ROBERT E. WHALEY*

DERIVATIVES

Abstract

The area of derivatives is arguably the most fascinating area within financial economics during the past thirty years. This chapter reviews the evolution of derivatives contract markets and derivatives research over the past thirty years. The chapter has six complementary sections. The first contains a brief history of contract markets. The most important innovations occurred in the 1970s and 1980s, when contracts written on financial contracts were introduced. Concurrent with these important industry innovations was the development of modern-day option valuation theory, which is reviewed in the second and third sections. The key contribution is seminal theoretical framework of the Black-Scholes (1973) and Merton (1973) ("BSM") model. The key economic insight of their model is that a risk-free hedge can be formed between a derivatives contract and its underlying asset. This implies that contract valuation is possible under the assumption of risk-neutrality without loss of generality.

The final three sections summarize the three main strands of empirical work in the derivatives area. In the first group are studies that focus on testing no-arbitrage pricing relations that link the prices of derivatives contracts with their underling asset and with each other. The second group contains studies that evaluate option empirical performance of option valuation models. The approaches used include investigating the in-sample properties of option values by examining pricing errors or patterns in implied volatilities, examining the performance of different option valuation models by simulating a trading strategy based on under- and over-pricing, and examining the informational content of the volatility implied by option prices. The final group focuses on the social costs and/or benefits that arise from derivatives trading. The main conclusion that can be drawn from the empirical work is that the BSM model is one of the most resilient in the history of financial economics.

Latest revision: August 27, 2002

Keywords: derivatives, options forwards, futures, swaps
JEL codes: G100, G120, G130, G140

*T. Austin Finch Foundation Professor of Business Administration, Fuqua School of Business, Duke University, Durham, NC 27706, E-mail: whaley@duke.edu. Prepared for the *Handbook of the Economics of Finance*, edited by George Constantinides, Milton Harris, and Rene Stulz. Comments and suggestions by Nick Bollen, George Constantinides, Jeff Fleming, Tom Smith, Rene Stulz, and Seth Wechsler are gratefully acknowledged.

DERIVATIVES

Arguably the most fascinating area within financial economics during the past thirty years is derivatives. With virtually no derivatives contracts written on financial assets at the beginning of the 1970s, the industry has grown to a level exceeding $100 trillion. This growth would not have been possible without the powerful theoretical contributions of Black-Scholes (1973) and Merton (1973). Their concept of forming a risk-free hedge between a derivatives contract and its underlying asset serves as the foundation for valuing an enormous array of different contract structures.

The purpose of this chapter is to provide an overview of the key contributions to the derivatives literature over the past thirty years. This review has six complementary sections. The first section contains a brief history of derivatives contracts and contract markets. Although the origin of derivatives use dates back thousands of years, the most important innovations occurred only recently, in the 1970s and 1980s. Not coincidently, these two decades are also the most important in terms of theoretical developments in the derivatives literature.

The second section is the first of two that focus on derivative contract valuation. The key assumption in the development of the valuation results presented in this section is the law of one price—two perfect substitutes will have the same price in a rationally-functioning marketplace. Under this seemingly innocuous assumption, a myriad of pricing relations can be developed for derivatives contracts including forwards, futures options and swaps.

The third section focuses particularly on the valuation of contingent claims. To value such claims, it is necessary to know the character of the asset price distribution at time(s) in the future as well as the appropriate discount rate to apply in bringing the expected future cash flows of the derivatives contract (which, of course, depend on the asset price distribution) back to the present. It is this style of claim that underlies the seminal theoretical framework of the Black-Scholes (1973) and Merton (1973) ("BSM") model. The key economic insight of their model is that a risk-free hedge can be formed between an option and its underlying asset. This implies that option valuation is possible without knowing investor risk preferences, hence the risk-free rate of interest can be used as the appropriate discount rate to apply to the expected future cash flows. To develop the expectations of future cash flows, BSM assume that the underlying asset price follows geometric Brownian motion with constant volatility. Among other things, this implies that, at any point in time in the future, the asset price will be log-normally distributed, eliminating the prospect of negative asset prices that had plagued earlier work. Not only did this framework provide BSM with the ability to value standard call and put options, it has provided other researchers with the ability to value thousands of differently structured agreements including caps, collars, floors, binary options, and quantos. Many of these contributions, as well as other extensions to the BSM model, are summarized.

The fourth through sixth sections of this chapter summarize empirical work that investigates the pricing and valuation of derivatives contracts and the efficiency of the markets within which they trade. The studies are divided into three groups. In the first group are studies that focus on testing no-arbitrage pricing conditions. These are contained in section 4. A review of tests of the no-arbitrage price relations between forwards and futures and their underlying assets as well as tests lower price bounds and put-call parity in the options markets is provided.

The second group contains studies that attempt to evaluate option empirical performance of option valuation models. Approaches differ. Some investigate the in-sample properties of option values by examining pricing errors or patterns in implied volatilities. Others examine the performance of different option valuation models by simulating a trading strategy based on under- and over-pricing. Yet others examine the informational content of the volatility implied by option prices. Discussions of each approach of study are included in Section 5.

The third and final group of studies focuses on the social costs and/or benefits that arise from derivatives trading. One sub-group examines whether the introduction of derivatives trading disrupts the market for the underlying asset by generating abnormal price movements and/or increased volatility. A second sub-group examines whether the expiration of derivatives groups disrupts the underlying asset market. A final sub-group examines the inter-temporal relation of price movements in the derivatives and asset markets to ascertain, among other things, where private information is being traded first. All of these discussions are contained in Section 6.

The final section contains a brief summary.

1. BACKGROUND

Derivatives, while seemingly new, have been used for thousands of years. In his treatise, *Politics*,[1] Aristotle tells the story of Thales, a philosopher (and reasonably good meteorologist). Based on studying the winter sky, Thales predicted an unusually large olive harvest. He was so confident of his prediction that he bought rights to rent all of the olive presses in the region for the following year. The fall arrived, and the harvest was unusually plentiful. The demand and price for the use of olive presses soared.

Thales' call options were early examples of *over-the-counter* (OTC) derivatives. OTC derivatives are private contracts negotiated between parties. Thales bought, and the olive press owners sold, call options. The prices of the options were negotiated, and Thales paid for them in the form of cash deposits. The chief advantage of OTC derivatives markets is limitless flexibility in contract design. The underlying asset can be anything, the size of the contract can be any amount, and the delivery can be made at any time and at any location. The only requirement of an OTC contract is a willing buyer and seller.

Among the disadvantages of OTC markets, however, is that willing buyers and sellers must spend time identifying each other. Thousands of years ago, before the advent of high-speed communication and computer technology, such searches were costly. Consequently, *centralized markets* evolved. The Romans organized commodity markets with specific locations and fixed times for trading. Medieval fairs in England and France during the 12th and 13th centuries served the same purpose. While centralized commodity markets were originally developed to facilitate immediate cash transactions, the practice of contracting for future delivery (i.e., forward transactions) was also introduced.

[1] See *Politics* by Aristotle (350 BC, Book 1, Part XI).

Another disadvantage of OTC derivatives is *credit risk*, that is, the risk that a counterparty will renege on his contractual obligation. Perhaps the most colorful example of this type of risk involves forward and option contracts on tulip bulbs. In what can be characterized as a *speculative bubble*, rare and beautiful tulips became collectors' items for the upper class in Holland in the early 17th century. Prices soared to incredible levels.[2] Homes, jewels, livestock—nothing was too precious that it could not be sacrificed for the purchase of tulips. In an attempt to cash-in on this craze, it was not uncommon for tulip bulb dealers to sell bulbs for future delivery. They did so based on call options provided by tulip bulb growers. In this way, if bulb prices rose significantly prior to delivery, the dealers would simply exercise their options and acquire the bulbs to be delivered on the forward commitments at a fixed (lower) price. The tulip bulb growers also engaged in risk management by buying put options from the dealers. In this way, if prices fell, the growers could exercise their puts and sell their bulbs at a price higher than that prevailing in the market. In retrospect, both the tulip bulb dealers and growers were managing the risk of their positions quite sensibly.

Everything could have worked out fine, except that the bubble burst in the winter of 1637 when a gathering of bulb merchants could not get the usual inflated prices for their bulbs. Panic ensued. Prices sank to levels of 1/100th of what they had once been. This set off an unfortunate chain of events. Individuals who had agreed to buy bulbs from dealers did not do so. Consequently, dealers did not have the cash necessary to buy the bulbs when the growers attempted to exercise their puts. Some legal attempts were made to enforce the contracts, however, the attempts were unsuccessful. These contract defaults left an indelible mark on OTC derivatives trading.

By the 1800s, the pendulum had swung from undisciplined derivatives trading in OTC markets toward more structured trading on organized exchanges. The first derivatives exchange in the U.S. was the Chicago Board of Trade (CBT). While the CBT was originally formed in 1848 as a centralized marketplace for exchanging grain, forward contracts were also negotiated. The earliest recorded forward contract trade was made on March 13, 1851 and called for 3,000 bushels of corn to be delivered in June at a price of

[2] Garber (2000) provides a detailed recount of tulip bulb price levels during this period.

one cent per bushel below the March 13th spot price.[3] Forward contracts had their drawbacks, however. They were not standardized according to quality or delivery time. In addition, like in the case of the tulip bulb fiasco, merchants and traders often did not fulfill their forward commitments.

In 1865, the CBT made three important changes to the structure of their grain trading market. First, they introduced the use of standardized contracts called *futures contracts*. Unlike forwards contracts in which the parties are free to choose the terms of the contract, the terms of futures contracts are set by the exchange and are standardized with respect to quality, quantity, and time and place of delivery for the underlying commodity. By concentrating hedging and speculative demands on fewer contracts, the depth and liquidity of the market are enhanced. This facilitates position unwinding. If a party to a trade wants to exit his position prior to the delivery date of the contract, he need only execute an opposite trade (i.e., *reverse* his trade) in the same contract. There is no need to seek out the counterparty of the original trade and attempt to negotiate the contract's termination.

The second and third changes were made in an effort to promote *market integrity*. The second was the introduction of a *clearinghouse* to stand between the buyer and the seller and guarantee the performance of each party. This crucial step eliminated the *counterparty risk* that had plagued OTC trading. In the event a buyer defaults, the clearinghouse "makes good" on the seller's position, and then holds the buyer's clearing firm liable for the consequences. The buyer's clearing firm, in turn, passes the liability onto the buyer's broker, and ultimately the buyer. Note that, at any point in time, the clearinghouse has no net position since there are as many long contracts outstanding as there are short. The third was the introduction of a *margining system*. When the buyer and seller enter a futures position, they are both required to deposit good-faith collateral designed to show that they can fulfill the terms of the contract.

From the late 1800s through the early 1980s, the majority of derivatives trading took place on exchanges. Futures contracts were the dominant contract design, and agricultural commodities were the dominant underlying asset. The list of contracts that

[3] See Chicago Board of Trade (1994, ch.1, p.14).

have been active for more than a century include the CBT's corn, oats, and wheat futures launched in 1865, the New York Cotton Exchange's cotton futures launched in 1870, the Chicago Produce Exchange[4] (which later became the Chicago Mercantile Exchange) was formed by a group of agricultural dealers to trade futures on butter and egg futures launched in 1874, and the Coffee Exchange's coffee futures launched in 1882.[5]

The move to non-agricultural commodities was slow. Indeed, 51 years elapsed before the Commodity Exchange (COMEX) in New York was formed to trade the first metals contract—silver futures. The New York Mercantile Exchange (NYMEX) followed with platinum futures in December 1956 and palladium futures in January 1968. The introduction of futures on livestock occurred in the 1960s. The Chicago Mercantile Exchange (CME) launched pork belly futures in September 1961, live cattle futures in November 1964, and live hog futures in February 1966. Futures contracts on energy products did not emerge until November 1978, at which time the NYMEX introduced the heating oil futures contract.

The pace of innovation in derivatives markets increased remarkably in the 1970s. Many of the important events occurring during this decade, as well as the next, are summarized in Table 1. The first major innovation occurred in February 1972, when the CME began trading futures on currencies in its International Monetary Market (IMM) division. This marked the first time a futures contract was written on anything other than a physical commodity. The second was in April 1973, when the CBT formed the Chicago Board Options Exchange (CBOE) to trade options on common stocks.[6] This marked the first time an option was traded on an exchange. The third major innovation occurred in October 1975, when the CBT introduced the first futures contract on an interest rate instrument—Government National Mortgage Association futures. In January 1976, the

---

[4] The Chicago Produce Exchange also traded futures on other perishable commodities. In 1898, the butter and egg dealers withdrew to form their own market, the Chicago Butter and Egg Board. In 1919, it was reorganized to trade other commodity futures and was renamed the Chicago Mercantile Exchange.
[5] In 1914, the Coffee Exchange expanded to include sugar futures, and, in 1916, it changed its name to the New York Coffee and Sugar Exchange. In 1979, it merged with the New York Cocoa Exchange to form today's Coffee, Sugar & Cocoa Exchange.
[6] Initially, only *call* options were listed in the U.S. Put option trading were not listed until June 1977, and, even then, only on an experimental basis.

CME launched Treasury bill futures and, in August 1977, the CBT launched Treasury bond futures.

The 1980s brought yet another round of important innovations. The first was the use of cash settlement. In December 1981, the IMM launched the first cash settlement contracts, the 3-month Eurodollar futures. At expiration, the Eurodollar futures is settled in cash based on the interest rate prevailing for a three-month Eurodollar time deposit.[7] Cash settlement made feasible the introduction of derivatives on stock index futures, the second major innovation of the 1980s. In February 1982, the Kansas City Board of Trade (KCBT) listed futures on the Value Line Composite stock index, and, in April 1982, the CME listed futures on the S&P 500. These contract introductions marked the first time that futures contracts were written on stock indexes. The third major innovation of the 1980s was the introduction of exchange-traded option contracts written on "underlyings"[8] other than individual common stocks.[9] The CBOE and AMEX listed interest rate options in October 1982 and the Philadelphia Stock Exchange (PHLX) listed currency options in December 1982. In the same year, options on futures appeared for the first time. In October 1982, the CBT began to list Treasury bond futures options, and the Coffee, Sugar, and Cocoa Exchange (CSCE) began to list options on sugar and gold futures. In January 1983, the CME and the New York Futures Exchange (NYFE) began to list options directly on stock index futures, and, in March 1983, the CBOE began to list options on stock indexes.

These two decades of innovation have transformed the nature of derivatives trading activity on exchanges. While derivatives exchanges were originally developed to help market participants manage the price risk of physical commodities, today's trading activity is focused on hedging the financial risks associated with unanticipated price movements in stocks, bonds, and currencies.

---

[7] A Eurodollar time deposit is a U.S. dollar deposit in a London bank, and the interest rate quoted on such deposits is called the London Interbank Offer Rate (i.e., the *LIBOR* rate). Since different banks may offer different rates on deposits of the same maturity, the settlement rate is based on an average of rates across banks.

[8] From this point forward, the term "underlying" refers to the asset or instrument that underlies the derivative contract.

[9] For a comprehensive review of these new option introductions and their economic purposes, see Stoll and Whaley (1985).

The 1980s also saw the re-emergence of OTC derivatives trading. As derivatives on financial assets became increasingly popular, investment banks began to think of new ways to tailor contracts to meet customer needs. Some innovations were minor changes in the standard terms of exchange-traded derivatives contracts on financial instruments (e.g., modifications to the expiration date and/or the contract denomination). In 1980, for example, the first OTC Treasury bond option was traded. Other contracts were new and seemingly different. They fall under the generic heading of "swaps." A *swap contract* is a contract to "swap" a series of periodic future cash flows, where the terms of the swap are usually set such that the up-front payment is zero. The first *interest rate swap* was in 1981, when the Student Loan Marketing Association (i.e., "Sallie Mae") swapped interest payments on intermediate-term fixed rate debt for floating-rate payments indexed to the three-month Treasury bill rate. The cash flows of the two legs of a swap can be linked to virtually any asset or index. A *basis rate swap*, for example, is an exchange of floating rate payments where the two floating rates are linked to, say, a three-month Treasury bill rate and a three-month Eurodollar time deposit rate. A *currency swap* is an exchange of interest payments (either fixed or floating) in one currency for payments (either fixed or floating) in another. An *equity swap* involves the exchange of an interest rate payment and a payment based on the performance of a stock index, while an *equity basis swap* involves an exchange of payments on two different indexes. Swap agreements may appear different from standard forward and option contracts, but they are not. Every swap can be decomposed into a portfolio of forwards and options. The benefit a swap provides is that several transactions are bundled into a single product.

2. NO-ARBITRAGE PRICING RELATIONS

A great deal can be learned about valuing derivatives under minimal assumptions. The sole necessary assumption is the *law of one price* (LOP). Stated simply, the LOP says that two perfect substitutes <u>must</u> have the same price. If they do not, a *costless arbitrage profit* can be earned by simultaneously buying the cheaper asset and selling the more expensive one.[10] Because the same asset is bought and sold simultaneously, the position is risk-free. This is the key attribute of an *arbitrage* strategy.[11] The fact that the strategy involves no initial cash outlay makes it *costless*. The absence of costless arbitrage opportunities is fundamental in derivatives contract valuation.

A second assumption is that markets are frictionless. Frictionless markets have a number of attributes including:

a) No trading costs.

b) No differential tax rates.

c) Unlimited borrowing and lending at the risk-free rate of interest.

d) Freedom to sell (short), with full use of any proceeds.

e) Can trade at any time and in any quantity.

The *frictionless market* assumption is made largely for convenience. By ignoring market frictions, pricing relations can more easily be identified. In most cases, the impact of considerations such as trading costs, taxes, and divergent borrowing and lending rates can be and have been introduced into the valuation framework straightforwardly. Indeed, the very presence of these market restrictions has caused many derivatives markets to thrive. The focus of this section is to describe some important no-arbitrage relations for derivative contract prices.

2.1 Carrying costs

Derivative contracts are written on four types of assets—stocks, bonds, foreign currencies and commodities. The derivatives literature contains seemingly independent developments of derivative valuation principles for each asset category. Generally speaking, however, the valuation principles are not asset-specific. The only distinction among assets is how carry costs are modeled.[12]

---

[10] This "law of one price" argument is the fundamental theoretical underpinning to the Nobel prize-winning corporate finance theory of Modigliani/Miller (1958) and Miller/Modigliani (1961).

[11] The term, arbitrage, is frequently misapplied. *Risk arbitrage*, for example, refers to a trading strategy in which the shares of a firm rumored to be on the verge of being acquired are purchased and the shares of the acquiring firm are simultaneously purchased. Since the merger may or may not take place and the stock prices may change, this activity is not arbitrage.

[12] See, for example, Stoll and Whaley (1986).

The *cost of carry* refers to the difference between the costs and the benefits that accrue while holding an asset. Suppose a breakfast cereal producer needs 5,000 bushels of wheat for processing in two months. To lock in the price of the wheat today, he can buy it and carry it for two months. One cost of this strategy is the opportunity cost of funds. To come up with the purchase price, he must either borrow money or reduce his earning assets by that amount. Beyond interest cost, however, carry costs vary depending upon the nature of the asset. For a *physical asset* such as wheat, he incurs storage costs (e.g., rent and insurance). At the same time, by storing wheat, he avoids the costs of possibly running out of his regular inventory before two months are up and having to pay extra for emergency deliveries. This benefit is called *convenience yield.* Thus, the cost of carry for a physical asset equals interest cost plus storage costs less convenience yield, that is,

$$\text{Carry costs} = \text{Cost of funds} + \text{storage cost} - \text{convenience yield}. \tag{1a}$$

For a *financial asset* such as a stock or a bond, storage costs are negligible. Moreover, income (yield) accrues in the form of quarterly cash dividends or semi-annual coupon payments. The cost of carry for a financial asset is

$$\text{Carry costs} = \text{Cost of funds} - \text{income}. \tag{1b}$$

Carry costs and benefits are modeled either as continuous rates or as discrete flows. Some costs/benefits such as the cost of funds (i.e., the risk-free interest rate) are best modeled continuously. The dividend yield on a broadly-based stock portfolio and the interest income on a foreign currency deposit also fall into this category. Other costs/benefits like quarterly cash dividends on individual common stocks, semi-annual coupons on bonds, and warehouse rent payments for holding an inventory of grain are best modeled as discrete cash flows. In the interest of brevity, only continuous costs are considered here.[13]

Dividend income from holding a broadly-based stock index portfolio or interest income from holding a foreign currency is typically modeled as a constant, continuous

[13] For a detailed discussion of the ways in which carrying costs can be modeled, see Whaley (2002).

rate.[14] The income, as it accrues, is re-invested in more units of the asset. In this way, buying e^{-iT} units of a stock index portfolio today grows to exactly one unit at time *T,* and produces a net terminal value of $\tilde{S}_T - Se^{(r-i)T}$. The cost of carry rate equals the difference between the risk-free rate of interest *r* and the dividend yield rate *i* for a stock index portfolio investment, and equals the difference between the domestic interest rate *r* and the foreign interest rate *i* for a foreign currency investment. The total cost of carry paid at time *T* is

$$\text{Carry costs} = S[e^{(r-i)T} - 1]. \quad (2)$$

2.2 Valuing forward/futures using the no-arbitrage principle

The *value of a forward contract* is inextricably linked to the cost of carry of the underlying asset. Since a forward contract requires its buyer to accept delivery of the underlying asset at time *T*, buying a forward contract today is a perfect substitute for buying the asset today and carrying it until time *T*. The present value of the payment obligation under the forward contract strategy is fe^{-rT}, and the present value of the latter strategy is Se^{-iT}. Since both strategies provide exactly one unit of the asset at time *T*, (i.e., $\tilde{S}_T$), their costs must be identical,

$$fe^{-rT} = Se^{-iT}. \quad (3a)$$

If the relation (3a) does not hold, costless arbitrage profits would be possible by selling the over-priced instrument and simultaneously buying the under-priced one. The relation (3a) is the present value version of the *cost of carry relation.* A more familiar version is the future value form,

$$f = Se^{(r-i)T}. \quad (3b)$$

When the prices of the forward and the asset are such that (3a) and/or (3b) hold exactly, the forward market is said to be *at full carry*. Unless costless arbitrage is somehow impeded, the forward market will always be at full carry. The difference between the forward (or futures) price and the asset price is frequently referred to as the *basis*.

