trading days in the year is 250. Historical annual $\sigma$ estimated from daily data for the 30 days prior to December 3, 1990, was 17.38%.

implied volatility for options maturing after the deadline, in February 1991, is 25.66%. The "forward" or incremental annualized volatility over the added 30 days amounts to 34.38%. This sharp increase reflects the uncertainty associated with the war, which would likely be resolved only after the U.N. deadline, since negotiated solutions are rarely reached prior to such deadlines.

**Implied Covariances**

As in the case of default-free fixed-income securities, there is a "hierarchy" in the extraction of information from option prices—a progression from the simple to the complex. Thus, one can use a multi-step procedure to extract information not only about the volatilities (standard deviations) of economic variables, but also about their covariances with other economic variables.

For example, consider the simplest case of an option on a portfolio of two stocks. By definition, the variance of the rate of return on the portfolio is given by the formula:

$$\sigma^2 = w_A^2 \sigma_A^2 + w_B^2 \sigma_B^2 + 2w_Aw_B\sigma_{AB}$$

where $w_A$ is the constant portfolio fraction in stock A, $w_B$ is the fraction in stock B, and $\sigma_{AB}$ is the covariance between the returns on stocks A and B.

From the prices of options on stock A and on stock B separately, one can infer the implied values of $\sigma_A$ and $\sigma_B$. That is the first step. The price

uncertainty of war is resolved on the deadline date requires a modified version of the Black-Scholes model [see Merton (1976)]. Implied volatilities reported in Table 6-2, however, are computed using the standard formula.
of an option on a portfolio of stocks A and B can similarly be used to infer the implied variance of the portfolio, \( \sigma^2 \). Given the estimates for \( \sigma^2 \), \( \sigma_A \), and \( \sigma_B \), then \( \sigma_{AB} \) is estimated by the formula:

\[
\sigma_{AB} = \frac{\sigma^2 - w^2 \sigma_A^2 - w^2 \sigma_B^2}{2w_A w_B}
\]

Other examples of contingent claims models that permit similar extraction of information about covariances from derivative security prices can be found in Merton (1973), Margrabe (1978), Fischer (1978), and Stulz (1982).

**Pricing Models for Fixed-Income Derivatives**

During the 1980s and 1990s a host of new fixed-income derivatives—both standardized exchange-traded instruments such as puts and calls and customized contracts such as caps, floors, and collars—were created to help firms manage their interest rate risk exposures. To price and hedge these instruments, financial services firms have adopted generalized fixed-income models that are "matched" frequently to replicate both the observed term structure of interest rates and the term structure of interest rate volatilities. The "implied" parameters of the models are "fitted" statically to the observed term structures in the same fashion that implied volatility is calculated in the Black-Scholes model. In practice, the models are recalibrated daily using the prices of publicly traded U.S. Treasury securities and standardized options on bond futures and swaps.

**Applications of Implied Volatility**

The prices of options on individual securities and on portfolios of securities can be used to infer the variance-covariance structure of returns on risky assets. These parameters can then be entered as inputs into Markowitz-type portfolio optimization models. Similarly, in corporate finance, one can also infer the parameters needed to make complex financial decisions from the prices of traded securities issued by the firm. As one example, consider the case of valuing executive stock options.

There has been considerable controversy in the United States over the recommendation that public corporations be required to report the value

---

31. Among the earliest stochastic interest rate models are ones by Vasicek (1977), and Cox, Ingersoll, and Ross (1985). More recently, the "matching" models are developed in Ho and Lee (1986), Black, Derman, and Toy (1990), and Heath, Jarrow, and Morton (1992). See Hull (1993, Chapter 15) for a survey description of all these models.
of stock options granted to employees as part of their compensation [see Financial Accounting Standards Board (1993)]. Since a market price for these options is not available, much of the debate surrounding the recommendation centers on the firm’s ability to value them accurately. The exposure draft recommends two alternative valuation techniques that use option pricing models and volatilities estimated from historical price data.

Another alternative is to use contingent claims analysis and implied volatilities. To illustrate the procedure, suppose the ABC Corporation is a publicly held company. A contingent claims analysis starts by specifying the terms of each of the security types issued by the firm. Suppose that the securities issued by ABC Corporation are: (1) bonds maturing in one year with a face value of $100 million and a current market price of $85 million, (2) 1 million shares of common stock with a market price of $20 million, and (3) 100,000 employee stock options expiring in one year with an exercise price of $20 per share (and no market price). Thus, the options entitle their holders to buy 1/11 of the firm’s value a year from now for a total price of $2 million. The risk-free interest rate \( r \) is 6% per year compounded continuously.