[14] The carry cost rates within this framework are deterministic. In the short-run, this assumption is

Futures contracts are like forward contracts, except that price movements are *marked-to-market* each day rather than receiving a single, once-and-for-all settlement on the contract's expiration day.[15] Obviously, the sum of the daily mark-to-market price moves over the life of the futures equals the overall price movement of a forward with the same maturity. With the futures position, however, the mark-to-market profits (losses) are invested (carried) at the risk-free interest rate until the futures expires. The value of the futures position at time *T*, therefore, may be greater or less than the terminal value of the forward position, depending on the path that futures price follows over the life of the contract.

Cox, Ingersoll and Ross (1981) (hereafter "CIR") and Jarrow and Oldfield (1981), among others, use no-arbitrage arguments to show the equivalence of forward and futures prices when interest rates are deterministic. To illustrate their argument, assume that the term structure of interest rates is flat and does not change through time. Also, assume that r is the continuously compounded interest rate on a daily basis. Now, consider a "rollover" futures position that begins, on day 0, with $e^{-r(T-1)}$ futures contracts and that increases the number of futures each day by a factor e^{r}. At the end of day 1, the position is marked-to-market, generating proceeds of $e^{-r(T-1)}(\widetilde{F}_1 - F)$. Assuming this gain/loss is carried forward until day *T*, the terminal gain/loss will be $e^{-r(T-1)}(\widetilde{F}_1 - F)e^{r(T-1)} = \widetilde{F}_1 - F$. For day 2, the position is increased by a factor e^{r} and is marked-to-market at $e^{-r(T-2)}(\widetilde{F}_2 - \widetilde{F}_1)$, generating proceeds of $e^{-r(T-2)}(\widetilde{F}_2 - \widetilde{F}_1)e^{r(T-2)} = \widetilde{F}_2 - \widetilde{F}_1$ on day *T*, and so on. Because the number of futures is chosen to exactly offset the accumulated interest factor on the daily mark-to-market gain/loss, the rollover futures position has exactly the same terminal value as the long forward position. Under the no-arbitrage assumption, the *valuation equation for a futures contract* is the same as that of the forward, that is,

$$F = f = Se^{(r-i)T}. \qquad (4)$$

CIR also use no-arbitrage arguments to show the relation between forward and futures prices when interest rates are stochastic. They find that the futures price will be

---

reasonable for most type of assets.

15 Recall that, in Section 1, we discussed the historical fact that the CBT changed from making markets in forward contracts to making markets in futures contracts in 1865 as a means of ensuring market integrity.

less (greater) than the forward price if (a) the price changes of the futures contract and the default-free discount bond are positively (negatively) correlated and/or (b) the variance of bond price changes is less than (exceeds) the covariance between spot price changes and bond price changes. They also show that if the covariance between spot price changes and bond price changes is positive, the futures price is less than the forward price, and the futures-forward price difference is a decreasing function of the expected forward-bond covariance.

2.3. Valuing options using the no-arbitrage principle

The no-arbitrage pricing results for options come in two primary forms—lower price bounds and put-call parity conditions. Each is discussed in turn.

2.3.1 Call options

The *lower price bound of a European-style call option* is

$$c \geq \max(0, Se^{-iT} - Xe^{-rT}), \tag{5}$$

where c is the price of a European-style call with exercise price X and time to expiration T. The price of the call must be greater than or equal to zero since it is a privilege. The reason the call price must exceed $Se^{-iT} - Xe^{-rT}$ is based on the following no-arbitrage argument. Suppose a portfolio is formed by selling e^{-iT} units of the underlying asset and buying a European-style call. To ensure that enough cash is on hand to exercise the call at expiration, Xe^{-rT} in risk-free securities are also purchased. At time T, the net value of the portfolio depends on whether the asset price is above or below the exercise price. If the asset price is below the exercise price, the call expires worthless. The risk-free securities (plus accrued interest) are used to buy a unit of the asset to cover the short sale obligation. A cash balance of $X - \tilde{S}_T$ remains. If the asset price is greater than the exercise price on day T, the call will be exercised. This requires a cash payment of X, which can be made exactly using the risk-free securities. The unit of the asset received upon exercising the call is used to retire the short sale obligation. Thus, if $S_T > X$, the net terminal value of the portfolio is certain to be 0. Considering both possible outcomes, this portfolio is *certain* to have a net terminal value of at least 0. This means that its initial value must be less than or equal to 0, otherwise a costless arbitrage opportunity would exist. With $Se^{-iT} - Xe^{-rT} - c \le 0$, the lower price bound or a European-style call is $c \ge Se^{-iT} - Xe^{-rT}$.

In general, the lower price bound of an option is called its *intrinsic value*, and the difference between the option's market price and its intrinsic value is called its *time value*. A European-style call has an intrinsic value of $\max(0, Se^{-iT} - Xe^{-rT})$ and a time value of $c - \max(0, Se^{-iT} - Xe^{-rT})$. No-arbitrage principle identifies the intrinsic value of an option. The determinants of time value are the focus of Section 3.

American-style options are like European-style options except that they can be exercised at any time up to and including the expiration day. Since this additional right cannot have a negative value, the relation between the prices of American-style and European-style call options is

$$C \ge c, \tag{6}$$

where the upper case C represents the price of an American-style call option with the same exercise price and time to expiration and on the same underlying asset as the European-style call. The *lower price bound of an American-style call option* is

$$C \ge \max(0, Se^{-iT} - Xe^{-rT}, S - X). \tag{7}$$

This is the same as (5), except that the term $S - X$ is added within the maximum value operator on the right-hand side since the American-style call cannot sell for less than its early exercise proceeds, $S - X$. If $C < S - X$, a costless arbitrage profit of $S - X - C$ can be earned by simultaneously buying the call (and exercising it) and selling the asset.

The structure of the lower price bound of the American-style call (7) provides important insight regarding the motivation (or lack thereof) for early exercise. The second term in the parentheses, $Se^{-iT} - Xe^{-rT}$, is the minimum price at which the call can be sold in the marketplace.[16] The third term is the value of the American-style if it is

[16] To exit a long position in an American-style call option, you have three alternatives. First, you can hold it to expiration, at which time you will (a) let it expire worthless if it is out of the money or (b) exercise it if it is in the money. Second, you can exercise it immediately, receiving the difference between the current

exercised immediately. If the value of the second term is greater than the third term (for a certain set of call options), the call's market price will always be greater than its exercise proceeds and it will never be optimal to exercise early.

To identify this set of calls, examine the conditions under which the relation

$$Se^{-iT} - Xe^{-rT} > S - X$$

or

$$S\left(e^{-iT} - 1\right) > -X\left(1 - e^{-rT}\right) \tag{8}$$

holds. Since the risk-free interest rate is positive, the expression on the right-hand side is negative. Hence, if the left-hand side is positive or zero, early exercise will never be optimal. This condition is met in cases in which $i \le 0$. If $i \le 0$, an American-style call will never optimally be exercised early, and the value of the American-style call is equal to the value of the European-style call, $C = c$. Merton (1973) was the first to identify this result and refers to the situation as the call being "worth more alive than dead."

The intuition underlying the "worth more alive than dead" result can be broken down into two components—interest cost, *r*, and non-interest cost, i. Holding other factors constant, a call option holder prefers to defer exercise. Immediate exercise requires a cash payment of *X* today. On the other hand, if exercise is deferred until the call's expiration, the cash is allowed to earn interest. The present value of the exercise cost is only Xe^{-rT}. With respect to non-interest cost, recall that $i < 0$ for physical assets that require storage. If a call on such an asset is exercised early, the asset is received immediately and storage costs begin to accrue. On the other hand, if exercise is deferred by continuing to hold the claim on the asset rather than the asset itself, storage costs are avoided. Note that, even if storage costs are zero (i.e., with $i = 0$), condition (8) holds because the interest cost incentive remains.

For American-style call options on assets with $i > 0$ (e.g., stock index portfolio paying dividend yield and foreign currencies paying foreign interest), early exercise <u>may</u> be optimal. The intuition is that, while there remains the incentive to defer exercise and

---

asset price and the exercise price. Third, you can sell it in the same marketplace. There is, after all, an earn interest on the exercise price, deferring exercise means forfeiting the income being generated on the underlying asset. The only way to capture this income is by exercising the call and taking delivery of the asset. For American-style call options on assets with $i > 0$, early exercise may be optimal and, therefore, $C > c$.

2.3.2 Put options

The *lower price bound of a European-style put option* is

$$p \geq \max(0, Xe^{-rT} - Se^{-iT}), \tag{9}$$

where p is the price of a put with exercise price X and time to expiration T. The reason that the put price must exceed $Xe^{-rT} - Se^{-iT}$ is based on a no-arbitrage portfolio involving a long position in the put, a long position of e^{-iT} units of the asset, and a short position of Xe^{-rT} in risk-free securities. If the asset price is less than or equal to the exercise price at the option's expiration, the put will be exercised. The cash proceeds from exercise are used to cover the risk-free borrowing. If the asset price is greater than the exercise price, the put expires worthless, and the asset is sold to cover the risk-free borrowing, leaving $\tilde{S}_T - X$ in cash. Since the net terminal value of the portfolio is always greater than or equal to zero, its present value must be less than or equal to zero.

An American-style put has an early exercise privilege, which means that the relation between the prices of American-style and European-style put options is

$$P \geq p, \tag{10}$$

where the upper case P represents the price of an American-style put option with the same exercise price, time to expiration and underlying asset as the European-style put. The *lower price bound of an American-style put option* is

$$p \geq \max(0, Xe^{-rT} - Se^{-iT}, X - S). \tag{11}$$

This is the same as (9), except that $X - S$ is added within the maximum value operator. If $P < X - S$, a costless arbitrage profit of $X - S - P$ can be earned by simultaneously buying the put (and exercising it) and buying the asset.

active secondary market for standard calls and puts.

In the case of an American-style call, early exercise is never optimal if the asset's income rate is less than or equal to zero (i.e., $i \leq 0$). In the case of an American-style put, no comparable condition exists;[17] there is always a possibility of early exercise depending on the value of *S*. To see this, suppose the asset price falls to 0. The put option holder will exercise *immediately* since (a) there is no chance that the asset price will fall further, and (b) deferring exercise means forfeiting the interest income that can be earned on the exercise proceeds. An American-style put is always worth more than the European-style put, $P > p$.

2.3.3 Put-call parity

Put-call parity uses trades in the call, the put, and the asset simultaneously to create a risk-free portfolio. *Put-call parity for European-style options* is given by

$$c - p = Se^{-iT} - Xe^{-rT}, \tag{12}$$

where the call and the put have the same exercise price and time to expiration, and are written on the same underlying asset. The pricing relation is driven by a no-arbitrage argument. In this case, the no-arbitrage portfolio consists of buying e^{-iT} units of the asset, buying a put, selling a call with the same exercise price, and borrowing Xe^{-rT}. It is straightforward to show that this portfolio will be worthless when the options expire at time *T* regardless of the relation between the asset price and the option's exercise price. Since no one would pay a positive amount to hold such a portfolio (or a portfolio with reverse investments), the put-call parity relation (12) must hold.

The set of trades used to derive put-call parity is called a *conversion*. If all of the trades are reversed (i.e., sell the asset, sell the put, buy the call, and buy risk-free securities), it is called a *reverse conversion*. These names arise from the fact that you can create any position in the asset, options, or risk-free securities by trading (or *converting*) the remaining securities. The concept of conversion/reverse conversion arbitrage was introduced into the academic literature about 30 years ago.[18] Some market participants were well aware of the concept decades earlier, however. Russell Sage, one of the great

[17] In the expression on the right-hand side of (11), the third term is greater than the second term over some range for *S*, independent of the level of *i*.

U.S. railroad speculators of the 1800s, used conversions to circumvent usury laws. Sage extended credit to individuals under three conditions: (a) they post collateral in the form of stock (with the loan amount capped at the current stock price, *S*), (b) they provide a written guarantee that Sage could sell back the stock at *S*, and (c) they pay a cash premium to Sage for the right to buy the stock (when the loan is repaid) at *S*. Ignoring the cash premium, the borrower has received an interest-free loan, borrowing *S* and then repaying *S*. In reality, however, the loan is anything but interest free. The cost of the call embeds the interest cost. Conveniently, the usury laws did not apply to implicit interest rates.

The early exercise feature of American-style options complicates the put-call parity relation. The specification of the relation depends on the non-interest carry cost, *i*. The *American-style put-call parity relations* are

$$S - X \le C - P \le Se^{-iT} - Xe^{-rT} \quad \text{if } i \le 0 \tag{13a}$$

and

$$Se^{-iT} - X \le C - P \le S - Xe^{-rT} \quad \text{if } i > 0 \ . \tag{13b}$$

Each inequality in (13a) and in (13b) has a separate set of no-arbitrage trades. Proofs are provided in Stoll and Whaley (1986).

2.3.4 Summary

The purpose of this section was to show some of the derivatives pricing relations that can be developed under the seemingly innocuous assumption that two perfect substitutes must have the same price. Some of these relations will be used in the next section to gather intuition about the specification of option valuation formulas. The relations also serve as the basis for the empirical investigations discussed in Section 4.

3. OPTION VALUATION

Valuing claims to uncertain income streams is one of the central problems in finance. The exercise is straightforward conceptually. First, the amount and the timing of

---

[18] The appearance of put-call parity in the academic literature is in Stoll (1969).

the expected cash flows from holding the claim must be identified. Next, the expected cash flows must be discounted to the present. The valuation of a European-style call option, therefore, requires the estimation of (a) the mean of the call option's payoff distribution on the day it expires, and (b) the risk-adjusted discount rate to apply to the option's expected terminal payoff.

In his dissertation, *Theory of Speculation*, Bachelier (1900) provides the first known valuation of the European-style call option. His valuation equation, which may be written

$$c = \int_X^\infty (S - X) f(S) dS, \tag{14}$$

shows that the option's value depends its expected terminal value. Bachelier assumes that the underlying asset price follows arithmetic Brownian motion,[19] which means $f(S)$ is a normal density function. Unfortunately, this assumption implies that asset prices can be negative.[20]

To circumvent this problem, Sprenkle (1961) and Samuelson (1965) value the call under the assumption that the asset price follows geometric Brownian motion. By letting asset prices have multiplicative, rather than additive, fluctuations through time, the asset price distribution at the option's expiration is lognormal, rather than normal, and the prospect of the asset price becoming negative is eliminated. Under lognormality, Sprenkle and Samuelson show that the call option valuation formula has the form,

$$c = e^{-\alpha_c T}[Se^{\alpha_S T} N(d_1) - XN(d_2)], \tag{15}$$

where

$$d_1 = \frac{\ln(S/X) + (\alpha_S + .5\sigma^2)T}{\sigma\sqrt{T}}, d_2 = d_1 - \sigma\sqrt{T},$$

[19] Many consider Bachelier to be the father of modern option pricing theory. For an interesting recount of Bachelier's life and his insights into option valuation, see Sullivan and Weithers (1991).

[20] Bachelier also assumes that the asset's expected price change is zero. This implies investors are risk-neutral and money has no time value.

α_S and α_c are the expected risk-adjusted rates of price appreciation for the asset and the call respectively, σ is the asset's volatility rate, S is the current asset price, X is the option's exercise price, and T is the option's time to expiration. The expression, $N(.)$, is the cumulative univariate normal probability function.

The structure of (15) shows that call option value is the present value of its expected terminal value. The expected terminal value depends on a number of factors including the expected growth rate of the asset price, α_S. The call is a claim to buy the asset, and the expected asset price at the option's expiration is $Se^{\alpha_S T}$. The expression, $Se^{\alpha_S T}N(d_1)$, is the expected asset price conditional on the asset price exceeding the exercise price at the option's expiration times the probability that the option will be exercised. The expression, $XN(d_2)$, is the expected exercise cost (i.e., the exercise price) times the probability that the option will be exercised.

As simple and elegant as formula (15) appears, it is not very useful. To implement the formula requires estimates of the risk-adjusted rates of price appreciation for both the asset and the option. The estimation of these values is difficult. In the case of the call, estimation is particularly troublesome because its return depends on the asset's rate of price appreciation as well as the passage of time.

3.1 The Black-Scholes/Merton option valuation theory

The breakthrough came in the early 1970s when Black/Scholes (1973) and Merton (1973) proved that a risk-free hedge could be formed between an option and its underlying asset. The intuition underlying their argument can be illustrated using a simple one-period binomial framework. Consider a European call option that allows its holder to buy one unit of an asset in one month at an exercise price of \$40. For the sake of simplicity, suppose that the current asset price is also \$40 and that, at the end one month, the asset price will be either \$45 or \$35. Now, consider selling call options against the unit investment in the asset. At expiration, each call will have a value of \$5 or \$0, depending on whether the asset price is \$45 or \$35. Under this scenario, selling two call options against each unit of the asset will create a terminal portfolio value of \$35, regardless of the level of asset price. Since the terminal portfolio value is certain, the value of the portfolio today must be \$35 discounted at the risk-free rate of interest. If the simple risk-free rate of interest is one percent over the life of the option, the current value of the portfolio must be \$34.65, and the current value of the call \$2.675 (i.e., (\$40.00-34.65)/2)). If the observed price of the call is above (below) its theoretical level of \$2.675, risk-free arbitrage profits are possible by selling the call and buying (selling) a portfolio consisting of a long position in a half unit of the asset and a short position of \$17.325 in risk-free bonds. In equilibrium, no such arbitrage opportunities can exist.

The Black-Scholes/Merton (hereafter "BSM") model is the continuous-time analogue of this illustration. First, asset price movements are assumed to follow the geometric Brownian motion,

$$dS = \alpha_S S dt + \sigma S dz \,. \tag{16}$$

That is, over the next infinitesimally small interval of time dt, the change in asset price, dS, equals an expected price increment (i.e., the product of the instantaneous expected rate of change in asset price, α_S, times the current asset price, *S*, times the length of the interval) plus a random increment proportional to the instantaneous standard deviation of the rate of change in asset price, σ, times the asset price. The term, dz, denotes an increment to a Wiener process. If the asset price follows the dynamics described by (16), it can be shown by Ito's lemma that derivative contracts written on the asset have price movements described by

$$df = \left(\frac{\partial f}{\partial S} \alpha_S S + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial f}{\partial S} \sigma S dz \,, \tag{17}$$

where f is the value of the derivatives contract. Note that the underlying source of uncertainty, *dz,* in (16) and (17) is the same.

The key insight of the BSM option valuation model is that, if the derivative contract and the underlying asset share the same source of risk, it is possible to create a risk-free hedge portfolio by buying $\partial f / \partial S$ units of the asset and selling the derivative contract (or vice versa). This portfolio has an initial value of

$$V = -f + \frac{\partial f}{\partial S} S \,.$$

Over the next instant in time, the portfolio value changes in response to changes in the prices of the derivative contract and the asset, as well as a result of collecting income on the asset at the constant, continuous rate, *i*. Algebraically,

$$dV = -df + \frac{\partial f}{\partial S} dS + \frac{\partial f}{\partial S} iSdt \,.$$

Substituting (17) and (16) for df and dS,

$$\begin{aligned} dV &= -\left(\frac{\partial f}{\partial S}\alpha_S S + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2 \right) dt - \frac{\partial f}{\partial S}\sigma S dz + \frac{\partial f}{\partial S}\left(\alpha_S S dt + \sigma S dz\right) + \frac{\partial f}{\partial S} iSdt \\ &= -\left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2 - \frac{\partial f}{\partial S} iS \right) dt \end{aligned}$$

Note that by constructing the portfolio in this manner, the only source of risk, dz, has been eliminated. Since the portfolio is risk-free and perfect substitutes must have the same price, holding this portfolio is equivalent to holding an equal dollar investment in risk-free bonds, that is,

$$-\left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2 - \frac{\partial f}{\partial S} iS \right) dt = r\left(-f + \frac{\partial f}{\partial S} S \right) dt \,. \tag{18}$$

By rearranging (18), the BSM partial differential equation is identified,

$$\frac{\partial f}{\partial t} + (r-i)S\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf \,. \tag{19}$$

Equation (19) is the *Black-Scholes/Merton model*, and should not be confused with the *Black-Scholes/Merton formula*. The latter is a special case of the model to be discussed shortly. The BSM model (19) applies to *all* derivatives written on *S* including calls, puts, European-style options, American-style options, caps, floors, and collars—any derivative contract for which it is appropriate to assume the asset price dynamics follow geometric Brownian motion.[21] What distinguishes each derivative is the set of boundary equations applied to (19). For a European-style call option, the boundary

[21] This, of course, eliminates many derivatives contracts written on interest rate instruments whose underlying asset price cannot rise above a certain level (e.g., an option on a Treasury bill). In these instances, it is more common to let the underlying source of uncertainty be the short-term interest rate.

condition is $f = \max(0, S - X)$ at time *T*. For a European-style put option, the boundary condition is $f = \max(0, X - S)$ at time *T*. For American-style calls and puts, the respective boundary conditions apply at all times between the current time 0 and the expiration date *T*. Sometimes the partial differential equation subject to a boundary condition has a solution that can be expressed as an analytical formula. This is true for European-style options, for example. At other times, no analytical formula is possible and approximation methods must be used.

3.2 Analytical formulas

For many types of derivatives contracts, analytical valuation formulas are possible, with the most well-known case being European-style options. These are the focus here.

3.2.1. The infamous "Black-Scholes/Merton formula"

The BSM formula for a European-style call can be derived from equation (19). It can also be obtained by applying the Black-Scholes/Merton insight to the Sprenkle/Samuelson valuation formula (15). More specifically, if it is possible to create a risk-free hedge portfolio by buying the asset and selling the option or vice versa, the option value will not depend on an individual's risk preferences. A risk-averse individual will value a European-style call at the same level as a risk-neutral individual. Consequently, for tractability, it is convenient to assume a risk-neutral world in which all assets (including options) have an expected rate of return equal to the risk-free interest rate, *r*. That is not to say that all assets have the same expected rate of price appreciation. Some assets pay out income in the form of dividends or coupon interest. With the asset's income modeled as a constant, continuous proportion of the asset price, *i*, the expected rate of price appreciation on the asset, α_S, equals the interest rate less the cash disbursement rate, *i*, that is, $\alpha_S = r - i$. On the other hand, some assets like the call option pay out nothing through time, in which case $\alpha_c = r$. Substituting the risk-neutral levels of α_S and α_c into (15), the "Black-Scholes/Merton formula" for the value of a European-style call option becomes

$$
\begin{aligned}
c &= e^{-rT}[Se^{(r-i)T}N(d_1) - XN(d_2)] \\
&= Se^{-iT}N(d_1) - Xe^{-rT}N(d_2)
\end{aligned}
\quad . \tag{20}
$$

where $d_1 = \dfrac{\ln(S/X) + (r - i + .5\sigma^2)T}{\sigma\sqrt{T}}$ and $d_2 = d_1 - \sigma\sqrt{T}$.

Several things about (20) are noteworthy. First, neither the risk premium of the call nor the risk premium of the asset appears in the formula, so the estimation problems that arose in applying the Sprenkle/Samuelson formula are eliminated. Second, where the underlying asset's volatility rate is zero (i.e., $\sigma = 0$), the formula (20) reduces to the lower price bound (5) in Section 2. This implies that the asset's return volatility over the life of the option, $\sigma\sqrt{T}$, drives the *time value* of the option. Finally, the value of the corresponding European-style put can be easily obtained by substituting (19) into the put-call parity relation (12) from Section 2.

Before proceeding with a description of other analytical formulas, it is worthwhile to show that the BSM call option formula (20) is a solution to (19). The partial derivatives of (20) are as follows:

delta: $\dfrac{\partial f}{\partial S} = e^{-iT}N(d_1)$, gamma: $\dfrac{\partial^2 f}{\partial S^2} = \dfrac{e^{-iT}n(d_1)}{S\sigma\sqrt{T}}$, and

theta: $\dfrac{\partial f}{\partial t} = Se^{-iT}n(d_1)\dfrac{\sigma}{2\sqrt{T}} - iSe^{-iT}N(d_1) + rXe^{-rT}N(d_2)$.

Substituting these expressions, as well as the valuation equation (20), into the partial differential equation (19), the expressions on the two sides of (19) are shown to be equal,[22]

$$
\begin{aligned}
-Se^{-iT}n(d_1)\frac{\sigma}{2\sqrt{T}} + iSe^{-iT}N(d_1) - rXe^{-rT}N(d_2) + (r-i)Se^{-iT}N(d_1) + \frac{1}{2}\sigma^2 S^2 \frac{e^{-iT}n(d_1)}{S\sigma\sqrt{T}} \\
= r\left[Se^{-iT}N(d_1) - Xe^{-rT}N(d_2)\right]
\end{aligned}
$$

[22] Note that the sign of theta needs to be reversed since the option's life is growing shorter through time.

3.2.2 Special cases of the Black-Scholes/Merton formula

The BSM formula covers a wide range of underlying assets. To show its versatility, first re-write the European-style call option formula as

$$c = e^{-rT}[Se^{bT}N(d_1) - XN(d_2)], \tag{21}$$

where $d_1 = \dfrac{\ln(S/X) + (b + .5\sigma^2)T}{\sigma\sqrt{T}}$, $d_2 = d_1 - \sigma\sqrt{T}$ and b is the asset's expected risk-neutral rate of price appreciation parameter.

Non-dividend-paying stock options

The most well-known option valuation problem is that of valuing options on non-dividend-paying stocks. This is, in fact, the valuation problem addressed by Black and Scholes (1973). With no dividends paid on the underlying stock, the expected price appreciation rate of the stock equals the risk-free rate of interest, and the call option valuation equation becomes the familiar Black/Scholes formula,

$$c = SN(d_1) - Xe^{-rT}N(d_2),$$

where $d_1 = \dfrac{\ln(S/X) + (r + .5\sigma^2)T}{\sigma\sqrt{T}}$, and $d_2 = d_1 - \sigma\sqrt{T}$.

Constant-dividend-yield stock options

Merton (1973) generalizes stock option valuation by assuming that stocks pay dividends at a constant, continuous dividend yield. The "Merton model," used for valuing many options on broad-based stock indexes, is equation (21), where i is the index's dividend yield rate.

Futures options

Black (1976) values options on futures. In a risk-neutral world with constant interest rates, the expected rate of price appreciation on a futures, because it involves no cash outlay, is zero. Substituting $b = 0$ into (21) provides what is commonly known as the "Black model,"

$$c = e^{-rT}\left[FN(d_1) - XN(d_2)\right],$$

where $d_1 = \dfrac{\ln(F/X) + .5\sigma^2 T}{\sigma\sqrt{T}}$, and $d_2 = d_1 - \sigma\sqrt{T}$.

Futures-style futures options

Following the work of Black, Asay (1982) values futures-style futures options. Such options trade on a number of exchanges including the London International Financial Futures Exchange (LIFFE) and the Sydney Futures Exchange (SFE). They have the distinguishing feature that the option premium is not paid up front. Instead, the option position is marked-to-market in the same manner as the underlying futures. To value this option, the cost of carry rates for the asset and the option are both set equal to zero. In a risk-neutral world, any security whose upfront investment is zero has an expected return equal to zero. With $b = 0$ and $r = 0$, the resulting formula, called the "Asay model," is

$$c = FN(d_1) - XN(d_2),$$

where $d_1 = \dfrac{\ln(F/X) + .5\sigma^2 T}{\sigma\sqrt{T}}$, and $d_2 = d_1 - \sigma\sqrt{T}$.

Foreign currency options

Finally, Garman and Kohlhagen (1983) develop a formula to value options on foreign currency. In this case, the expected rate of price appreciation of a foreign currency equals the domestic rate of interest less the foreign rate of interest. The "Garman-Kohlhagen model" is specified exactly in the manner of the Merton model, except that r represents the domestic risk-free interest rate and i represents the foreign risk-free interest rate.

3.2.3 *Valuation by replication*

The key contribution of the BSM model is the recognition that a risk-free hedge can be formed between an option and its underlying asset. Consequently, the payoffs of a call option can be replicated with a portfolio consisting of the asset and some risk-free bonds. The BSM formula provides the composition of the asset/bond portfolio that mimics the payoffs of the call. A long call position can be replicated by buying $e^{-iT}N(d_1)$ units of the asset (each unit with price, S) and selling $N(d_2)$ units of risk-free bonds (each unit with price, Xe^{-rT}). As time passes and as the asset price moves, the units invested in the asset and risk-free bonds change. Nonetheless, with continuous rebalancing, the portfolio's payoffs will be identical to those of the call.

Dynamic portfolio insurance[23]

Dynamic replication is at the heart of one of the most popular financial products of the 1980s—dynamic portfolio insurance. Because long-term index put options were not traded at the time, stock portfolio managers had to create their own insurance by dynamically rebalancing a portfolio consisting of stocks and risk-free bonds. The mechanism for identifying the portfolio weights is given by the BSM put option formula,

$$p = Xe^{-rT}N(-d_2) - Se^{-iT}N(-d_1).$$

The objective is to create an "insured" portfolio whose payoffs mimic the portfolio, $Se^{-iT} + p$. Substituting the BSM put formula, we find

$$\begin{aligned} Se^{-iT} + p &= Se^{-iT} + Xe^{-rT}N(-d_2) - Se^{-iT}N(-d_1) \\ &= Se^{-iT}N(d_1) + Xe^{-rT}N(-d_2). \end{aligned}$$

Hence, a dynamically insured portfolio has $e^{-iT}N(d_1)$ units of stocks and $N(-d_2)$ units of risk-free bonds. The weights show that as stock prices rise, funds are transferred from bonds to stocks and vice versa.

[23] For a lucid description of portfolio insurance, see Rubinstein (1985).

Static replication

The valuation-by-replication technique can also be applied in a static context. Many multiple contingency financial products such as caps, collars, and floors (so-called "exotic" options) are valued as portfolios of standard options. Even a standard call option can be valued in this manner. Consider a portfolio that consists of (a) a long position in an asset-or-nothing call that pays the asset price at expiration if the asset price exceeds *X* and (b) a short position in a cash-or-nothing call that pays *X* if the asset price exceeds *X*.[24] Under the assumptions of risk-neutrality and lognormally distributed asset prices, the value of the asset-or-nothing call is $Se^{-iT}N(d_1)$, and the value of the cash-or-nothing call option is $Xe^{-rT}N(d_2)$. Combining these option values produces the BSM formula (20).

3.2.4 Extensions: Single underlying asset

The BSM option valuation framework has been extended in several important ways. Some involve the valuation of more complex claims on a single underlying asset. Others involve claims on two or more underlying assets. The extensions involving a single underlying asset are discussed first.

Compound options

An important extension of the BSM model that falls in the single underlying asset category is the compound option valuation theory developed by Geske (1979a). Compound options are options on options. A call on a call, for example, provides its holder with the right to buy a call on the underlying asset at some future date. Geske shows that, if these options are European-style, valuation formulas can be derived.

American-style call options on dividend-paying stocks

The Geske (1979a) compound option model has been applied in other contexts. Roll (1977), Geske (1979b), and Whaley (1981), for example, develop a formula for valuing an American-style call option on a stock with known discrete dividends. If a stock pays a cash dividend during the call's life, it may be optimal to exercise the call early, just prior to dividend payment. An American-style call on a dividend-paying stock, therefore, can be modeled as a compound option providing its holder with the right, on the ex-dividend date, either to exercise early and collect the dividend, or to leave the position open. In this application, the stock price, net of the present value of the promised dividends is assumed to follow geometric Brownian motion.[25]

Chooser options

Rubinstein (1991) uses the compound option framework to value the "chooser" or "as-you-like-it" options traded in the OTC market. The holder of a chooser option has the right to decide at some future date whether the option is a call or a put. The call and the put usually have the same exercise price and the same time remaining to expiration.

Reset options

Gray and Whaley (1997) use the compound option framework to value yet another type of contingent claim. S&P 500 bear market warrants with a periodic reset trade at the CBOE and the NYSE. The warrants are originally issued as at-the-money put options, however, they have the distinguishing feature that if the underlying index level is above the original exercise on some pre-specified future date, the exercise price of the warrant is reset at the then prevailing index level. These warrants offer an intriguing form of portfolio insurance whose floor value adjusts automatically as the index level rises. The structure of the valuation problem is again a compound option.

Lookback options

A lookback option is another exotic with only one underlying source of price uncertainty. A lookback option is an option whose exercise price is determined at the end of the option's life. For a call, the exercise price is set equal to the lowest price that the asset reached during the life of the option, and, for a put, the exercise price equals the highest asset price. These "buy at the low" and "sell at the high" options can be valued analytically. Formulas are provided in Goldman, Sosin, and Gatto (1979).

---

[24] Asset-or-nothing and cash-or-nothing options are commonly referred to as "binary" or "digital" options, and, themselves, are generally considered to be "exotics."

[25] Equivalently, the forward price of the stock is assumed to follow geometric Brownian motion.

<u>Barrier options</u>

Barrier options are the final type of option in this category to be discussed. Barrier options are options that either cease to exist or come into existence when some pre-defined asset price barrier is hit during the option's life. A down-and-out call, for example, is a call that gets "knocked out" when the asset price falls to some pre-specified level prior to the option's expiration. Rubinstein and Reiner (1991) provide valuation equations for a large family of barrier options.

3.2.5 Extensions: Multiple underlying assets

The BSM option valuation framework has also been extended to include multiple underlying assets. As long as each asset is traded, the BSM risk-free hedge argument remains intact and risk-neutral valuation is permitted without loss of generality.

<u>Exchange options</u>

The first important development along this line was by Margrabe (1978). He derives a valuation formula for an exchange option. An exchange option gives its holder the right to exchange one risky asset or asset for another. The BSM formula is a special case of the Margrabe formula in the sense that if the call is in the money at expiration the option holder exchanges risk-free bonds for the asset.

<u>Options on the minimum and the maximum</u>

Stulz (1982) and Johnson (1987) derive valuation formulas for options on the maximum and the minimum of two or more risky assets. Many of the exchange-traded futures contracts can be valued as an option on the minimum. The CBT's T-bond futures, for example, provides the seller with the right to deliver the cheapest of a number of deliverable T-bond issues.

3.3 Approximation Methods

Many option valuation problems do not have explicit closed-form solutions. Probably the best known example is the valuation of standard American-style options. With American-style options, the option holder has an infinite number of exercise opportunities between the current date and the option's expiration date, making the problem intractable from a mathematical standpoint.[26] But, many other examples also exist. Hundreds of different types of exotic options trade in the OTC market, and many, if not most, of these options do not have analytical formulas. Nonetheless, all of them can be valued accurately using the BSM model. If a risk-free hedge can be formed between the option and the underlying asset, the BSM risk-neutral valuation theory can be applied, albeit through the use of numerical methods. Below three types of commonly-applied approximation methods are described.[27]

3.3.1 Lattice-based methods

A number of numerical methods for valuing options are lattice-based. These methods replace the BSM assumption that asset price moves smoothly and continuously through time with an assumption that the asset price moves in discrete jumps over discrete intervals during the option's life.

Binomial method

Perhaps the best-known lattice-based method is the binomial method, developed independently by Cox, Ross and Rubinstein (1979) and Rendleman and Bartter (1979). Given the importance of the role that this approximation method plays within the derivatives industry, its development and relation to the BSM model are described more fully.

To develop the binomial method, it is more convenient to use the dynamics of the logarithm of asset price rather than asset price. Under the BSM model, asset price follows the geometric Brownian motion described by (16). It can be shown by Ito's lemma that, if asset price follows (16), the logarithm of asset price follows

$$d \ln S = \mu dt + \sigma dz, \tag{22}$$

[26] An exception is, of course, an American-style call option on an asset where $i \le 0$, as was discussed in Section 2.

[27] The techniques included in this discussion are lattice-based methods, Monte Carlo simulation methods, and quasi-analytical methods. A less traveled route to valuing American-style options is numerical integration. See, for example, Parkinson (1977).

where $\mu = \alpha - \frac{\sigma^2}{2}$ and the subscript on α has been suppressed. Since the binomial method replaces the assumption of continuous asset price movements with price movements over a discrete interval, (22) is re-written as

$$\Delta \ln S = \mu \Delta t + \sigma \varepsilon \sqrt{\Delta t}\ , \tag{23}$$

where ε is a normally distributed random variable with zero mean and unit standard deviation. Under the binomial method, the option's life is divided into fixed length time steps, and, in each time step, the asset price is allowed to jump up or down. If n is the number of time steps, each time increment has length $\Delta t = T / n$, where T is the time to expiration of the option.

The binomial distribution is characterized by the size of its price steps and their probabilities. The parameters are chosen in such a way that the mean and the variance of the discrete binomial distribution are consistent with the mean and the variance of the continuous log-normal distribution underlying the BSM model. Under the BSM assumptions, the logarithm of the asset price at the end of the time increment Δt is normally distributed with mean $\ln S + \mu \Delta t$ and variance $\sigma^2 \Delta t$. First, set the mean of the binomial distribution equal to the mean of the logarithm of asset price distribution, that is,

$$p(\ln S + v) + (1-p)(\ln S + w) = \ln S + \mu \Delta t\ . \tag{24}$$

In (24), p is the probability that the logarithm of asset price changes by v, and $1-p$ is the probability that the logarithm of asset price changes by w. What remains is

$$pv + (1-p)w = \mu \Delta t\ . \tag{25}$$

Next, set the variance of the binomial distribution equal to the variance of the logarithm of asset price distribution,

$$p(\ln S + v - (\ln S + \mu \Delta t))^2 + (1-p)(\ln S + w - (\ln S - \mu \Delta t))^2 = \sigma^2 \Delta t\ . \tag{26}$$

The $\ln S$ terms are again irrelevant, and, with a little additional algebra, equation (26) becomes

$$pv^2 + (1-p)w^2 = \sigma^2 \Delta t + \mu^2 \Delta t^2\ . \tag{27a}$$

Equation (27a) is a little unusual in the sense that it has a term that includes the time increment squared, Δt^2. In applying the binomial method to value options, however, a large number of time steps is usually used, so Δt is very small. Consequently, terms with higher order values of Δt can safely be ignored. Ignoring the higher order term, (27a) can be written

$$pv^2 + (1-p)w^2 = \sigma^2 \Delta t \,. \tag{27b}$$

Note that the values on the right-hand side of (25) and (27) are known. They are parameters of the normal distribution of the logarithm of asset prices. The objective is to find the values of *v*, *w*, and *p,* which characterize the binomial distribution. With two equations (i.e., (25) and either (27a) or (27b)) and three unknowns, we cannot solve for the parameters *v*, *w*, and *p* uniquely, so another constraint must be imposed. Below, the constraints used in two well-known implementations of the binomial method are discussed.

Cox, Ross and Rubinstein (1979) (hereafter "CRR") impose the symmetry constraint, $w = -v$, where *v* is a positive increment. This implies that the asset price will either rise to a level, $\ln S + v$, or fall to a level, $\ln S - v$ over the next increment in time Δt. CRR use (27b) to tie the variance of the binomial distribution to the variance of the logarithm of asset prices. The value of *v* becomes

$$v = \sigma\sqrt{\Delta t} \,. \tag{28}$$

With *v* and, hence $w \; (= -v)$, known, only the level of probability, *p*, remains to be identified. Substituting into (28) into (26) and rearranging,

$$p = \frac{1}{2} + \frac{1}{2}\left(\frac{\mu}{\sigma}\right)\sqrt{\Delta t} \,. \tag{29}$$

Substituting the relation between the mean continuously compounded rate of price appreciation, μ, and the continuously compounded mean rate of price appreciation, *b*, that is, $\mu = b - .5\sigma^2$,

$$p = \frac{1}{2} + \frac{1}{2}\left(\frac{b - .5\sigma^2}{\sigma}\right)\sqrt{\Delta t} \,. \tag{30}$$

In another well-known implementation of the binomial method, Jarrow and Rudd (1983) (hereafter "JR") impose the constraint that the up-step and the down-step probabilities are both equal to ½ . This means that the relation between the mean of the binomial distribution and the mean of the change in the logarithm of prices (25) may be written

$$v + w = 2\mu\Delta t . \tag{31}$$

To express the relation between the variances, JR use (27a). The variance relation can be re-written as

$$v^2 + w^2 = 2\sigma^2\Delta t + \frac{1}{2}\left(4\mu^2\Delta t^2\right). \tag{32}$$

Substituting the square of (31) into the parentheses on the right-hand side of (32), rearranging, factoring, taking the square root and then simplifying,

$$v - w = \sigma\sqrt{\Delta t} . \tag{33}$$

Equations (31) and (33) can now be used to identify *u* and *v*. With the probability set equal to ½, the up-step coefficient is

$$v = \mu\Delta t + \sigma\sqrt{\Delta t} = \left(b - .5\sigma^2\right)\Delta t + \sigma\sqrt{\Delta t} \tag{34a}$$

and the down-step coefficient is

$$w = \mu\Delta t - \sigma\sqrt{\Delta t} = \left(b - .5\sigma^2\right)\Delta t - \sigma\sqrt{\Delta t} . \tag{34b}$$

The distinction between the two approaches is that the CRR method handles the rate of drift in the asset price through the up-step and down-step probabilities, while the JR method handles the drift through the step sizes.

With the probabilities and step sizes linked to the parameters of the BSM lognormal price distribution, the steps of the approximation method are now described. The first step is to enumerate the possible paths that the asset price may take between now and the option's expiration. The user chooses the number of time steps, *n*, and thereby sets the time increment, $\Delta t = T / n$, and step sizes (i.e., (28) under the CRR approach, and (34a) and (34b) under the JR approach). Generally, the implementations use an asset price lattice rather than the logarithm of asset price, so the jumps at each step are multiplicative rather than additive. In the CRR method, for example, the initial asset price S moves to uS or dS at the end of the first time step, where the up-step coefficient is $u = e^{v} = e^{\sigma\sqrt{\Delta t}}$ and the down-step coefficient is $d = 1/u$. At the end of the second time step, the asset prices are uuS, S, and ddS, and so on. With n time steps, there will be $n+1$ terminal asset price nodes. The greater the number of time steps, the more precise the method. The cost of the increased precision, however, is computational speed. With n time steps, 2^{n} asset price paths over the life of the option are considered. With 20 time steps, this means over a million paths.

The second step of the binomial method is to value the option at expiration at each of the possible asset price levels. At expiration, the option value at each asset price node is simply the option's intrinsic value, that is, $\max(0, S_j - X)$ for a call option and $\max(0, X - S_j)$ for a put, where j represents the j-th node. Once the option values at all nodes at time n are identified, the procedure steps backward one time step.

The third step is recursive. At time $n-1$, the value of the option at each node is computed by taking the present value of the expected future value of the option. The expected future value is simply the probability of an up-step times the option's value if the asset price steps up plus the probability of a down-step times the option's value if the asset price steps down. The discount rate in the present value computation is the risk-free rate of interest. Before proceeding backward another step in time, it is necessary to determine whether any of the computed options values at time $n-1$ are affected by a feature of the contract. If the option is American-style, for example, the computed option value at each node must be compared with its early-exercise proceeds. If the early exercise proceeds exceed the computed value, the computed value is replaced by the amount of the exercise proceeds. The interpretation is, of course, that if the option holder finds himself standing at that time in the option's life, with the underlying asset priced at that level, he will exercise his option. If proceeds are less, the option "is worth more alive than dead", and the computed value is left undisturbed. Note that, if the check of the early exercise condition is not performed, the binomial method will produce an approximate value for a European-style option.[28] The procedure now takes another step back in time, repeats the computations of all nodes, and then checks for early exercise. The procedure is repeated again and again until only a single node remains at time 0. This node will contain the value of the American-style option, as approximated by the binomial method.

The binomial method has wide applicability. Aside from the American-style option feature, which is easily incorporated within the framework, the binomial method can be used to value many types of exotic options. Knockout options, for example, can be valued using this technique. A different check on the computed option values at the nodes of the intermediate time steps between 0 and n is imposed. If the underlying asset price falls below the option's barrier, the option value at that node is set equal to zero. The method can also be extended to handle multiple sources of asset price uncertainty. Boyle (1988) and Boyle, Evnine and Gibbs (1989) adapt the binomial procedure to handle exotics with multiple sources of uncertainty including options on the minimum and maximum, spread options,[29] and so on.