Consider the possible payoff pattern to the holders of ABC’s debt at its maturity date one year from now. If the future value of ABC’s assets exceeds the face value of its debt (i.e., if \( V_1 > $100 \) million), the bondholders receive the promised payment of $100 million. If the value of the assets falls short of $100 million, however, the firm will default on the debt, and the bondholders will receive the firm’s assets.

Now consider the payoffs to the optionholders. If the value of the firm’s assets is less than $100 million a year from now, then neither the stockholders nor the optionholders get anything—the bondholders get it all. If the value of the firm’s assets exceeds $100 million, the optionholders have the right to buy a fraction (1/11) of the firm’s assets for $2 million. If they choose to exercise that option, they will have 1/11 of \( (V_1 - \$100 \text{ million} + \$2 \text{ million}) \). The value of \( V_1 \) at which it would just pay to exercise the option is therefore:

\[
\$2 \text{ million} = (V_1 - \$100 \text{ million} + \$2 \text{ million})/11 \\
V_1 = \$120 \text{ million}
\]

Knowing the market value of the debt ($85 million) and the stock ($20 million), one can derive implied estimates both for the value of the entire firm (including the nontraded options) and for the future volatility of that firm value. As with standard implied volatility estimates, the market prices of the traded securities are assumed to satisfy the pricing formulas.
The equation for the market value of the bonds satisfies:

\[ 85 = [1 - N(d_1)]V + N(d_2)100e^{-rT} \]

\[ d_1 = \frac{\ln \left( \frac{V}{100} \right) + \left( r + \frac{\sigma^2}{2} \right)T}{\sigma\sqrt{T}} \]

\[ d_2 = d_1 - \sigma\sqrt{T} \]

where \(N(\cdot)\) is the cumulative normal density function; \(V\) is the (implicit) value of the firm as a whole (i.e., the combined total value of the bonds, the stocks, and the options), and \(\sigma\) is the volatility of this total combined value.

The equation for the market value of the stock satisfies:

\[ 20 = -N(d_3)V - 100N(d_4)e^{-rT} - \frac{1}{11}[N(d_3)V - 120N(d_4)e^{-rT}] \]

\[ d_1 = \frac{\ln \left( \frac{V}{120} \right) + \left( r + \frac{\sigma^2}{2} \right)T}{\sigma\sqrt{T}} \]

\[ d_2 = d_1 - \sigma\sqrt{T} \]

\[ d_3 = \frac{\ln \left( \frac{V}{120} \right) + \left( r + \frac{\sigma^2}{2} \right)T}{\sigma\sqrt{T}} \]

\[ d_4 = d_3 - \sigma\sqrt{T} \]

Solving these two equations simultaneously for \(V\) and \(\sigma\), we get \(V = \$106.14\) million and \(\sigma = 0.3621\). Therefore, the market value of the employee stock options is:

\[
V - \text{Value of the Bonds} - \text{Value of the Stock} = \\
\$106.14 \text{ million} - \$85 \text{ million} - \$20 \text{ million} = \$1.14 \text{ million}
\]

or $11.40 per option.

---

32. The contingent claims pricing formula for the debt is derived in Black and Scholes (1973) and Merton (1973, 1974, 1992).

33. The solution is found through an iterative procedure using the Gauss computer software program.
Information Extraction in the Future

As we have seen, the more complete and diverse the set of financial instruments traded in financial markets, the more information that can be extracted from their prices. Information extraction is an integral part of the financial-innovation spiral discussed in Chapter 1. The pace of financial innovation to manage risks, which accelerated in the 1970s and 1980s, is likely to continue into the future. As products such as futures, options, swaps, and securitized loans become standardized, and move from intermediaries to markets, the proliferation of new trading markets in those instruments facilitates the creation of new custom-designed financial products that improve "market completeness."

To hedge their exposures on those products, the producers (typically, financial intermediaries) trade in these new markets and volume expands. Increased volume reduces marginal transactions costs and thereby makes possible implementation of more new products and trading strategies by intermediaries, which in turn leads to still more volume. The success of these trading markets and custom products encourages investment in creating additional markets and products—so on it goes, spiraling toward the theoretically limiting case of zero marginal transactions costs and dynamically complete markets.

As new markets are created, they may not at first be informationally efficient. All the while, however, incentives are created for market participants to develop new or improved pricing models to exploit the profit opportunities created by any mispricing. Better models for valuing traded financial assets lead to more reliable information extraction, which in turn leads to better calibration of models for the pricing of new financial products. This richer information set will facilitate more efficient resource allocation decisions.

References