Trinomial method

The trinomial method is another popular lattice-based method. The trinomial method, as outlined by Kamrad and Ritchken (1991), allows the asset to move up, down, or stay the same at each time increment. Again, the parameters of the discrete distribution are chosen in a manner consistent with the lognormal distribution, and the procedure begins at the end of the option's life and works backward. By having three branches as opposed to two, the trinomial method provides greater accuracy than the binomial method for a given number of time steps. The cost is, of course, the greater the number of branches, the slower the computational speed. The trinomial method is also useful in valuing options that depend on the prices of two underlying assets.

---

[28] Indeed, a useful way to gauge the approximation error of the various numerical methods is to implement them on valuation problems for which there is an analytical formula.

[29] A spread option is an option whose underlying source of uncertainty is the difference between two asset prices. Since the difference between two log-normally distributed variables is not log-normal, valuing spread options is not merely a matter of applying the BSM model to the difference in asset prices. For an application of the binomial method in valuing spread options, see Whaley (1996).

<u>Finite difference methods</u>

Finite difference methods solve the BSM differential equation (19) by converting it into a set of difference equations and then solving the difference equations iteratively. The simplest finite difference method, and, indeed, the first application of a lattice-based procedure to value an option, is the *explicit method*, applied by Schwartz (1977) to value warrants and by Brennan and Schwartz (1977) to value American-style put options on stocks. The explicit finite difference method is the functional equivalent of the trinomial method in the sense that the asset price moves up, down, or stays the same at each time step during the option's life. The difference in the techniques arises only from how the price increments and the probabilities are set. Once the lattice is traced out, the valuation computations begin at the end of the option's life and work backwards. The *implicit method* is computationally more robust than the explicit method (i.e. converges to the differential equation as the asset price and time increments approach zero), however, requires simultaneous solution to the difference equations and therefore considerably more computational time. The chief advantage of the implicit method is its accuracy.

3.3.2 Monte Carlo methods

Boyle (1977) introduced Monte Carlo simulation as a means of valuing options. Like the lattice-based procedures, the technique involves simulating possible paths that the asset price may take over the life of the option. And, again, the simulation is performed in a manner consistent with the lognormal asset price process. To value a European-style option, each sample run is used to produce a terminal asset price, which, in turn, is used to determine the terminal option value. Over the course of many sample runs, a distribution of terminal option values is obtained. The mean of the distribution is then discounted to the present to value the option. An advantage of the Monte Carlo method is that the degree of valuation error can be assessed directly using the standard error of the estimate. The standard error equals the standard deviation of the terminal option values divided by the square root of the number of trials.

Another advantage of the Monte Carlo technique is its flexibility. Since the path of the asset price beginning at time 0 and continuing through the life of the option is observed, the technique is well-suited for handling barrier-style options, Asian-style options, Bermuda-style options, and other exotics. Moreover, it can be easily adapted to handle multiple sources of price uncertainty. The technique's chief disadvantage is that it can only be applied when the option payout does not depend on its value at future points in time. This eliminates the possibility of applying the technique to American-style option valuation, where the decision to exercise early depends on the value of the option that will be forfeited. In addition, a large number of trials are required to get the level of valuation accuracy to a reasonable level.

3.3.3 Quasi-analytical methods

In valuing American-style options, the chief difficulty lies identifying a simple expression for the optimal early exercise boundary. Within the lattice-based procedures, the matter is handled by brute force, and the approximation procedure proceeds backwards through the option's life, comparing computed "alive" values of the option with the option's early exercise proceeds. Quasi-analytical methods make different simplifying assumptions regarding the optimal early exercise boundary and then proceed analytically. Below the compound option and quadratic approximations are discussed.[30]

Compound option approximation

Geske and Johnson (1984) use the Geske (1979a) compound option valuation model to develop an approximate value of an American-style option. The approach is intuitively appealing. An American-style option is, after all, a compound option with an infinite number of early exercise opportunities. While valuing an option in this manner makes intuitive sense, the problem is intractable from a computational standpoint. The Geske/Johnson insight is that, although an option with an infinite number of early exercise opportunities cannot be value analytically, its value can be extrapolated from the values of a sequence of "pseudo-American" options with zero, one, two, and perhaps more early exercise opportunities at discrete, equally-spaced, intervals during the option's life. The advantage of this approach is that each of these pseudo-American options can be valued analytically. Unfortunately, with each new option added to the sequence, the valuation of a higher-order multivariate normal integral is required. With no early

[30] Carr (1998) develops a new approach for determining American-style option values and exercise boundaries based on a technique called *randomization*.

exercise opportunities, only a univariate function is required; however, with one early exercise opportunity, a bivariate, with two opportunities, a trivariate, and so on. The more of these options used in the series, the greater the precision in approximating the limiting value of the sequence. The cost of increased precision is that higher-order multivariate integral valuations are time-consuming computationally.

Quadratic approximation

Barone-Adesi and Whaley (1987) present a quadratic approximation for valuing American-style options. Their approach, based on the work of MacMillan (1986), separates the value of an American-style option into two components: the European-style option value and an early exercise premium. Since the BSM formula provides the value of the European-style option, they focus on approximating the value of the early exercise premium. By imposing a subtle change to the BSM partial differential equation, they obtain an analytical expression for the early exercise premium, which they then add to the European-style option value, thereby providing an approximation of the American-style option value. The advantages of the quadratic approximation method are speed and accuracy.

3.3.4 Moore's law

For many years, the search for quasi-analytical approximations was an important research pursuit. Using lattice-based procedures or Monte Carlo simulation were impractical in real-time applications. This pursuit has become much less critical, thanks to Moore's Law. In April 1965, Gordon Moore, an engineer and co-founder of Intel predicted that integrated circuit complexity would double every two years. The prediction has been surprisingly accurate. In the late 1970's, when the lattice-based and Monte Carlo simulation methods were first applied to option valuation problems, Intel's most advanced microprocessor technology was the 8086 chip. Today, the Pentium IV microprocessor is more than 2,000 times faster, and the impracticality of lattice-based and simulation-based methods has been substantially reduced.

3.4 Generalizations

The generalizations of the BSM option valuation theory focus mostly on relaxing the constant volatility assumption.[31] Some valuation models assume that the local volatility rate as a deterministic function of the asset price or time or both. Others assume that volatility, like asset price, is stochastic.

3.4.1 Deterministic volatility functions

The BSM risk-free hedge mechanics are preserved under the assumption that the local volatility rate is a deterministic function of time or the asset price, so risk-neutral valuation remains possible. The simplest in this class of models is the case where the local volatility rate is a deterministic function of time. The asset price follows the process,

$$dS = \alpha S dt + \sigma(t) S dz .$$

Under this assumption, Merton (1973) shows that the valuation equation for a European-style call option is the BSM formula (20), where the volatility parameter is the average local volatility rate over the life of the option.

Other models focus on the relation between asset price and volatility. These models attempt to account for the empirical fact that, in at least some markets, volatility varies inversely with the level of asset price. One such model is the constant elasticity of variance (CEV) model proposed by Cox and Ross (1976). The CEV asset price dynamics are

$$dS = \alpha dt + \delta S^{\theta/2} dz ,$$

where the instantaneous variance of the stock price is $\delta^2 S^\theta$, the elasticity of variance with respect to stock price equals θ, and $0 \le \theta \le 2$. The instantaneous variance of return, σ^2, is given by $\sigma^2 = \delta^2 S^{\theta-2}$. If the value of θ equals 2, the instantaneous variance of return is constant, which is the assumption underlying Black and Scholes. If $\theta = 0$,

[31] The assumption of constant interest rates is relaxed in Merton (1973), Bailey and Stulz (1989), and Amin and Jarrow (1992), among others. These studies have attracted less attention than work exclusively on stochastic volatility because empirical investigations focus on exchanged-traded options, exchange-traded

volatility is inversely proportional to the asset price, and a European-style call option can also be valued analytically using a formula called the "absolute diffusion model."[32] For the general case in which $0 \le \theta \le 2$, analytical solutions are not possible. Valuation can be handled straightforwardly, however, using lattice-based or Monte Carlo simulation procedures.

Recently, Derman and Kani (1994), Dupire (1994), and Rubinstein (1994) developed a valuation framework in which the local volatility rate is a deterministic, but unspecified, function of asset price and time,

$$dS = \alpha S dt + \sigma(S,t) S dz \,.$$

Rather than positing a structural form for their deterministic volatility function (DVF), they search for a binomial or trinomial lattice that achieves an *exact* cross-sectional fit of reported option prices.[33] Rubinstein uses an "implied binomial tree" whose branches at each node are designed (either by choice of up-and-down increment sizes or probabilities) to reflect the time variation of volatility.

3.4.2 Stochastic volatility functions

The effects of stochastic volatility on option valuation are modeled by either superimposing jumps on the asset price process, or allowing volatility to have its own diffusion process, or both. Merton (1976), for example, adds a jump term to the usual geometric Brownian motion governing asset price dynamics, that is,

$$dS = (\alpha - \lambda k) S dt + \sigma S dz + S dq \,,$$

where dq is the Poisson process generating the jumps and dz and dq are independent. The parameter α is the instantaneous expected return on the asset, λ is the mean number of arrivals per unit time, and $k \equiv E(Y-1)$, where $(Y-1)$ is the random variable percentage change in asset price if the Poisson event occurs. By assuming that the jump

---

options are generally short-term, and short-term options are relatively insensitive to interest rates (and the assumed interest rate process).

[32] See Cox and Ross (1976).

[33] In contrast, Dumas, Fleming and Whaley (1998) implemented the deterministic volatility function option valuation model by expanding the local volatility rate function in a Taylor series and estimating the parameters of the function directly.

component of an asset's return is unsystematic, Merton creates a risk-free portfolio in the BSM sense and applies risk-neutral valuation for European-style options.

Merton's case is an exception to the rule, however. In general, risk-neutral valuation is not possible where volatility is stochastic because volatility is not a traded asset and, consequently, the BSM risk-free hedging argument does not apply. In studies of option valuation under stochastic volatility, asset price and asset price volatility are modeled as separate, but correlated, diffusion processes. Asset price is usually assumed to follow geometric Brownian motion with a stochastic volatility rate. The assumptions governing volatility vary. Hull and White (1987) assume volatility follows geometric Brownian motion. Scott (1987) models volatility using a mean-reverting process, and Wiggins (1987) uses a general Wiener process. Bates (1996) combines both jump and volatility diffusions in valuing foreign currency options.[34] For the asset price process, he uses the Merton (1976) assumption. For volatility movements, he assumes that variance follows the mean reverting, square root process. In all of these pairings of price and volatility process assumptions, however, the resulting differential equation describing the option price dynamics is utility-dependent and, therefore, difficult to implement. Heston (1993) derives a closed-form solution for valuing options with stochastic volatility by assuming that the risk premium is proportional to the volatility rate.

4. STUDIES OF NO-ARBITRAGE PRICE RELATIONS

In section 2, a number of no-arbitrage price relations were described. These relations are based on the absence of costless arbitrage opportunities in an efficiently functioning market. This section reviews some of the studies that have empirically tested the no-arbitrage bounds on the prices of derivatives contracts. The studies are divided into two categories—forward/futures and options.

[34] The listed studies are by no means exhaustive. Other examples include Melino and Turnbull (1990, 1995) and Stein and Stein (1991).

4.1 Forward/futures prices

4.1.1 Cost of carry relation

Studies of the cost of carry relation (equations (3a) and (3b) in Section 2) using futures prices are few.[35] There are two primary reasons. First, a no-arbitrage price relation is a *relative* pricing relation, which means that the prices of the derivatives contract and the asset must be observed simultaneously. This *price synchronization* requirement eliminates the possibility of meaningful empirical investigation using daily closing price data for most futures markets including bonds, currencies, and commodities. The CBT's T-bond futures market, for example, closes at 2:00 PM CST while the underlying cash Treasury bond market closes at 4:30 PM. Using daily closing prices from these markets to examine the empirical performance of the cost of carry model would produce frequent (and large) violations of the cost of carry relation. Second, for many asset classes, the cost of carry relation is mis-specified. Physical commodities, for example, cannot be sold short. This means that the cost of carry relation will hold only as a weak inequality in which the futures price will be less than or equal to the asset price plus the cost of carry. Another reason is that many futures contracts have embedded options. The grain contracts traded on the CBT, for example, allow the short to deliver one of a number of different qualities of the underlying asset at one of a number of different delivery locations. Naturally, the short will choose the least expensive. Other contracts provide the short a great deal of flexibility regarding the timing of delivery during the delivery month. This timing option may also have significant value. Certain financial futures like the CBT's T-bond and T-note futures have both options. Since the purchaser of the futures will want payment for the options being provided to the seller, the futures price will lie below the asset price by the cumulative value of these options.[36]

A market ideally suited for empirical examination of the cost of carry relation is the stock index futures market. Nearly synchronous price data are easily accessible, and the contract design is unencumbered by embedded options. Stock index futures began

---

[35] This ignores, of course, a large literature examining forward price relations such as *interest rate parity*.

[36] The values of these options can, of course, be modeled. The "cheapest-to-deliver" option, for example, is an option on the minimum of several assets and can be valued using the Stulz (1982) and Johnson (1987) framework.

trading in 1982. The Kansas City Board of Trade (KCBT) was the first to launch such a market introducing the Value Line index futures in February 1982. The Chicago Mercantile Exchange (CME) followed two months later with the S&P 500 index futures. The price synchronization problem is not as a serious problem as for other markets, since the futures market closes at 3:15PM CST, only fifteen minutes after the stock market. Index futures contracts are cash settled, with no embedded quality or timing options. In addition, unlike commodities markets, stocks can usually be sold short freely with full use of proceeds (at least by market professionals such as index arbitrageurs). This leaves only two sources for price deviations in the cost of carry relation—trading costs and staleness in the reported index level. The *staleness* or *infrequent trading issue* has to do with the fact that the reported index level is an amalgam of the last trade prices of individual stocks, some of which may not have traded for several hours. Hence, the reported index is always "stale," so to speak.[37] On average, however, neither trading costs nor infrequent trading should cause the actual futures price to be different from its theoretical level.

Empirical studies of stock index futures generally focus on the CME's S&P 500 futures contract. It is by far the most active index futures contract in the world. The earliest study of the cost of carry relation for the S&P 500 futures is Figlewski (1984). During the period June 1982 through September 1983 (essentially the first fifteen months of trading of the S&P 500 futures), he finds that the average futures price is too low relative to the index level. The behavior is temporary, however. Using fifteen-minute (simultaneous) price observations, MacKinlay and Ramaswamy (1988) examine the difference between the actual futures price and theoretical futures price for all S&P 500 futures during the period April 1982 through June 1987. They find that, while the average deviation is positive and as high as .78 index points for contracts as recent as December 1984, the average deviation is consistently below .20 for all contract maturities after the September 1985 contract. The evidence indicates that in the early days of trading, index arbitrageurs had not yet fully developed effective mechanisms for short selling the basket

[37] Miller, Muthuswamy, and Whaley (1994) modeled the effects that infrequent trading has on the time-series properties of the theoretical basis in the S&P 500 futures market.

of stocks underlying the S&P 500. Consequently, the futures price did not rise to its proper theoretical level.

4.1.2 Forward/futures price relation

The difference between forward and futures prices is driven by the marking-to-market practice of futures markets. When interest rates are known, the forward and futures prices will be equal, as was demonstrated in Section 2. When interest rate are uncertain, however, the price differential between the forward and the futures is driven by the covariance between futures price changes and discount bond price changes. A positive (negative) covariance implies that the futures price is less (greater) than the forward price, as was discussed in Section 2. Cornell and Reinganum (1981) compare the daily closing prices of five different FX futures contracts traded on the Chicago Mercantile Exchange (CME) with FX forward rates quoted at the same time by the Continental Illinois Bank during the period June 1974 through June 1979. They conclude that any differences are statistically and economically insignificant. Not surprisingly, they also document that the sizes of the covariances are all very small, on order of 1.0×10^{-7}. Chang and Chang (1990) argue that the Cornell and Reinganum results may be undermined by (a) the mismatches in the delivery dates of the futures and forwards, and (b) a sample period that is dominated by an economic cycle in which interest rates are volatile but currency rates are not. They adjust for the mismatch problem in the 1974-1979 period as well as investigate the differences in prices for the 1979-1987 period. They, too, conclude that there is no meaningful difference in the levels of forward and futures prices for currencies.

Early tests of the forward/futures price relation in bond markets focused on the CME's T-bill futures market and forward prices implied from spreads in the cash T-bill market. Rendleman and Carabini (1979), for example, examine daily closing price data during the first two years of market operation—January 1976 through March 1978. They report frequent violations of the no-arbitrage relation (of equal prices), but that the potential arbitrage gains are not worth exploiting due to trading costs, monitoring costs, maturity mismatches, and so on. Cornell (1981) and Viswanath (1989) investigate whether the differential tax treatment of cash T-bills and T-bill futures gains/losses in the early years of the market is the cause. A more recent study by Meulbroek (1992) focuses on the more liquid Eurodollar futures market using daily data during the period March 1982 through June 1987. She finds strong support of the CIR predictions. Among other things, she finds that the covariance between futures and bond price changes and the covariance of forward price changes and bond price changes are positive. Under these conditions, the futures price should be (and is documented as being) less than the forward price.

The relation between forward and futures prices has also been examined using commodity futures prices. Using copper and silver data, French (1983) finds that the forward-futures price differential has the sign predicted by the CIR model, but not the magnitude. French ascribes the latter result to measurement error. Park and Chen (1985) examine daily price data for forwards and futures written on six physical commodities during the period July 1977 through December 1981 and find that the differences between futures and forward prices are positive and significant in a statistical sense and are consistent with the CIR predictions regarding the difference between the variance of the bond price changes and the covariance of the bond prices changes with the price changes of the underlying commodity. They conclude that differences between futures and forward prices can be attributed to the marking-to-market process.

4.2 Option prices

Empirical tests of no-arbitrage option price relations generally fall into two groups: those that examine violations of intrinsic value relations and those that examine violations of put-call parity relations.[38] In both cases, however, the structure of the

[38] A third no-arbitrage condition called the convexity relation is occasionally examined. The convexity relation, as it applies to call options, is $C(X_2) \leq qC(X_1) + (1-q)C(X_3)$, where the option exercise prices have the order $X_1 < X_2 < X_3$ and q is defined by $X_2 = qX_1 + (1-q)X_3$. In the event the convexity relation is violated, a costless arbitrage profit may be earned by engaging in a "butterfly spread", that is, by selling the call with exercise price X_2, buying q and 1-q units of the calls with exercise prices X_1 and X_3, respectively. Galai (1979) examines CBOE stock options traded during the period April to October 1973. When closing prices are used, he reported violations in 24 of 1,000 cases. When intraday trade prices are used, however, all of these violations disappear. Similarly, Bhattacharya (1983) examines 1,006 triplets of CBOE call options written on the same underlying stock during the period August 1976 through June 1977 and finds no case in which the convexity condition is violated.

experiment is the same. Are there costless arbitrage opportunities available in the options market?

4.2.1 Intrinsic value relations

Intrinsic value relations were derived in Section 2. The earliest investigations of intrinsic value violations were conducted on prices from the CBOE's stock option market. Galai (1978) examines the prices of call options[39] on 32 stocks during the period April through November 1973, the first five months of CBOE operations. Using daily closing prices, he finds frequent violations. A trading strategy that exploits these violations generates positive profits, but the magnitude of the profits is small. Bhattacharya (1983) uses intraday trade price to examine call options on 58 stocks during the period August 1976 through June 1977, 86,137 option records in all. Of these, 1,304 quotes of 54,375 violate the European-style option intrinsic value and 442 quotes of 32,432 violated an early exercise lower bound. The average mis-pricings for the violations are small—\$9.88 per contract[40] and \$10.85 per contract, respectively. After reasonable tradings costs, profitable arbitrage trading opportunities virtually disappear.

The CBOE introduced the first stock index option contract on March 11, 1983, on the S&P 100 index. The American Exchange (AMEX) launched the Major Market Index contract on April 29th of the same year. Evnine and Rudd (1985) examine the trade prices of these contracts using on-the-hour data over the period June 1984 through August 1984. They report 30 violations of the immediate exercise bound in a sample of 1,091 (2.7%) for the S&P 100 calls; and 11 violations in a sample of 707 observations (1.6%) for MMI calls. Interestingly, all violations occur during the first week of August 1984, which was a particularly turbulent time in the stock market. In such a period, the likelihood of reporting errors is higher than normal. Furthermore, at the time, the indexes were not traded contracts. To exploit such violations, it is necessary to short sell the index portfolio. Finally, as noted earlier, the reported index level is always a "stale" indicator of the true level of the index.

---

[39] The exclusive focus on call options should not be surprising. Recall that in Section 1 it was noted the trading in put options on stocks did not commence until June 3, 1977, and, even then, only on an experimental basis.

[40] A contract is for 100 shares. The profit per share is \$.0988.

The Philadelphia Exchange (PHLX) launched trading in American-style options on five different currencies in December 1982. Bodurtha and Courtadon (1986) compare option prices to their immediate exercise proceeds during the period February 1983 through September 1984. When end-of-day prices for the option and the underlying exchange rate are used, frequent violations are found. When option trade prices are matched against the currency rate prevailing at the time (from Telerate bid/ask quotations provided by the PHLX), they find that only .9% of call option transaction prices and 6.7% of put option transaction prices violate the immediate-exercise lower bounds. Finally, the percentages drop to .03 % for calls and .2% for puts when transaction costs are taken into account. Ogden and Tucker (1987) perform a similar experiment using the American-style British pound, Deutschemark, and Swiss franc futures options traded on the CME during the calendar year 1986. They carefully match each futures option trade price with the price of the underlying futures at the time of its nearest preceding trade, eliminating those with time differentials greater than thirty minutes. In all, 81,257 call trades and 44,362 put trades are identified. Before trading costs, 1,756 (2.2%) call option violations and 251 (.6%) put violations are reported. After trading costs, the percentages drop to 1.8% and .5% for calls and puts, respectively.

The Chicago Board of Trade (CBT) launched T-bond futures option trading on October 1, 1982. Blomeyer and Boyd (1988) examine the immediate exercise proceeds bounds for call and put options trades during the period October 1982 through June 1983. They examine both *ex post* and *ex ante* trading strategies. An ex post opportunity is signaled only if the arbitrage trade is profitable after transaction costs. An ex ante trade is predicated on an ex post signal, and the trade is executed at the next available prices for the option and the futures. For calls, only 377 of 50,477 (.7%) trades signal an ex post opportunity, and, if the ex ante strategy is executed, the trader incurs an average loss of $38 per trade. Of the 30,065 put trades, 90 (.3%) satisfy the ex post requirement and the average ex ante loss is $42 per trade.

4.2.2 Put-call parity relations

Recall from Section 2 that put-call parity arises from conversion (and reverse conversion) arbitrage. Borrowing to buy a put and its underlying asset, for example, is tantamount to buying a call. The earliest systematic empirical examination of the put-call parity relation appears in Stoll (1969). Stoll uses OTC option price data that the Put and Call Dealers Association provided the Securities and Exchange Commission on a weekly basis during the two-year period 1966-67. The put-call parity relation that Stoll tests does not include an expression for cash dividends since the OTC stock options that traded at the time were dividend-protected.[41] Rather than examining violations of put-call parity *per se*, Stoll uses a pooled time-series cross-sectional regression framework. Although the coefficients in the regression model differ slightly from their theoretical predictions, Stoll concludes that the evidence supports the theory of put-call parity.[42]

Klemkosky and Resnick (1979) perform the first test of put-call parity using exchange-traded options. They collect monthly observations for calls and puts traded on 15 underlying stocks during the period July 1977 through June 1978[43] including options traded on the CBOE, AMEX and PHLX. These data are obtained from Francis Emery Fitch, Inc. and include the price, volume, and time of each trade of each option and its underlying stock. To ensure simultaneity of prices, they require that the call, put, and underlying stock trade within one minute of each other. They conclude that their test results are consistent with put-call parity and the efficiency of the option market.

Evnine and Rudd (1985) examine the put-call parity for S&P 100 and MMI options. In general, they find more violations than are common in studies of options in other markets. Considering the put-call parity relation does not apply when trading in the underlying asset is restricted, this is not surprising. The most common violation is where the call appears overpriced relative to the put and the underlying index. This is surprising considering that this is the easiest arbitrage to execute (i.e., the stock index portfolio is bought rather than sold short). One possible explanation for this result is that the market generally trended upward during this period. Given infrequent trading of the stocks comprising the index, the reported index level always lags the “true” value. In an upward

---

[41] At the time of a cash dividend payment during the option’s life, the exercise price is reduced by the dividend amount.

[42] Gould and Galai (1974) re-examine put-call parity using the American-style option relation and reach a similar conclusion.

trending market, index call prices are reacting more quickly than the prices of <u>all</u> of the stocks in the index portfolio. This explanation is further supported by the fact that the relative frequency of over-priced calls is lower for the 20-stock MMI than the 100-stock S&P 100 index.

In addition to the intrinsic value tests discussed earlier, Bodurtha and Courtadon (1986) examine the American-style put-call parity relation using PHLX options using simultaneous spot and option prices during the period February 1983 through Septmeber 1984. Across the five currencies, they observe only 25 violations of put-call parity in 8,509 tests (0.3%), all but one of which disappear when reasonable trading costs are considered. Using approximately simultaneous trade prices of CME options on three currency futures during the calendar year 1986, Ogden and Tucker (1987) report 466 violations in 29,288 tests (1.6%) net of transaction costs.

4.3 Summary and analysis

The conclusion that must be drawn from the empirical investigations discussed in this section is that it is difficult, if not impossible, to earn abnormal profits from violation of no-arbitrage price relations. Violations, where they have been reported, are usually in the early stages of the market's development. The fact that virtually no violations appear in established markets is reassuring, since, were they to occur, the fundamental financial-economics underpinning that individuals prefer more wealth to less would have to be reconsidered. Reported violations of no-arbitrage relations in today's markets are most likely the result of (a) stale or non-synchronous prices, (b) data recording problems, and/or (c) mis-measured trading costs.

5. STUDIES OF OPTION VALUATION MODELS

The empirical performance of competing option valuation models has been evaluated using three different types of methodology. The first type of methodology focuses in-sample on either deviations of observed prices from model values (i.e., pricing

---

[43] Beginning on June 3, 1977, the Securities Exchange Commission allowed the five stock options exchanges to begin trading put options on five different stocks each. It is unclear why the put options traded on the Midwest Exchange and the Pacific Exchange were not included in the sample.

errors) or on systematic patterns in implied volatilities. In the pricing error tests, a single volatility estimate is used to value all option series within the class,[44] and then deviations of the prices of the individual series from their model values are tabulated. In the implied volatility tests, the volatility level of each option series is inferred (or implied) by setting its price equal to the model value, and then the implied volatilities for the different option series in the class are tabulated. These studies are labeled "Pricing error/implied volatility anomalies" and are discussed first.

Another methodology for investigating the performance of different option valuation models is to simulate a trading strategy. To understand the structure of these investigations, recall that, according to the Black-Scholes/Merton model, a risk-free hedge can be formed between an option and its underlying asset (and that the return of this portfolio will therefore be equal to the risk-free rate of interest). If the BSM assumptions hold and the BSM formula identifies a particular call option as being over-priced, then a portfolio formed by selling the call and "delta-hedging" continuously over its remaining life should produce a rate of return in excess of the risk-free rate. Option valuation studies using strategies such as this are discussed under the heading, "Trading simulations."

The final category is called "Informational content of implied volatility." Studies falling in this category are characterized by their focus on understanding the predictive power of implied volatility. Some tests are cross-sectional in nature, comparing implied volatility to the realized volatility of the underlying asset's daily returns during the option's life. Such cross-sectional investigations are possible only in cases where there are multiple option classes on the same type of asset. This limits this type of analysis to stock option markets. Consequently, most studies of implied volatility are time-series in nature. They tend to focus on stock index options. In the early 1980s, a number of index options were launched, with varying degrees of success. Three contract markets have dominated in terms of trading volume—the CBOE's S&P 100 and S&P 500 index options, and the CME's S&P 500 futures options. The fact that intraday data for these

[44] An *option class* refers to all options written on the same underlying asset (e.g., all options written on the shares of IBM). An *option series* is a single type of option written on the asset, and is identified by three attributes: (a) exercise price, (b) expiration date, and (c) call or put.

option classes are widely available and that, historically, there has been intense empirical interest in the phenomenon of "market volatility" has made this category of study the most voluminous.

5.1 Pricing errors/implied volatility anomalies

The first empirical studies in this category appeared in the late 1970s and early 1980s. Black (1975) reports that the BSM formula systematically under-priced deep out-of-the-money calls and overpriced deep in-the-money calls during the first two years of stock option trading (i.e., 1973-1975) at the CBOE. One reason that this "moneyness bias" may appear is that the BSM formula values European-style options while the stock options traded on the CBOE are American-style. If the stocks underlying the options paid no dividends, this would be a non-issue,[45] however, most of the stocks at the time were 'blue-chip' stocks with generous dividend payouts. Whaley (1982) uses weekly closing prices for CBOE call options during the period January 1975 through February 1978 to examine whether the American-style call option valuation formula of Roll (1977)-Geske (1979b)-Whaley (1981) eliminates the moneyness bias. He finds that explicit recognition of dividends and the early exercise premium reduces, but does not eliminate, the moneyness bias.

MacBeth and Merville (1980) examine the daily closing prices of CBOE call options on six stocks during the calendar year 1976 and find that the BSM formula produces option values that are too high for out-of-the-money call options and too low for in-the-money call options, exactly opposite the bias reported by Black and Whaley. Such behavior, they contend, may be driven by the fact that the BSM model assumes a constant volatility rate throughout the life of the option. They suggest that stock price dynamics should be modeled as a constant elasticity of variance process described in Section 3. They cite anecdotal evidence that suggests that return volatility falls when stock prices rise to support their claim. When they use $\theta < 2$ and examine the pricing

[45] Recall that, in Section 2, it was shown that a call option on a non-dividend-paying stock will never optimally be exercised early.

errors of the CEV option valuation model, they find it fits better than the BSM formula and that the moneyness bias is reduced.[46]

Apparently perplexed by the reversal in sign of the moneyness bias, Emanuel and MacBeth (1982) gather an updated sample (calendar year 1978) of daily closing price data for the options on the same six stocks as MacBeth and Merville. Interestingly enough, they find that the moneyness bias reversed itself yet once again, with the pattern in pricing errors reverting back to that described by Black (1975).[47] The flip-flopping of the moneyness bias lead Emanuel and MacBeth to conclude that the CEV model with stationary parameters cannot explain the mispricing of call options any better than BSM. While the CEV model fits better than BSM formula in-sample due to the presence of an extra parameter, the movement in the parameter estimates undermines the model's usefulness.

Like the stock option prices (and as we will see shortly, currency option prices), stock index option prices exhibit moneyness biases, and the biases are not stationary through time. Whaley (1986) examines the prices of call and put options written on the S&P 500 futures option during the period January 1983 through December 1983 (i.e., the first calendar year of trading). Using the Barone-Adesi and Whaley (1987) approximation method to value these American-style options, he finds that out-of-the-money calls have model values that are too high and that in-the-money call options have model values that are too low.[48] Translated into implied volatility terms, this means implied volatility is a decreasing function of the option's exercise price. Sheikh (1991) examines BSM implied volatility patterns for S&P 100 index call options[49] during the period July 1983 through

---

[46] Based on the daily returns of 47 stocks during the period September 1972 through September 1977 (1,254 trading days), Beckers (1980) concludes that the constant elasticity of variance provided a better descriptor of stock price behavior than does the constant variance lognormal model.

[47] Rubinstein (1985) also documents the moneyness patterns found in the MacBeth and Merville (1980) and Emanuel and MacBeth (1982) studies.

[48] Naturally, due to put-call parity, Whaley documents the exact opposite bias for put options written on the S&P 500 futures.

[49] On one hand, using price data for S&P 100 index options is the best of the available alternatives, since, at the time, they were by far the most actively traded index options. On the other hand, computing accurately implied volatilities from S&P 100 index options requires handling multiple discrete cash dividends during the option's life (see Harvey and Whaley (1991)). Such an exercise, while computationally burdensome, is mandatory. Moreover, the S&P 100 index options have a wildcard feature that allows the option holder to wait until 3:15PM to decide upon exercise while the settlement proceeds

December 1985. He divides his sample into three sub-periods and documents a different moneyness bias in each. In the first sub-period, for example, he finds that call option's implied volatility falls monotonically with its exercise price across option maturities. In the second sub-period, the relation is more complex: short-term options had a "smile" shape, with the at-the-money option having the lowest implied volatility, while longer term options' volatilities decreased with exercise price. Finally, in the third sub-period, period, he finds an implied volatility smile for all option maturities.

The moneyness biases for foreign currency options are also non-stationary. Bodurtha and Courtadon (1987) examine pricing errors of five sets of foreign currency options traded on the PHLX during the period February 1983 through March 1985. They find that in-the-money calls (out-of-the-money puts) are over-priced and out-of-the-money calls (in-the-money puts) are under-priced. This means that option-implied volatility is a decreasing function of the option's exercise price. Shastri and Tandon (1986a) find similar evidence for the CME's call option on Deutschemark futures during the period February 1984 through December 1984. Hsieh and Manas-Anton (1988), on the other hand, find evidence of a U-shape relation between implied volatility and exercise price for Deutschemark futures option prices during the period January 23, 1984 through October 10, 1984. Similarly, Bates (1996) provides evidence a "volatility smile" for Deutschemark options in sample periods after 1988. Bollen and Rasiel (2002) measure explicitly time variation in the slope of the implied volatility function derived from OTC currency option quotes from 1998. In this single year, options on British Pounds Sterling and Japanese Yen exhibit both symmetric and asymmetric patterns, with significant changes from week to week.

5.2 Trading simulations

Under the BSM option valuation assumptions, a mis-priced option, delta-hedged over its remaining life, should provide a risk-free return different from the prevailing risk-free rate. This proposition underlies all of the trading simulation tests of option

---

are established at 3:00PM. Fleming and Whaley (1994) show that the wildcard option may have significant value, and increases in value as the option goes further in the money. Sheihk's use of the BSM formula to compute implied volatilities most assuredly accounts for some part of the positive relation between implied volatility and moneyness. It does not, however, explain the smile observed in the last sub-period.

valuation models, including that of Black and Scholes (1972). The strategies take on various forms, as discussed below. The basic procedure involves selling over-priced and buying under-priced options, simultaneously hedging the positions with a position in the underlying asset so that the net portfolio delta equals zero. The position is then held until the option's expiration, with the zero-delta hedge being maintained at the close each day by adjusting the asset position. At the end of the option's life, profits before and after transactions costs are tallied.[50]

Black and Scholes (1972) design the first trading simulation test of an option valuation model. They use a sample of 2,039 six-month call option transactions on 545 NYSE securities taken from the diaries of an option broker for the period May 1966 through July 1969 (766 trading days).[51] BSM formula values are computed each day. The volatility parameter is based on the daily returns of the underlying stock over the past year. The six-month commercial paper rate is used as a proxy for the risk-free rate of interest. In all, they implement four trading strategies, each one involving delta-hedging the mis-priced option using the underlying stock. To assess whether the valuation model values are, on average, too high or too low, they conduct a strategy whereby all calls are purchased at model values. To assess whether the option writer's premium is too high or too low on average, they conduct a second strategy whereby all calls are purchased at market prices. To test whether or not the model should be used to value contracts, they simulate a third strategy whereby under-priced calls are purchased and over-priced calls are sold at model values. Finally, to test whether or not profit opportunities existed in the option market over the sample period, they conduct a final test whereby under-priced

---

[50] Trading simulations cannot exactly portray the BSM world in that (a) portfolios must be rebalanced discretely rather than continuously, (b) investors face significant trading costs, (c) contracts are indivisible, and (d) future volatility is not known. Nonetheless, tests can manage the effects of these market imperfections. First, the contract-indivisibility constraint usually arises because researchers typically buy or sell a single mis-priced option and hedge using a fractional number of units of the underlying asset. The remedy is simple—buy or sell more option contracts in the simulation. Second, while discrete-rebalancing and imperfect volatility foresight can and do undermine hedging effectiveness, researchers can and do handle the problem by risk-adjusting reported profits based on the realized hedge portfolio volatility. Finally, trading costs are known, and are usually implemented directly into simulation analysis. Figlewski (1989) carries out an extensive set of Monte Carlo simulations showing the effects of each of these factors on the expected return and risk of the hedge portfolio. Boyle and Emanuel (1980) show how discrete-rebalancing affects the properties of the hedge portfolio return distribution.

[51] Recall the Chicago Board Options Exchange did not begin trading stock options until April 1973.

calls are purchased and over-priced calls are sold at market prices. Most subsequent empirical studies implement only the last procedure.

Their results are interesting in a number of respects. First, buying options at model values produces trading profits that are not significantly different from zero. In other words, the model produces unbiased estimates of option prices. This implies, of course, that the BSM implied volatility is an unbiased predictor of future realized volatility for the stocks in their sample. Second, buying options at market prices produces significant losses. This stands to reason since the trades contained in the sample are only sales by the option broker (i.e. customer buys). The option broker, of course, earned significant profits. The fourth strategy of buying under-priced and selling over-priced calls using market prices produces significant profits.

The third strategy of buying under-priced calls and selling over-priced calls at model prices produces significant losses. The explanation for this is subtle. The estimates of volatility based on historical returns have measurement error. Holding other factors constant, calls on high-volatility stocks tend to appear under-priced because their volatility estimates are higher than they should be. Buying these options at market prices will tend to produce losses. Conversely, calls on low-volatility stocks will tend to appear over-priced because their volatility estimates are lower than they should be. Selling such calls at market prices will tend to produce losses. To test this explanation, Black and Scholes re-run the strategy using the realized rate of return volatility over the life of the contract. The profits from the third strategy become insignificantly different from zero.[52] Interestingly, using the realized volatility over the life of the call made the profits from the fourth strategy larger and more significant. This is analogous to saying the implied volatility from the Black-Scholes formula is a better predictor of future realized volatility than is past realized volatility.

Whaley (1982) simulates a similar trading strategy using closing prices of CBOE call options written on 91 different dividend-paying stocks during a 160-week period from January 1975 through February 1978. Prices for options and stocks, as well as T-bill

---

[52] Karolyi (1993) documents the same type of behavior for a sample of CBOE call options written on 74 different stocks during the period January 1984 through December 1985. His test results show that the mean squared prediction error is reduced when a Bayesian shrinkage estimator is used.

rates (which proxy for the risk-free rate of return), are drawn from the *Wall Street Journal*. Dividend information and stock returns are drawn from the CRSP daily files. All options are valued using an average implied volatility for the option class from the previous week. All over-priced options on all stocks are sold, and all under-priced options on all stocks are purchased. Weights are assigned to the options in each portfolio so that both the dollar investment and the systematic risk of the option portfolios are the same.[53] Under this scheme, the expected rate of return of the hedge portfolio is zero (i.e., the equilibrium return on a portfolio with no risk and no capital investment). The positions are liquidated at the end of each week, and new positions are established. Over the sample period, the mean hedge portfolio return is 2.46% per week and is statistically significant. After trading costs equal to the bid/ask spread are applied, however, the profits disappear.

Trading simulation tests have also been performed for index options as well as foreign currency options. Whaley (1986) conducts a trading simulation using all S&P 500 futures option trades reported in the CME's trade and quote file during the period January 1983 through December 1983 (i.e., the first year of trading in this contract market). His strategy involves selling all over-priced options and buying all under-priced options. The volatility used in valuing the options is the at-the-money implied volatility from the previous day. Each option position is then delta-hedged and held to expiration. Any subsequent movement in the delta is corrected at the close each day by re-aligning the number of futures in the hedge portfolio. Overall, the before transaction-cost profits from the trading strategy are both positive and statistically significant. He also computes the breakeven trading cost rate and finds that, while it is below the rate a retail customer would face, it is well above the rates market makers would face. It is also interesting to note that when he categorizes the options by moneyness and option type (i.e., call or put), the out-of-the-money puts have by far the largest trading profits, followed by the in-the-money calls. This is not surprising considering the downward-sloping implied volatility function usually found in the S&P 500 futures options market.

[53] Whaley uses the Sharpe (1964)-Lintner (1965) capital asset pricing model to create his hedge portfolio. To estimate the call option's beta, he takes the product of the estimated beta of the underlying stock and the option's elasticity with respect to stock price.

Shastri and Tandon (1986a) also use the CME's trade and quote information to perform a trading strategy test for German mark futures options during the period February 1984 through December 1984. They find that significant excess trading profits, even after the incorporation of modest trading costs. Concerned that the trade may not be executable at the prices that signaled the profit opportunity, they repeat the simulation exercise by executing the trade at the next available trade prices. This delay causes the trading profits to disappear. Using the next trade price likely overstates the case, however, since the trade and quote data contain only prices that have changed from the previously recorded prices. Successive trades at the same price do not appear. Nonetheless, the evidence clearly indicates that the profit opportunities are fleeting.

Bollen, Gray and Whaley (2000) examine the trading profit opportunities in the PHLX's markets for options on British pounds, German marks, and Japanese yen during the period February 1983 through May 1996. They conclude that in these currency option markets, trading profits are better identified using a regime-switching option valuation model than using the standard models such as BSM. Reasonable trading costs mitigate the viability of the trading strategy, however.

5.3 Informational content of implied volatility

A good deal of energy has been focused on the information content of implied volatility in the empirical options literature. While not talking about implied volatility *per se*, Black and Scholes (1972) observe that their option valuation formula better described the cross-sectional structure of observed option prices when realized volatility over the option's life is used as the volatility parameter in the model rather than the realized volatility over the year preceding the cross section. This implies that the BSM implied volatility is a better predictor of future realized return volatility than is past realized return volatility. A number of direct investigations of this phenomenon ensued.

The first study to appear is by Latane and Rendleman (1976). They use closing option and stock prices from the *Wall Street Journal* for 24 firms whose options traded on the CBOE during the period October 1973 through June 1974. Weekly price observations are used, and each option class is required to have prices for at least two option series each week. Because options on the same stock but at different exercise prices have different implied volatilities, Latane and Rendleman create a weighted implied standard deviation (WISD) by weighting each series' estimate by its "vega" (i.e., the partial derivative of the option price with respect to volatility). They then compare the WISDs across stocks with estimates of realized volatilities for the period before, during, and after the sample period. They discover that implied volatility is more highly correlated with concurrent and subsequent realized volatility than historical volatility.

Chiras and Manaster (1978) examine the predictive power of stock option implied volatility for all stock options traded on the CBOE each month during the period June 1973 through April 1975. On day *t*, they calculate a weighted-average implied volatility from all options series on a particular stock. Next they compute the standard deviations of realized returns over the past twenty trading days, and the next twenty trading days. Realized volatility over the next twenty days serves as the dependent variable in a cross-sectional regression. The weighted average implied volatility and the historical realized volatility serve as regressors. They conclude that the cross-sectional implied volatility has become more informative over time based on the increasing R^2 values in the regression of future volatility on implied volatility. Based on the regressions that include both implied volatility and historical volatility as independent variables, they conclude that historical volatility is insignificant.

Beckers (1981) performs a similar cross-sectional regression using daily closing price data on CBOE stock options over a 75-trading day period in October 1975 through January 1976. His independent variables include historical volatility, the Latane and Rendlemen WISD, and the implied volatility of the at-the-money option. He concludes that at-the-money implied volatility predicts as well or better than the other alternatives. In addition, in contrast to Chiras and Manaster, Beckers find that historical volatility provides incremental explanatory power when included in the same regression with at-the-money implied volatility.

For options on assets other than common stocks, the information content of implied volatility is assessed in a time-series fashion. Day and Lewis (1992) compare the implied volatility of S&P 100 index option prices to GARCH and EGARCH models of conditional volatility over a 319-week period from November 1983 to

December 1989. They conclude that S&P 100 options provide unbiased forecasts of future volatility but that the inclusion of GARCH and EGARCH volatility assessments contains additional information. Harvey and Whaley (1992) develop and test a conditional market volatility prediction using S&P 100 index option prices during the period March 1983 through December 1989. They find that, although volatility changes are predictable in a statistical sense, the profits generated by simulated trading of S&P 100 index options based on the volatility predictions are insignificant after trading costs.[54] Fleming (1993, 1998) regresses first-differenced realized volatility (options' lifetime and 28-day) on first-differenced implied volatility using daily transaction data over the period October 1985 through April 1992, excluding the 1987 crash period. He concludes that implied volatility is a biased but substantially informative forecast of future volatility. He also examines the profits from trading volatility straddles (i.e., buying an at-the-money call and put) on the S&P 100 index, and reports that apparent trading profits disappear after reasonable trading costs are imposed. In other words, the bias in implied volatility is not economically significant.

The CBOE computes intraday levels of implied volatilities for S&P 100 and NASDAQ 100 index options. They are disseminated under the ticker symbols "VIX" and "VXN," respectively. The methodology used to create the volatility indexes is described in Whaley (1993). Fleming, Ostdiek and Whaley (1995) examine the time series properties of the VIX over the seven-year period 1986 through 1992 on a daily and weekly basis. They report that changes in the VIX have a strong inverse and asymetric contemporaneous association with the returns of the S&P 100 index. On days that the S&P 100 rises the VIX falls, but on days that the S&P 100 index falls the VIX rises by even more in absolute terms. This evidence is consistent with the work of Schwert (1989, 1990), who finds asymetry in the relation between stock returs and expected volattility, that is, the increase in expected volatility corresponding to a given negative stock market return is larger than the decrease in expected volatility corresponding to a similar size positive return. Schwert (2002) examines the contemporaneous relation between changes in the VIX and the VXN, the dynamics of which he ascribes is due, in part, to IPO activity.

---

[54] Lamoureux and Lastrapes (1993) conduct a similar analysis using stock option implied volatilities.

Shastri and Tandon (1986b) compare the volatility predictions using daily closing prices British pound, German mark, Japanese yen, and Swiss franc options traded on the PHLX during the period December 1982 through February 1984. To predict future realized volatility, they use: (a) the historical realized volatility, (b) the weighted-average implied volatility advocated by Latane and Rendleman (1976), and (c) the implied volatility of the at-the-money option. For British pounds and Japanese yen, the WISD performed best using a goodness-of-fit criterion, the historical estimator for German marks, and the at-the-money implied volatility for Swiss francs. The results indicate that the implied volatility is biased, however. In contrast, Lyons (1988) uses weekly price observations (from a transaction data base of PHLX currency options) to compare the at-the-money implied volatility to historical volatility of exchange rates for the Deutschemark, pound and yen during a sample period July 1983 through May 1986. He finds no difference on average between the levels of implied volatility and historical realized volatility. Jorion (1995) examines Deutschemark options over the period January 1985 to February 1992. He finds that implied volatilities are almost unbiased forecasts of the next day's absolute return, but are slightly biased forecasts of the volatility over the option's life. He also finds that neither historical volatility nor GARCH-based volatility assessments provide additional forecast power, however.

Finally, a recent study by Ederington and Guan (2000) examines the informational content of implied volatilities at different exercise prices. Using S&P 500 futures options over the period January 1988 through April 1998, they regress return volatility estimated over the option's life separately on implied volatilities of options with different exercise prices. They find that the "best" predictor of realized volatility is out-of-money calls/in-the-money puts, as evidenced by the least amount of bias and the highest correlation. Recall that earlier we discussed that fact that over the past ten or fifteen years, the implied volatilities of S&P 500 futures options have been a decreasing function of exercise price, but the slope of the relation changes through time. This suggests that the prices of out-of-the-money puts/in-the-money calls are driven by factors other than volatility.

5.4 Summary and analysis

The studies described in this section provide a number of stylized facts regarding the performance of option valuation models.

1) The shape of the BSM implied volatility function[55] (IVF) is not stationary through time for stocks, stock indexes, or currencies. Sometimes it appears as a smile, with at-the-money options having the lowest implied volatility. At other times it is downward sloping in exercise price. Yet, at other times still, the relation between implied volatility and exercise price is different for different option maturities.

2) Certain categories of options appear to generate higher abnormal trading profits than other categories of options. In particular, selling out-of-the-money index puts has been shown to consistently generate significant risk-adjusted profits before trading costs, even well before the October 1987 market crash.

3) Implied volatility appears to be an upward biased estimate of future volatility for index options, and, perhaps, for stock options. The economic significance of this bias appears small, however. For foreign currency options, implied volatility appears to be unbiased.

Of these results, the absence of a flat IVF is probably the most perplexing. A number of studies argue that the "volatility smiles" have appeared because asset prices do not follow the assumed geometric Brownian motion with constant volatility. Under geometric Brownian motion, the conditional risk-neutral density function of the underlying asset price is lognormal (or, alternatively, the distribution of asset returns is normal). The fact that the IVF is not a horizontal line, it is argued, indicates that the risk-neutral density function has different skewness and kurtosis parameters than the lognormal distribution. A downward sloping IVF indicates that the risk-neutral distribution is more negatively skewed than the BSM model assumes. A smile-shaped IVF indicates that the risk-neutral density functions is leptokurtic (i.e., has fatter tails). Under this explanation, the theoretical challenge is to find a stochastic process capable of generating the moments of the risk-neutral distribution that match those implied by option prices.[56]

One family of models advanced under this explanation are the *deterministic volatility function* (DVF) models. Derman and Kani (1994), Dupire (1994), and Rubinstein (1994) develop variations of a model that assumes the *local volatility rate* of the index is a function of the index level and time. Rather than positing a structural form for their DVF, they search for a binomial or trinomial lattice that achieves an *exact* cross-sectional fit of reported option prices (or, alternatively, an exact match of the moments of the risk-neutral distribution). This model cannot be tested in-sample, since all pricing errors are equal to zero. To test such a model, it is necessary to use out-of-sample data.

Dumas, Fleming, and Whaley (1998) perform such an experiment for S&P 500 index options during the period June 1988 through December 1993. They conclude that, although there is unlimited flexibility in specifying the DVF and it is always possible to describe exactly the observed structure of option prices, a parsimonious model works best using in-sample according to the Akaike Information Criterion. More importantly, they also show that, when the fitted volatility function is used to value options one week later, the DVF model's prediction errors grow large as the volatility function becomes less parsimonious. These results imply that models such as the DVF model are vulnerable to over-fitting the data.

In an attempt to evaluate the economic content of the DVF parameter estimates, Dumas *et al* evaluate the predictive performance of an *ad hoc* implementation of the BSM model that smoothes the implied volatilities of the S&P 500 options across exercise prices and times to expiration,[57] and then uses the estimated IVF to calculate option

---

[55] The *implied volatility function* is defined as the relation between an option's BSM implied volatility and the option's exercise price and time to expiration.

[56] This argument presupposes that there are no trading costs or other impediments to hedging one option against another. A number of studies provide methods for estimating the moments of risk-neutral distribution for arbitrary stochastic processes, an idea that originally appears in Breeden and Litzenberger (1978). Jarrow and Rudd (1982) are the first to attempt such estimations, and Corrado and Su (1996) improve upon the Jarrow and Rudd methodology. Jackwerth and Rubinstein (1996) and Bondarenko (2000) recover risk-neutral densities using nonparametric approaches. Bakshi, Kapadia, and Madan (2001) and Dennis and Mayhem (2002) use the technique for inferring the moments of risk-neutral stock return distributions.

[57] This "model" is intended to mimic the practice of market makers.

values one week later. Surprisingly, the *ad hoc* model outperforms all of the DVF models that they consider.[58]

Another family of models that can explain the absence of a flat IVF is option valuation models based on stochastic volatility. Stochastic volatility models can generate a downward sloping implied volatility function through volatility innovations that are negatively correlated with index returns. Bakshi, Cao, and Chen (1997) advocate the use of a stochastic volatility model with jumps for valuing S&P 500 index options. They conduct a comprehensive empirical study on the relative merits of competing option valuation models based on stochastic volatility (SV), stochastic interest rates (SI), and random jumps in the asset price (J). In all, their "SVSI-J" model has eleven parameters, which they estimate each day using a cross-section of S&P 500 index option prices. The sample period extends from June 1988 through May 1991. As in the experiments of Dumas, Fleming, and Whaley, the parameter estimates are obtained by minimizing the sum of squared errors between the last bid-ask quote midpoint and the model value.[59] In-sample pricing errors, out-of-sample pricing errors, and hedging errors are tabulated. Overall, their results appear to support the claim that a model with stochastic volatility and random jumps is a better alternative to the BSM formula with constant volatility across exercise prices.

The results of the Bakshi *et al* study are not conclusive, however. First, the fact that the more complicated models fit the cross-section of index option prices better than the BSM formula in-sample is simply a result of using more parameters. The Bakshi *et al* tests do not penalize the goodness-of-fit for the addition of more parameters. Second, out-of-sample pricing and hedging errors are computed only one day after the parameters are estimated. As such, they are practically in-sample and are subject to the same criticisms. In all likelihood, the *ad hoc* BSM model used by Dumas *et al* would have performed as well in such a test design. Both approaches are doing nothing more than indirectly smoothing a function through the cross-section of option prices, and the estimated surface happens to be fairly stable over short intervals of time. More frequent recalibration, in

[58] A recent study by Brandt and Wu (2002) reaches similar conclusions using cross-sectional data for European- and American style FT-SE 100 index options.

principle, should bring the prediction errors and hedging errors of the two methods closer together. This is not the issue, however. The objective of the study is to identify the stochastic process governing index movements through time.

Third, the fact that the implied volatility smile is so steeply sloped (and changes through time) cannot be reconciled with the parameters of the empirical distribution or risk aversion. Bakshi *et al* find that the volatility of volatility coefficient implied from options differs significantly from the one estimated directly from returns. Similarly, Bates (2000) examines the ability of a stochastic volatility model, with and without jumps, to generate the negative skewness consistent with a steep IVF. He finds that the inclusion of a jump process improves the model's ability to generate IVFs consistent with market prices, but that the parameter values are unreasonable. Searching for an economic explanation, Jackwerth (2000) attempts to recover risk aversion functions from S&P 500 index option prices, but winds up concluding that they are "irreconcilable with a representative investor." Using a new trading strategy test methodology, Bondarenko (2001) examines prices of out-of-the-money puts written on the S&P 500 futures during the period 1988 through 2000 and concludes the market is inefficient.

Finally, it is also worth noting that different estimation methods for identifying the parameters of the risk-neutral distribution can provide dramatically different parameter estimates. Campa, Chang, and Reider (1998) use three methods to estimate the parameters of the risk-neutral distribution from OTC currency option prices: (a) cublic splines based on the volatility-smoothing approach of Shimko (1993), (b) the implied binomial tree approach of Rubinstein (1994) (both untrimmed and trimmed), and (c) the mixture of lognormals approach of Melick and Thomas (1997). Using one-month ITL/DEM options in the period April 1996 through March 1997, for example, they find that the average implied kurtosis is 2.379 using the cublic spline method, 16.23 and 1.346 using the untrimmed and trimmed binomial trees, and 2.192 using the mixture of lognormals.

---

[59] An important implementation issue that goes largely unnoted is that available numerical methods cannot (and do not) guarantee a global minimum in an estimation problem with so many parameters.

Based on these anomalies, spending additional resources developing more elaborate theoretical models (with even more parameters) and more sophisticated computational techniques seems imprudent, at least in the short-run. A more promising avenue of investigation, perhaps, is the study of the option market participants' supply and demand for different option series in different option markets. One way to think of the IVF is as a series of market clearing option prices quoted in terms of BSM implied volatilities. In the BSM model, dynamic replication ensures that the supply curve for all option series in a given class is a horizontal line. No matter how large the demand for buying options, price and implied volatility are unaffected. In reality, however, a market maker will not stand ready to sell an unlimited number of contracts in a particular option series. As his position grows large, so do his expected hedging costs, not only in the form of direct trading costs such as the bid/ask spread of other options he will need to hedge, but also in the very availability of the other option series needed to hedge, so-called "limits to arbitrage."[60]

To help distinguish between the "stochastic process" and the "buying pressure" explanations for the shape of the IVF, consider Figure 1 in which the average daily implied volatility for S&P 500 index options and the average daily implied volatility for twenty stock options over the period January 1995 through December 2000 is plotted by moneyness category. The data are drawn from Bollen and Whaley (2002). The stock option classes considered are the twenty most active that traded continuously on the CBOE during the sample period. The underlying stocks have both high market-capitalization and highly liquid markets. Category 1 includes deep out-out-of-the-money puts and deep in-the-money calls. Moneyness is based on the option's delta.[61] Category 1 contains puts with deltas above -.125 and calls with deltas above .875. Calls and puts are included in the same category since there implied volatilities are linked through put-call parity. Category 2 include out-of-the-money puts (-.125 to -.375)/in-the-money calls (.625 to .875), category 3 contains at-the-money puts (-.375 to -.625) and calls (.375 to .625), and so on.

[60] See Shleifer and Vishny (1997).

[61] In comparing IVFs across option series and across option classes, it is necessary to account for differing times to expiration and volatility rates in the definition of moneyness.

With respect to distinguishing between the competing hypotheses, Figure 1 has two important features. First, the familiar IVF sneer appears for S&P 500 index options. Index option implied volatility decreases monotonically as the exercise price rises. They range from 26.2% for category 1 options to 16.9% for category 5 options, about 970 basis points in total. Second, the IVF for stock options appears as a smile, and its range is less than 400 basis points—from 35.6% for Category 3 (at-the-money) options to 39.4% for category 5 options. Since the shape of the IVF is tied to the moments of the risk-neutral distribution of the underlying asset, the IVFs shown in Figure 1 suggest that the return distribution for the S&P 500 index is highly skewed to the left while the return distribution for typical stocks options is leptokurtic (i.e., has fat tails). As Bollen and Whaley show (and is reproduced in Figure 2), however, the cumulative standardized empirical return distributions of the S&P 500 index and the individual stocks are <u>not</u> discernibly different from one another in terms of skewness, although both seem to have fatter tails than the normal. The subtle differences shown in Figure 2 can be translated into BSM implied volatilities by computing hypothetical risk-neutral put option prices based on the empirical distribution and then BSM implied volatilities from the hypothetical put prices. Figure 3 shows the resulting IVFs. The figure clearly shows the effects of the leptokurtosis in the empirical distributions as the IVFs are smiled-shaped. On the left-hand side of the effects, the slight differences in the skewness of the distribution appear. Overall, however, the proposition that the shape of the IVF is driven by a failure to identify the appropriate stochastic processes governing the movements of the asset price and volatility seems unjustified.

Figure 4 provides more clues about what may be the root cause. Plotted are the average differences between the implied volatilities of Figure 1 and the realized historical volatility for each asset over the most recent sixty trading days. Again, the data are drawn from Bollen and Whaley. Interestingly, all of the differences between implied and historical volatility for S&P 500 index options are greater than zero, with the greatest difference being for category 1 options (deep out-of-the-money puts/deep in-the-money calls). This finding is consistent with Longstaff (1995), who finds that the implied index level from S&P 100 call option prices exceeds the observed level of the S&P 100 index. For stock options, on the other hand, the deviations across all moneyness categories average about zero. For the individual categories, the differences are positive only for Category 1 and Category 5 options and negative for the rest. One possible explanation for these results is the supply and demand conditions differ for different option series on different underlying assets. It is well known, for example, that institutional investors buy S&P 500 puts for portfolio insurance. Unfortunately, there are no natural counterparties to these trades, and market makers must step in and absorb the imbalance, sometimes taking ever larger positions in a particular option series. Since hedging these positions is more and more costly as the size of his position increases, the market maker has no choice but to raise prices. The fact that all S&P 500 implied volatilities are higher than historical volatility suggests that trading activity in index options is largely buyer-initiated. The fact that out-of-the-money puts have higher implied volatilities than at-the-money puts reveals the institutional preference for out-of-the-money puts. In contrast, the flatness of the stock option IVF in Figure 1 and the small differences between implied volatility and historical volatility for stock options in Figure 3 suggest that the public order flow for stock options is much more evenly balanced between buyers and sellers, with market makers absorbing less of the imbalance.

The above supply-demand argument relates to the average demand over time. If buying pressure explains the behavior of the IVF, the shape of the IVF should change through time as the demands for options at different exercise prices change. Such evidence has already appeared. Longstaff (1995), for example, finds that differences between the call option-implied and observed S&P 100 index level is related to trading costs and the level of trading activity, and Bates (1996) finds a strong relation between the slope of the IVF (implied risk-neutral skewness) and the relative trading activity in calls versus puts. Bollen and Whaley (2002) document that movements in the implied volatilities of S&P 500 index option and stock option series are strongly correlated with their net buying pressure.[62]

The evidence indicates that net buying pressure along with upward sloping supply curves affects the average shape of IVFs as well as its movements through time. Until

[62] Net buying pressure is defined as the total index-equivalent contract volume executed at the ask price less the total index-equivalent contract volume executed at the bid price.

these effects are better understood and quantified, the practice of using option prices to deduce the stochastic process of the underlying asset will produce questionable results.

6. SOCIAL COSTS/BENEFITS OF DERIVATIVES TRADING

A number of empirical studies have attempted to address issues related to the social cost/benefits of derivatives contracts. These studies fall into three categories. The first category includes studies that examine movements (either price or volatility) in the underlying asset market when a derivatives contract market is introduced, and the second category includes studies that examine movements in the underlying asset market when a derivatives contract expires. The final category examines the inter-temporal relation between price changes in the derivative and asset markets.

6.1 Contract introductions

In complete and frictionless markets, any new security can be synthesized from existing securities. Consequently, the introduction of derivative contracts should have no effect on the price or return volatility of the underlying asset. The studies that examine the effects of contract introduction focus primarily on stock options. The reason is that the sample of stock option introduction events is large, permitting more reliable statistical inference. Options on more than 2,000 stocks have been introduced at different times over a period now spanning nearly 30 years. In contrast, the number of futures contracts on different financial assets is quite small, and there is only one event associated with each contract introduction. If an investigator wanted to assess the effect of the introduction of long-term interest rate futures contracts on interest rate volatility, the sample would consist of only one observation, the CBT's introduction of the T-bond futures in August 1977. Separating the effects of contract introduction from other contemporaneous market events in such circumstances is virtually impossible.

The effects of stock option introductions are measured stock return volatility and/or on stock price movements. Below, studies in each category are discussed in turn. Before doing so, however, it is useful to consider certain institutional details regarding the stock options market in the U.S., as they have an influence on how to interpret the results.

6.1.1 Institutional factors

The Chicago Board of Trade launched the Chicago Board Options Exchange on April 26, 1973, making it the first exchange-traded options market in the world. At the time, only call options were listed on sixteen NYSE stocks. Between April 1973 and December 1976 four more option exchanges began making markets in options, however, again, only calls. Puts on 25 stocks began trading in June 1977, but, even then, only on an experimental basis. The Securities Exchange Commission (SEC) allowed each of the five stock options exchanges in the U.S. to trade puts on five stocks. Later in the year, however, the SEC declared a moratorium on option introductions while it reviewed the practices of the option exchanges and considered the economic impact of option trading. About three years later, the SEC lifted its moratorium and adopted a lottery system called the "Stock Allocation Plan." According to this plan, exchanges randomly selected options from a pool of stocks and were each granted an excusive franchise to trade options in those chosen stocks. The first options traded under the plan were listed June 2, 1980. On June 3, 1985, the first options on NASDAQ stocks were listed. These options were exempt from the SEC's allocation plan and have always been eligible for multiple listing. On January 22, 1990, the SEC abolished its allocation plan and allowed all options to be multiple-listed. A program to "roll out" the grandfathered option classes began in November 1992 and went quarterly through February 1995.

A second set of institutional factors that may have a bearing on the empirical results are the stock option listing criteria. The CBOE, for example, requires that the firm has (a) a minimum of seven million shares outstanding not including those held by insiders, and (b) a minimum of 2,000 shareholders. In addition, it requires that the stock (c) traded at least 2,400,000 shares in the last twelve months, and (d) closed at a market price of at least $7.50 per share for the majority of the business days during the last three months.[63] Note that the first criterion requires that the prospective stock have at least seven million shares outstanding *excluding* those held by insiders. This helps ensure the availability of shares to short sell the underlying stocks, which market makers may have to do to hedge an open option position. To identify candidates for listing, the CBOE

[63] Chicago Board Options Exchange *Constitution and Rules* (May 1995), Paragraph 2113.

monitors the trading activity of all stocks satisfying the listing criteria. Among the factors considered in gauging the market's potential interest are the stock's trading volume and return volatility. The higher the trading volume and the greater the return volatility, the greater the projected option activity. Once the CBOE decides to list options on a particular stock, it registers with the SEC. Trading begins a few days later. As a matter of courtesy, the exchange sends a letter informing the firm of its decision.

6.1.2 Changes in return volatility

A number of theoretical arguments on how option contract introduction affects the underlying stock's return volatility have been advanced. A report by Nathan Associates (1969), for example, suggests that opening stock option markets might divert speculative trading volume away from the stock market, thereby reducing stock market liquidity and increasing stock return volatility. Others argue that return volatility will fall, albeit for different reasons. One explanation is based on selection bias. Exchanges use historical return volatility as one of the selection criteria for choosing stock options to list (i.e., the higher the volatility, the greater the prospective option trading activity). Since stock return volatility tends to be mean-reverting, exchanges may be systematically listing options on stocks whose volatilities are temporarily higher than their long-run levels. Similarly, the use of this criterion will pick off stocks whose return volatility is high due to sampling error. In both cases, return volatility is expected to fall after option introduction.

A second measurement-based explanation is that, if the introduction of stock option markets allows stock market makers to hedge their inventories, the bid-ask spreads in the stock market should narrow after options listing due reduced inventory holding costs. With smaller spreads, there will be less bid-ask price bounce included in the measurement of stock return volatility.

Yet another measurement-based explanation is based on "noise." Noise, as defined in the context of Black (1986), is the difference between a security's observed price and its intrinsic (but unobservable) value. One source of noise, as has just been discussed, is the market maker's bid/ask spread. Another is that observed market prices do not convey everyone's opinions because the market is incomplete. To the extent that the introduction of options affords potential market participants new trading opportunities (i.e., certain investors may be attracted by the availability of increased financial leverage at low transaction costs and/or the ability to short sell) and that arbitrageurs link the prices of options to the underlying stocks, stock prices will become less noisy and return volatility will be reduced.[64]

None of the above arguments is particularly compelling. Moreover, they have little empirical support. Studies by Bansal, Pruitt and Wei (1989), Skinner (1989), and Damadoran and Lim (1991) report increased trading volume in the stock market following option listing, refuting the notion in the Nathan (1969) report that stock market liquidity would be reduced. In addition, Stephan and Whaley (1990) report that price changes in the stock market lead the option market by as much as fifteen minutes, which suggests that informed traders prefer depth and anonymity of the stock market. Fleming, Ostdiek, and Whaley (1996) show that the relative illiquidity of the option market creates a price effect for block trades, such that an informed trader would rationally prefer the stock market. Fedenia and Grammatikos (1992) document that bid-ask spreads in stock markets narrow after options listing. They do not, however, directly assess the degree to whether variance reduction was caused by some true economic phenomenon or by the reduction in spread.

The early empirical evidence examining the effects of option introduction on return volatility finds post-listing reductions. Trennepohl and Dukes (1979) estimate betas for optioned and non-optioned stocks for the periods 1970-73 and 1973-1976 and find that betas for optioned stocks decrease more than the betas of non-optioned stocks. Skinner (1989) finds that the total stock return variance falls by an average of 4.8% after options are introduced. Conrad (1989) estimates that the excess return variance decreases from 2.29% for the 200 days prior to listing to 1.79% for the 200 days after listing. Bansal, Pruitt and Wei (1989) find that variance drops by 6.4% after options are listed, while Damadoran and Lim (1991) detect a 20% drop.

---

[64] Grossman (1988) offers a similar argument regarding the usefulness of exhange-traded puts versus portfolio insurance. With dynamic portfolio insurance, the put is synthesized by trading the stocks in the underlying portfolio or an index futures. The absence of real put option trading prevents the dissemination to market participants of important information regarding expected future price volatility. The less the

More recent evidence suggests that there is no change in stock return volatility when options are introduced. Bollen (1998) examines a sample of 1,010 stock option listings ending in December 1992. Calls and puts are included, as well as exchange-traded and NASDAQ stocks. He carefully matches each stock with a control stock in the same industry, where each member of the control group has no options listed during the sample period. Over the full sample period, Bollen shows that there is no significant difference between the return variances of optioned stocks versus non-optioned stocks. Furthermore, Bollen examines the NYSE-AMEX subsample used by Skinner and finds that, while return variance fell for optioned stocks after introduction, it also fell for the control group stocks, with the difference between the return variances being insignificantly different from zero. For the post-1986 period, Bollen documents increased return variance for the optioned stocks as well as non-optioned stocks, with no significant difference between the groups.

6.1.3 Price effects

The primary argument supporting the notion that the introduction of stock option will induce a price reaction is one of market completeness. Suppose that the introduction of call options allows a subset of investors to take leveraged positions in an underlying stock that they were earlier impeded from doing. Assuming all of these investors step into the market at once and buy call options, their aggregate demand to buy would be met by market makers, who would delta-hedge their inventory by buying the stocks. If the aggregate demand for calls is large enough, stock prices will rise. Similarly, suppose that a subset of investors cannot short sell stocks or at least find it costly to do so. If put option markets are introduced, the excess demand to short sell will finally come to fruition through the purchase of puts. Market makers will then have to hedge their short put positions by short selling the stock, and the stock price will fall.

Conrad (1989) examines the price effect of option introduction from 1974 to 1980. Using CBOE and AMEX stock option introduction dates gathered from the *Wall Street Journal* during the period 1974 through 1980, she analyzes abnormal stock returns and finds that the introduction of individual stock options caused a permanent price

---

information being transmitted to liquidity providers, the more difficult for the market to absorb the trades increase of about 2% in the underlying security, beginning approximately three days before the introduction. While the size of the increase is statistically significant, it is not large from an economic standpoint. Round-trip trading costs at the time would have almost certainly have eliminated trading profits.[65]

The fact that the average return is positive, however, is consistent with the market completeness argument. Given her sample period, the option introductions falls squarely within the period in which the option exchanges listed only call options. The only put options in the sample are the puts introduced between June 1977. Thus, a potential explanation for her reported stock price increase is price pressure exerted by the market maker as a result of hedging short call option positions. Sorescu (2000) confirms Conrad's findings for option introductions prior to 1980. For introductions after 1980, however, he finds a significant decrease. An explanation for the Sorescu results is that after 1980, call and put option introductions, for the most part, took place at the same time. Assuming the pent-up demand to short sell stocks supercedes the demand to buy, purchases of puts will exceed those of calls and the stock price will fall.[66]

In related work, Figlewski and Webb (1993) examine the introduction of put options in June 1977. Here the market completeness argument is probably stronger than in the case of calls since a significant number of investors cannot short sell stocks in a cost-effective manner. While Figlewski and Webb do not examine abnormal price effects *per se*, they report that (a) optionable stocks exhibit a significantly higher level of short interest than stocks without options, and (b) a stock's short interest increases after option listing.

6.2 Contract expirations

Studies of the effects of derivative contract expirations were motivated largely by public criticism of derivatives markets in the mid-1980s. This criticism was a consequence of abnormal price movements in the stock market during the "triple-witching" hour. The term, "triple," in triple-witching refers to the fact that, at the time, stock index futures, stock index options, and stock options all expired at the close of

---

implied by dynamic hedging strategies, and the higher the return volatility.

[65] The trading strategy would involve buying the stock and shorting the stock index portfolio.

trading on the third Friday of the contract month. Why stock options are included is not obvious since they are settled by delivery.[67] Stock index futures and options, on the other hand, are cash-settled, a feature that lies at the heart of the allegedly "abnormal" price movements.

To understand the phenomenon, consider the S&P 500 index futures contract. Over the life of a typical contract, index arbitrageurs accumulate large positions in the S&P 500 futures and the stocks comprising the S&P 500 index portfolio. These positions are usually carried into expiration, at which time the futures contract is cash-settled at the reported index level. But, in order for the index arbitrageur to exit his stock positions at the cash settlement value of the index, he must place market-on-close orders to buy or short the stocks in the S&P 500 index portfolio, depending upon whether his net stock position is short or long. When arbitrageurs exit these positions in unison on the quarterly expirations of the S&P 500 futures market, the reported level of the index moves.

Stoll and Whaley (1987) investigate these price movements for quarterly and monthly contract expirations for all contract expirations since contract market inception. They find that, while the reported index levels did move unpredictably in one direction or the other in the last half-hour of trading on the quarterly expirations, the size of the move is generally no more than should be expected given the size of the bid/ask spreads of the underlying stocks. A half-hour before the market close, the reported level of the index is based on the last trades of the index's component stocks. Absent significant news in the marketplace, the last trades of about half the stocks occurred at a bid price and the last trades of about half the stocks occurred at an ask price, so the reported index level is trading at about a midpoint. At the market close, index arbitrageurs must unwind with market orders to buy or to sell, driving trade prices of all of the stocks in the index to an ask or bid. This movement from the midpoint to one-side of the spread constitutes a .25% price move for the S&P 500 index more than half of the average price move observed during the last half hour of trading on expiration days, .40%. The incremental .15% is an

---

[66] See also Danielson and Sorescu (2001).

[67] Klemkosky (1978) examines the stock returns in the week before and the week after stock option introductions.

additional liquidity cost, no different in spirit (or in size) from the liquidity cost incurred in a large block trade.

In spite of the fact that at the time the evidence documented abnormal trading volume but no abnormal price movements at the close on expiration day, the CME changed the cash settlement from the close on expiration day to the open beginning with the June 1987 contract. Stoll and Whaley (1991) investigate the impact of this change and find the size price reversal did not change relative to their earlier study—only its location during the day. Karolyi (1996) examines Nikkei 225 futures contract expirations, and, like Stoll and Whaley, concludes that the expiration of the Nikkei 225 futures induces abnormal trading volume but economically insignificant price effects. Stoll and Whaley (1997) find similar results for the Australian All Ordinaries Share Price Index futures and options expirations, as do Bollen and Whaley (1999) for the Hang Seng index futures in Hong Kong.[68]

6.3 Market synchronization

In frictionless markets, derivative contracts are redundant securities. Hence, the price changes of derivative contracts should be perfectly positively correlated with the price changes of their underlying assets. With friction, however, the situation may change.

6.3.1 Stock market versus option market

Early evidence in the stock option market indicates that price changes in the option market tend to lead those in the stock market. Manaster and Rendleman (1982), for example, analyze close-to-close returns of portfolios based on the relative difference between stock prices and stock prices implied by option prices. They conclude that closing option prices contain information that is not contained in closing stock prices. Furthermore, they claim that it takes up to one day for the stock market to adjust. This evidence is consistent with the view that the extreme leverage provided by stock option contracts makes them the preferred trading vehicle for informed traders. The use of closing price data, however, seriously undermines the interpretation of Manaster and

[68] Day and Lewis (1988) examine changes in implied volatility of index options in the days surrounding contract expiration.

Rendleman results. The stock option market closes at 3:10PM (CST), ten minutes after the close of the stock market. Information that Manaster and Rendlemen ascribe to option prices and not stock prices may only be information that has disseminated into the marketplace between closing times in the two markets.

Bhattacharya (1987) uses observed intraday bid/ask call prices to compute implied bid/ask stock prices. These prices are then compared to actual bid/ask stock prices in order to identify arbitrage opportunities. The stock is considered under-priced (over-priced) if the implied bid (ask) is higher (lower) than the actual ask (bid). A simulated trading strategy based on these arbitrage signals indicates that profits are insufficient overcome trading costs for all intraday holding periods. He also confirms the Manaster and Rendleman results by documenting statistically significant excess returns for overnight holding periods (after the costs of transacting but before information search and exchange seat opportunity costs).

A flaw in Bhattacharya's test design is that it can only detect whether the option market leads the stock market and not vice versa.[69] In order to test for the possibility of stock prices leading option prices, he should have used observed bid/ask stock prices to compute implied bid/ask call prices to identify over- (under)-pricing. Although he shows that option price changes have some predictive power, his results do not preclude the possibility that the stock price changes have enormous predictive power on option price changes.

Anthony (1988) uses daily data to examine whether trading in the option market causes trading in the stock market (or vice versa).[70] He concludes "…trading in call options leads trading in the underlying shares, with one day lag." His results are subject to the same caveats as Manaster and Rendleman due to the non-simultaneity of closing times in the two markets. Moreover, his evidence is not overwhelming. He finds that (a) option volume leads stock volume for 13 firms, (b) stock volume leads option volume for 4 firms, and (c) there is no unambiguous direction of causality for the remaining 8 firms.

---

[69] Bhattacharya recognizes this problem but does not perform the simulations in the reverse way.
[70] Anthony uses the terms "cause" and "causality" in the sense of Granger (1969).

Stephan and Whaley (1990) use intraday transaction data to re-examine the price change and trading volume results. Their sample consists of all trades of CBOE call options on individual stocks during the first quarter of 1986, and all individual stock trades from the NYSE. From the trade price files, they compute five-minute stock price changes using the NYSE data, and five-minute implied stock price changes using the CBOE data. They also aggregate trading volume within each five-minute interval. When implied stock price changes are regressed on leading, contemporaneous, and lagged stock price changes, the stock market appears to lead the option market by as much as fifteen to twenty minutes on average. They divide the sample by intraday stock return and by option delta to see if option price changes tend to lead on days with large positive or negative stock returns due to the release of information. The results for the different sub-samples are qualitatively the same. Stephan and Whaley also perform the lead/lag test using number of trades and number of contracts/shares traded in each interval and find that the stock market's lead can be as much as 30 minutes or more.

Chan, Chung, and Johnson (1993) extend the Stephan and Whaley results in an important way. They use the same sample period, but, in addition to trade data, they use bid/ask quote data for the options as well as the stocks. Consistent with Stephan and Whaley, they find that when trade prices are used to generate five-minute price changes, the stock market leads the option market by as much as fifteen minutes or more. When bid/ask midpoints are used to generate five-minute price changes, however, the prices in the two markets appear to move simultaneously. Taken together, these two results suggest that the documented lead of stock market is largely driven by stock options trading less frequently than stocks. The fact that the lead disappears when quote midpoints are used is not surprising considering that option market makers condition quotes on the prevailing stock price and, when the stock price changes, the quotes on all option series are updated automatically using an option valuation model.

6.3.2 Stock market versus index derivatives markets

More recent work on market synchronization has tended to focus on index derivatives. Stoll and Whaley (1990), for example, examine five-minute returns for the S&P 500 futures in relation to the returns the underlying S&P 500 index. On face appearance, the stock market appears to lag the futures market by about five minutes on average, but occasionally as long as ten minutes or more. But this lag is illusory. The observed level of the S&P 500 index is an amalgam of 500 last trade prices, where many stocks have not traded in some time. The observed index level, therefore, is always a stale indicator of the true index level.

This "infrequent trading" effect induces positive serial correlation in observed intraday index returns that does not appear in the returns of the index futures. To purge the infrequent trading effect, Stoll and Whaley model the observed index returns as an ARMA(2,3) process, and then correlate the S&P 500 index return innovations with the returns of the S&P 500 futures. They find pronounced lag that appears in raw returns disappears using return innovations and conclude the returns in the futures market and the stock market are, for the most part, contemporaneous. The mild evidence that continues to show that the futures market leads even after accounting for the effects of infrequent trading they ascribe to being support for the price discovery hypothesis that new market information disseminates in the market where it is least expensive to trade—in this case, the futures market. A number of subsequent empirical investigations re-examine the relation using more recent data, more refined measurement of the infrequent trading effect, and different index futures market domestically and internationally, and find that it continues to persist.[71]

The evidence regarding the lead-lag relation between stock index options and the stock market is less plentiful because of the need to convert index option price movements into index movements. Kleidon and Whaley (1992) examine intraday price movements of the S&P 500 index, the S&P 500 futures, and the S&P 100 index options in the days surrounding the October 1987 crash. They use five-minute price changes. To compute the five-minute price changes for the S&P 100 index options, they take all index option trade prices in a given five-minute interval and perform a non-linear regression to solve of the implied index level and volatility rate. The implied index levels are used to compute S&P 100 index returns, and the implied S&P 100 index returns are compared to the returns of the S&P 500 index and index futures. They find that in the first half of

October 1987 the return behavior of stocks, index and index options behave in a very similar fashion. During the October 1987 crash, however, the usual integration broke down but the problem lay in the stock market. The usual links between the futures and options markets lay largely intact. By contrast, both the futures and options markets were delinked from the cash market, which showed price levels much higher than those in either of the other markets. This result is not surprising considering that order queues in printers at the specialists' posts were up to seventy-five minutes long by noon on October 19th. On this day, the index derivatives prices were better reflections of the level of the stock market than the stock market itself.

6.4 Summary and analysis

The fact that equity derivatives markets are so active attests to the social benefits they provide. Call options offer the market participants increased leverage and limited liability. Put options offer similar advantages, but in addition provide some market participants (such as retail customers) with the ability to short sell without prohibitive restrictions. In addition, index futures and options are inexpensive ways for the public to trade portfolios of stocks.

Of the categories of studies in this section, the most promising area for future empirical research is in the area market completeness/trading restrictions. Studies of contract expirations reach the common conclusion that while trading volume is high, price movements are no larger than one would expect in block trades, and studies of market synchronization reach the conclusion that prices in derivatives and asset markets are closely integrated. The studies on stock option contract introductions, however, appear to indicate that stock prices may move as a result of option introduction. While this evidence supports the notion that stock options help complete the market, the evidence is surprising considering that (a) stock option trading activity pales by comparison to stock market trading in the best of times, and (b) stock options are seldom highly active from the outset. For stock prices to move significantly, stock option trading must be substantial. Showing the level of option trading (and open interest) relative to

---

[71] See, for example, Chan (1992), Wahab and Lashgari (1933), Shyy, Vijayraghavan, and Scott-Quinn (1996), Fleming, Ostdiek and Whaley (1996), Booth, So and Tse (1999), Frino, Walter, and West (2000),

stock trading (and shares outstanding) just after options are introduced would add to the credibility of the reported price movement results.

Along the same line, examining the time series behavior of differences between the put and call option-implied volatilities (or, equivalently, violations of put-call parity) in stock option markets also seems a worthwhile pursuit. As shown in D'Avolio (2002), stocks are frequently difficult to borrow, and rebate rates on borrowed stocks are frequently negative. In such circumstances, the implied volatilities of puts will exceed those of calls since the market maker is faced with extraordinary costs when shorting stocks to hedge short put positions. This fact was noted recently by Lamont and Thaler (2001) for stock options written on Palm after its spin off from 3Com, and is examined more systematically in Ofek, Richardson, and Whitelaw (2002). Much work remains to be done, however.

7. SUMMARY

The purpose of this chapter is to provide an overview of the evolution of derivatives contract markets and derivatives research over the past thirty years. The chapter has six complementary sections. The first section contains a brief history of derivatives contracts and contract markets. Although the origin of derivatives use dates back thousands of years, the most important innovations occurred only recently, in the 1970s and 1980s. Concurrent with these industry innovations was the development of modern-day option valuation theory. These advances are reviewed in the second and third sections. The key contribution is seminal theoretical framework of the Black-Scholes (1973) and Merton (1973) ("BSM") model. The key economic insight of their model is that a risk-free hedge can be formed between a derivatives contract and its underlying asset. This implies that contract valuation is possible under the assumption of risk-neutrality without loss of generality. Not only does this framework provide BSM with the ability to value standard call and put options, it has provided other researchers with the ability to value thousands of different derivatives contract structures such as caps, collars, floors, binary options, and quantos. Many of these contributions, as well as other extensions to the BSM model, are summarized.

The final three sections of this chapter summarize empirical work that investigates the pricing and valuation of derivatives contracts and the efficiency of the markets within which they trade. The studies are divided into three groups. In the first group are studies that focus on testing no-arbitrage pricing conditions. A review of tests of the no-arbitrage price relations between forwards and futures and their underlying assets as well as tests lower price bounds and put-call parity in the options markets is provided. The second group contains studies that attempt to evaluate option empirical performance of option valuation models. The approaches used include investigating the in-sample properties of option values by examining pricing errors or patterns in implied volatilities, examining the performance of different option valuation models by simulating a trading strategy based on under- and over-pricing, and examining the informational content of the volatility implied by option prices. The third and final group of studies focuses on the social costs and/or benefits that arise from derivatives trading. Through this review, one important fact emerges—the BSM model is one of the most resilient in the history of financial economics.

References

Abhyankar, A., 1998, Linear and nonlinear Granger causality: Evidence from the U.K. stock index futures market, *Journal of Futures Markets* 18, 519-540.

Anthony, J. H., 1988, The interrelation of stock and options market trading-volume data, *Journal of Finance* 43, 949-964.

Aristotle, 350 B.C. *Politics*. Benjamin Jowett edition. (Boston, MA: http://classics.mit.edu/Aristotle/politics.1.one.html).

Asay, Michael R., 1982, A note on the design of commodity option contracts, *Journal of Futures Markets* 2 (Spring), 1-8.

Bachelier, Louis, 1900, Theory of speculation, in English version appears in Cootner, Paul H. 1964, ed.: *The Random Character of Stock Market Prices*, (MIT Press, Cambridge, Massachusetts).

Bailey, Warren and Stulz, Rene, 1989, The pricing of stock index options in a general equilibrium model, *Journal of Financial and Quantitative Analysis* 24, 1-12.

Bakshi, Gurdip, Cao, Charles, and Chen, Zhiwu, 1997, Empirical performance of alternative option pricing models, *Journal of Finance* 52 (December).

Bakshi, Gurdip, Kapadia, Nikunj, and Madan, Dilip, 2001, Stock return characteristics, skewness laws, and the differential pricing of individual equity options, *Review of Financial Studies*.

Bansal, Vipal K. and Pruitt, Stephen W. Wei K. C. John, 1989, An empirical examination of the impact of CBOE option initiation on the volatility and trading volume of the underlying equities: 1973-1986, *Financial Review* 24, 19-29.

Barone-Adesi, Giovanni and Whaley, Robert E., 1987, Efficient analytic approximation of American option values, *Journal of Finance* 42 (June), 301-320.

Bates, David S., 1996, Dollar jump fears -- 1984-1992: Distributional abnormalities implicit in currency futures options, *Journal of International Money and Finance* 15, 65-93.

Bates, David S., 2000, Post-'87 crash fears in the S&P 500 futures options market, *Journal of Econometrics* 94, 181-238.

Beckers, Stan, 1981, Standard deviations implied in option prices as predictors of future stock price variability, *Journal of Banking and Finance* 5, 363-381.

Beckers, Stan, 1980, The constant elasticity of variance model and its implications for option pricing, *The Journal of Finance* 35 (3), 661-673.

Bhattacharya, Mihir, 1987, Price changes of related securities: The case of call options and stocks, *Journal of Financial and Quantitative Analysis* 22, 1-15.

Bhattacharya, Mihir, 1983, Transaction data tests of efficiency of the Chicago Board Options Exchange, *Journal of Financial Economics* 12, 161-185.

Black, Fischer, 1975, Fact and fantasy in the use of options, *Financial Analysts Journal* 31 (July/August), 36-41 and 61-72.

Black, Fischer, 1986, Noise, *Journal of Finance* 41, 529-543.

Black, Fischer, 1976, Studies of stock price volatility changes, *Proceedings of the 1976 Meetings of the Business and Economics Section, American Statistical Association*, 177-181.

Black, Fischer, 1976, The pricing of commodity contracts, *Journal of Financial Economics* 3 (March), 167-179.

Black, Fischer and Scholes, Myron, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637-659.

Black, Fischer and Scholes, Myron, 1972, The valuation of option contracts and a test of market efficiency, *Journal of Finance* 27 (2), 399-417.

Blomeyer, Edward C. and Boyd, James C., 1988, Empirical Tests of boundary conditions for options on Treasury bond futures, *Journal of Futures Markets* 4, 185-198.

Bodurtha, James N. and Courtadon, George R., 1986, Efficiency tests of the foreign currency options market, *Journal of Finance* 41, 151-162.

Bodurtha, James N. and Courtadon, George R., 1987, Tests of an American pricing model on the foreign currency options market, *Journal of Financial and Quantitative Analysis* 22 (June), 153-167.

Bollen, Nicolas P. B., 1998, A note on the impact of options on stock return volatility, *Journal of Banking and Finance* 22, 1181-1191.

Bollen, Nicolas P. B., Gray, Stephan A. and Whaley, Robert E., 2000, Regime-switching in foreign exchange rates: Evidence from currency option prices, *Journal of Econometrics* 94, 239-276.

Bollen, Nicolas P. B. and Rasiel, Emma, 2002, The performance of alternative valuation models in the OTC currency options market, *Journal of International Money and Finance*.

Bollen, Nicolas P. B. and Whaley, Robert E., 1999, Do expirations of Hang Seng index derivatives affect stock market volatility?, *Pacific-Basin Journal* 7, 453-470.

Bollen, Nicolas P. B. and Whaley, Robert E., 2002. Does price pressure affect the shape of implied volatility functions? Duke University.

Bondarenko, Oleg, 2001. On market efficiency and joint hypothesis. University of Illinois at Chicago.

Bondarenko, Oleg, 2000. Recovering risk-neutral densities: A new nonparametric approach. University of Illinois at Chicago.

Booth, G., So. R., and Tse, Y., 1999, Price discovery in the German equity index derivatives markets, *Journal of Futures Markets* 19, 619-643.

Boyle, Phelim P., 1988, A lattice framework for option pricing with two state variables, *Journal of Financial and Quantitative Analysis* 23 (March), 1-12.

Boyle, Phelim P., 1977, Options: A Monte Carlo approach, *Journal of Financial Economics* 4, 323-338.

Boyle, Phelim P and Emanuel, David, 1980, Discretely-adjusted option hedges, *Journal of Financial Economics* 8, 259-82.

Boyle, Phelim P., Evnine, J., and Gibbs, S., 1989, Numerical evaluation of multivariate contingent claims, *Review of Financial Studies* 2, 241-250.

Brandt, Michael W. and Wu, Tao, 2002, Cross-sectional tests of deterministic volatility functions, *Journal of Empirical Finance* (forthcoming).

Breeden, Douglas T. and Litzenberger, Robert H., 1978, Prices of state-contingent claims implicit in option prices, *Journal of Business* 51, 621-651.

Brennan, Michael J. and Schwartz, Eduardo S., 1977, The valuation of American put options, *Journal of Finance* 32, 449-462.

Campa, J., Chang, K., and Reider, R., 1998, Implied exchange rate distributions: Evidence from the OTC option markets, *Journal of International Money and Finance* 17, 117-160.

Carr, Peter, 1998, Randomization and the American put, *Review of Financial Studies* 11, 597-626.

Chan, K., 1992, A further analysis of the lead/lag relationship between the cash market and stock index futures market, *Review of Financial Studies* 5, 123-152.

Chan, Kalok, Chung, Y. Peter, and Johnson, Herb, 1993, Why option prices lag stock prices: A trading-based explanation, *Journal of Finance* 48, 1957-1967.

Chang, Carolyn W. and Chang, Jack S. K., 1990, Forward and futures prices: Evidence from the foreign exchange markets , *Journal of Finance* 45 (September), 1333-1336.

Chicago Board of Trade, 1994. *Commodities Trading Manual*. Chicago, IL: Chicago Board of Trade).

Chicago Board Options Exchange, 1995. *Constitution and Rules*. Chicago, Illinois: Chicago Board options Exchange).

Chiras, D. P. and Manaster, Steve, 1978, The information content of option prices and a test of market efficiency, *Journal of Financial Economics* 6, 213-234.

Conrad, Jennifer, 1989, The price effect of option introduction, *Journal of Finance* 44, 487-498.

Cornell, Bradford, 1981, The relationship between volume and price variability in futures markets, *Journal of Futures Markets* 1 (Fall), 303-316.

Cornell, Bradford and Reinganum, Mark R., 1981, Forward and futures prices: Evidence from the foreign exchange market, *Journal of Finance* 36, 1035-1045.

Corrado, Charles J. and Su, Tie, 1996, Skewness and kurtosis in S&P 500 index returns implied by option prices, *Journal of Financial Research* 19 (Summer), 175-192.

Cox, John C., Ingersoll, John E., and Ross, Stephen A., 1981, The relation between forward and futures prices, *Journal of Financial Economics* 9, 321-346.

Cox, John C. and Ross, Stephen A., 1976, The valuation of options for alternative stochastic processes, *Journal of Financial Economics* 3, 145-166.

Cox, John C., Ross, Stephen A., and Rubinstein, Mark, 1979, Option pricing: A simplified approach, *Journal of Financial Economics* 7 (September), 229-264.

D'Avolio, Gene, 2001. The market for borrowing stock. Harvard University .

Damadoran, Aswath and Lim, Joseph, 1991, The effects of option listing on the underlying stocks' return process, *Journal of Banking and Finance* 15, 647-664.

Danielsen, Bartley R. and Sorescu, Sorin M., 2001, Why do option introudctions depress stock prices? A study of diminsishing short sale constraints, *Journal of Financial and Qunatitative Analysis* (Forthcoming).

Day, Ted E. and Lewis, Craig M., 1992, Stock market volatility and the information content of stock index options, *Journal of Econometrics* 52, 267-287.

Day, Ted E. and Lewis, Craig M., 1988, The behavior of the volatility implicit in the prices of stock index options, *Journal of Financial Economics* 22, 103-122.

Dennis, Patrick and Mayhew, Stewart, 2002, Risk-neutral skewness: Evidence from stock options, *Journal of Financial and Quantitative Analysis* (forthcoming).

Derman, Emanuel and Kani, Iraj, 1994, Riding on the smile, *Risk* 7, 32-39.

Derman, Emanuel and Kani, Iraj, 1994. The volatility smile and its implied tree. Goldman Sachs.

Dumas, Bernard, Fleming, Jeff, and Whaley, Robert E., 1998, Implied volatility functions: Empirical tests, *Journal of Finance* 53, 2059-2106.

Dupire, Bruno, 1994, Pricing with a smile, *Risk* 7, 18-20.

Ederington, Louis H. and Guan, Wei, 2000. The information content of implied volatility and the "frown" in option prices. University of Oklahoma.

Emanuel, David C. and MacBeth, James D., 1982, Further results on the constant elasticity of variance call option pricing model, *Journal of Financial and Quantitative Analysis* 4 (November), 533-554.

Engle, R. F., Kane, A., and Noh, J., 1994, Forecasting volatility and option prices of the S&P 500 index, *Journal of Derivatives* 2, 17-30.

Evnine, J. and Rudd, A., 1985, Index options: The early evidence, *Journal of Finance* 40, 743-756.

Fedenia, Mark and Grammatikos, THeoharry, 1992, Options trading and the bid-ask spread of the underlying stocks, *Journal of Business* 65 (3), 335-351.

Figlewski, Stephen, 1984, Hedging performance and basis risk in stock index futures, *Journal of Finance* 39 (July), 657-669.

Figlewski, Stephen, 1989, Option arbitrage in imperfect markets, *Journal of Finance* 44 (December), 1289-1311.

Figlewski, Stephen and Webb, Gwendolyn P., 1993, Options, short sales, and market completeness, *Journal of Finance* 48 (June), 761-777.

Fleming, Jeff, 1998, The quality of market volatility forecasts implied by S&P 100 index option prices, *Journal of Empirical Finance* 5, 317-345.

Fleming, Jeff, 1993. The valuation and information content of S&P 100 index option prices. Ph.D. dissertation, Duke University.

Fleming, Jeff, Ostdiek, Barbara, and Whaley, Robert E., 1995, Predicting stock market volatility: A new measure, *Journal of Futures Markets* 15 (3), 265-302.

Fleming, Jeff, Ostdiek, Barbara, and Whaley, Robert E., 1996, Trading costs and the relative rates of price discovery in stock, futures, and options markets, *Journal of Futures Markets* 16 (June), 353-387.

Fleming, Jeff and Whaley, Robert E., 1994, The value of wildcard options, *Journal of Finance* 49, 215-236.

French, Kenneth R., 1983, A comparison of forward and futures prices, *Journal of Financial Economics* 12, 311-342.

Frino, A., Walter, T., and West, A., 2000, The lead-lag relationship between equities and stock index futures around information releases, *Jounral of Futures Markets* 20, 467-487.

Galai, Dan, 1979, A convexity test for traded options, *Quarterly Review of Economics and Business* 19 (Summer), 83-90.

Galai, Dan, 1978, Empirical tests of boundary conditions for CBOE options, *Journal of Financial Economics* 6, 187-211.

Garber, Peter M., 2000. *Famous First Bubbles: The Fundamentals of Early Manias*. Cambridge, Massachusetts: The MIT Press).

Garman, Mark B. and Kohlhagen, S. W., 1983, Foreign currency option values, *Journal of International Money and Finance* 2 (December), 231-237.

Geske, Robert, 1979, A note on an analytical formula for unprotected American call options on stocks with known dividends, *Journal of Financial Economics* 7 (October), 375-380.

Geske, Robert, 1979, The valuation of compound options, *Journal of Financial Economics* 7, 63-81.

Geske, Robert and Johnson, Herbert E., 1984, The American put valued analytically, *Journal of Finance* 39 (December), 1511-1524.

Goldman, B., Sosin, H., and Gatto, M. A., 1979, Path dependent options: Buy at the low, sell at the high, *Journal of Finance* 34 (December), 1111-1127.

Gould, John P and Galai, Dan, 1974, Transaction costs and the relationship between put and call prices, *Journal of Financial Economics* 1, 105-29.

Granger, C. W. J., 1969, Investigating causal relations by econometric models and cross spectral methods, *Econmetrica* 37, 424-438.

Gray, Stephan F. and Whaley, Robert E., 1999, Reset put options: Valuation, risk characteristics, and an application, *Australian Journal of Management* 24 (1), 1-20.

Grossman, Sanford, 1988, An analysis of the implications for stock and futures price volatility of program trading and dynamic hedging strategies, *Journal of Business* 61, 275-298.

Hammer, J. A., 1989, On biases reported in studies of the Black-Scholes option pricing model, *Journal of Economics and Business* 41, 153-169.

Harvey, Campbell R. and Whaley, Robert E., 1992, Market volatility prediction and the efficiency of the S&P 100 index option market, *Journal of Financial Economics* 30, 43-73.

Harvey, Campbell R. and Whaley, Robert E., 1991, S&P 100 index option volatility, *Journal of Finance* 46, 1551-1561.

Heston, Stephen L., 1993, A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* 6, 327-343.

Hsieh, David A. and Manas-Anton, Louis, 1988, Empirical regulatirites in the Deutschemark futures options, *Advances in Futures and Options Research* 3, 183-208.

Hull, John and White, Alan, 1987, The pricing of options on assets with stochastic volatilities, *Journal of Finance* 42, 281-300.

Jarrow, Robert A. and Oldfield, George S., 1981, Forward contracts and futures contracts, *Journal of Financial Economics* 9 (December), 373-82.

Jarrow, Robert A. and Rudd, Andrew, 1982, Approximate option valuation for arbitrary stochastic processes, *Journal of Financial Economics* 10, 347-69.

Jarrow, Robert A. and Rudd, Andrew, 1983. *Option Pricing*. Homewood, Illinois: Irwin).

Jeckwerth, Jens C., 2000, Recovering risk aversion from option prices and realized returns, *Review of Finacial Studies* 13, 433-451.

Jeckwerth, Jens C. and Rubinstein, Mark, 1996, Recovering probability distributions from option prices, *Journal of Finance* 51, 1611-1631.

Johnson, Herbert E., 1987, Options on the minimum and maximum of several assets, *Journal of Financial and Quantitative Analysis* 22 (September), 277-283.

Jorion, Philippe, 1995, Predicting volatility in the foreign exchange market, *Journal of Finance* 50, 507-528.

Kamrad, Bardia and Ritchken, Peter, 1991, Multinomial approximating models for options with k state variables, *Management Science* 37 (12), 1640-1652.

Karolyi, G. Andrew, 1993, A Bayesian approach to modeling stock return volatility for option valuation, *Journal of Financial and Quantitative Analysis* 28 (December), 579-594.

Karolyi, G. Andrew, 1996, Stock market volatility around expiration days in Japan, *Journal of Derivatives* 4 (Winter), 23-43.

Kleidon, Allan W. and Whaley, Robert E., 1992, One market? stocks, futures, and options during October 1997, *Journal of Finance* 47 (3), 851-877.

Klemkosky, Robert C., 1978, The impact of option expirations on stock prices, *Journal of Financial and Quantitative Analysis* 13, 507-518.

Klemkosky, Robert C. and Lasser, D. J., 1985, An efficiency analysis of the T-bond futures market, *Journal of Futures Markets* 5, 607-620.

Klemkosky, Robert C. and Resnick, Bruce G., 1979, Put-call parity and market efficiency, *Journal of Finance* 34 (December), 1141-1155.

Lamont, Owen and Thaler, Richard, 2001. Can the market add and subtract? Mispricing in tech stock carve-outs. Graduate School of Business, University of Chicago.

Lamoureux, Chris G. and Lastrapes, W. D., 1993, Forecasting stock-return variance: Toward understanding stochastic implied volatilities, *Review of Financial Studies* 6, 293-326.

Latane, Henry A. and Rendleman, Richard J., 1976, Standard deviations of stock price ratios implied in option prices, *Journal of Finance* 31, 369-381.

Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13-37.

Longstaff, Francis A., 1995, Option pricing and the martingale restriction, *Review of Financial Studies* 8 (Winter), 1091-1124.

Lyons, Richard K., 1988, Tests of the foreign exchange risk premium using the expected second moments implied by option pricing, *Journal of International Money and Finance* 7, 91-108.

MacBeth, James D. and Merville, Larry J., 1980, Tests of the Black/Scholes and Cox call option valuation models, *Journal of Finance* 35, 285-301.

MacKinley, Craig A. and Ramaswamy, Krishna, 1988, Index futures arbitrage and the behavior of stock index futures prices, *Review of Financial Studies* 1, 137-158.

MacMillan, Lionel W., 1986, Analytic approximation for the American put option, *Advances in Futures and Options Research* 1, 119-139.

Manaster, Steven and Rendleman, Richard, 1982, Option prices as predictors of equilibrium stock prices, *Journal of Finance* 37, 1043-1048.

Margrabe, William, 1978, The value of an option to exchange one asset for another, *Journal of Finance* 33 (March), 177-186.

Melick, William R. and Thomas, Charles P., 1997, Recovering an asset's implied probability density function from option prices: An application to oil prices during the Gulf crisis, *Journal of Financial and Quantitative Analysis* 32 (March), 91-115.

Melino, Angelo and Turnbull, Stuart, 1990, Pricing foreign currency options with stochastic volatility, *Journal of Econometrics* 45, 239-265.

Merton, Robert C., 1976, Option pricing when underlying stock returns are discontinuous, *Journal of Financial Economics* 3, 125-143.

Merton, Robert C., 1973, Theory of rational option pricing, *Bell Journal of Economics and Management Science*, 141-183.

Meulbroek, Lisa, 1992, A comparison of forward and futures prices of an interest-rate sensitive financial asset, *Journal of Finance* 47 (March), 381-396.

Miller, Merton H. and Modigliani, Franco, 1961, Dividend policy, growth and the valuation of shares, *Journal of Business* 34 (October), 411-433.

Miller, Merton H., Muthuswamy, Jay, and Whaley, Robert E., 1994, Mean reversion in S&P 500 index basis changes: Arbitrage-induced or statistical illusion?, *Journal of Finance* 49 (June), 479-513.

Modigliani, Franco and Miller, Merton H., 1958, The cost of capital, corporation finance and the theory of investment, *American Economic Review* 48 (June), 261-297.

Nathan Associates Inc., Robert R., 1969. Public Policy of a Futures-Type Market in Options on Securities. Prepared for the Chicago Board of Trade.

Ofek, Eli, Richardson, Matthew, and Whitelaw, Robert F., 2002. Limited arbitrage and short sales restrictions: Evidence from the options market. New York University.

Ogden, J. P. and Tucker, Alan L., 1987, Empirical tests of the efficiency of the currency futures option markets, *Journal of Futures Markets* 7, 695-703.

Park, H. Y. and Chen, A. Y., 1985, Differences between futures and forward prices: A further investigation of marking to market effects, *Journal of Futures Markets* 5 (February), 77-88.

Parkinson, Michael, 1977, Option pricing: The American put, *Journal of Business* 50 (January), 21-36.

Parkinson, Michael, 1980, The extreme value method for estimating the variance of the rate of return, *Journal of Business* 53 (1), 61-65.

Phillips, Susan and Smith Jr., Clifford W., 1980, Trading costs for listed options: The implications for market efficiency, *Journal of Financial Economics* 8, 179-201.

Rendleman Jr., Richard J. and Bartter, Brit J., 1979, Two-state option pricing, *Journal of Finance* 34 (December), 1093-1110.

Rendleman Jr., Richard J. and Carabini, Christopher, 1979, The efficiency of the Treasury bill futures markets, *Journal of Finance* 34 (September), 895-914.

Roll, Richard, 1977, An analytic valuation formula for unprotected American call options on stocks with known dividends, *Journal of Financial Economics* 5 (November), 251-258.

Rubinstein, Mark, 1985, Alternative paths to portfolio insurance, *Financial Analysts Journal* 41 (July-August), 42-52.

Rubinstein, Mark, 1994, Implied binomial trees, *Journal of Finance* 49, 771-818.

Rubinstein, Mark, 1985, Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active CBOE option classes from August 23, 1976 through August 31, 1978, *Journal of Finance* 40, 455-480.

Rubinstein, Mark, 1991, Pay now, choose later, *RISK* 4 (February).

Rubinstein, Mark and Reiner, Eric, 1991, Breaking down the barriers, *RISK* 4 (September), 31-35.

Samuelson, Paul A., 1965, Rational theory of warrant pricing, *Industrial Management*

Review 10 (Winter), 13-31.

Schmalensee, Robert and Trippi, Robert R., 1978, Common stock volatility expectations implied by option premia, *Journal of Finance* 33, 129-147.

Schwartz, Eduardo S., 1977, The valuation of warrants: Implementing a new approach, *Journal of Financial Economics* 4, 79-93.

Schwert, G. William, 1990, Stock volatility and the crash of '87, *Review of Financial Studies* 3, 77-102.

Schwert, G. William, 2002, Stock volatility in the new millenium: How wacky is NASDAQ?, *Journal of Monetary Economics* 49, 3-26.

Schwert, G. William, 1989, Why does stock market volatility change over time?, *Journal of Finance* 44, 1115-1154.

Scott, Louis O., 1987, Option pricing when the variance changes randomly: Theory estimation, and an application, *Journal of Financial and Quantitative Analysis* 22, 419-438.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425-442.

Shastri, Kuldeep and Tandon, Kishore, 1986a, An empirical test of a valuation model for American options on futures contracts, *Journal of Financial and Quantitative Analysis* 21 (December), 377-392.

Shastri, Kuldeep and Tandon, Kishore, 1986b, Valuation of foreign currency options: Some empirical tests, *Journal of Financial and Quantitative Analysis* 21 (June), 145-160.

Sheikh, Aamir M., 1991, Transaction data tests of S&P 100 call option pricing, *Journal of Financial and Quantitative Analysis* 26 (December), 459-475.

Shimko, David, 1993, Bounds of probability, *Risk* 6, 33-37.

Shleifer, Andrei and Vishny, Robert, 1997, The limits of arbitrage, *Journal of Finance* 52, 35-55.

Shyy, G., Vijayraghavan, V., and Scott-Quinn, B. , 1996, A further investigation of the lead-lag relationship between the cash market and stock index futures market with the use of bid/ask quotes: The case of France, *Journal of Futures Markets* 16, 405-420.

Skinner, Douglas, 1989, Options markets and stock return volatility, *Journal of Financial Economics* 23, 61-78.

Sorescu, Sorin M., 2000, The effect of options on stock prices: 1973 to 1995, *Journal of Finance* 55 (1), 487-514.

Sprenkle, Case M., 1961, Warrant prices as indicators of expectations and preferences, *Yale Economic Essays* 1, 172-231.

Stein, Elias M. and Stein, Jeremy C., 1991, Stock price distributions with stochastic volatility: An analytic approach, *Review of Financial Studies* 4, 727-752.

Stephan, Jens A. and Whaley, Robert E., 1990, Intraday price change and trading volume relations in the stock and stock option markets, *Journal of Finance* 45, 191-220.

Stoll, Hans R., 1969, The relationship between put and call prices, *Journal of Finance* 24, 802-24.

Stoll, Hans R. and Whaley, Robert E., 1986. Expiration day effects of index options and futures. Vanderbilt University, Duke University.

Stoll, Hans R. and Whaley, Robert E., 1997, Expiration-day effects of the All Ordinaries Share Price Futures: Empirical evidence and alternative settlement procedures, *Australian Journal of Management* (December).

Stoll, Hans R. and Whaley, Robert E., 1991, Expiration-day effects: What has changed?, *Financial Analysts Journal* 47, 58-72.

Stoll, Hans R. and Whaley, Robert E., 1986, New option instruments: Arbitrageable linkages and valuation, *Advances in Futures and Options Research* 1 (A), 25-62.

Stoll, Hans R. and Whaley, Robert E., 1987, Program trading and expiration-day effects, *Financial Analysts Journal* 43, 16-28.

Stoll, Hans R. and Whaley, Robert E., 1990, Program trading and individual stock returns: ingredients of the triple-witching brew, *Journal of Business* 63 (1), S165-S192.

Stoll, Hans R. and Whaley, Robert E., 1990, The dynamics of stock index and stock index futures returns, *Journal of Financial and Quantitative Analysis* 25, 441-468.

Stoll, Hans R. and Whaley, Robert E., 1985, The new options markets, in Peck, Anne, ed.: *Futures Markets: Their Economic Role*, (American Enterprise Institute, Washington, D.C.).

Stulz, Rene, 1982, Options on the minimum or maximum of two assets, *Journal of Financial Economics* 10, 161-185.

Sullivan, Edward J. and Weithers, Timothy M., 1991, Louis Bachelier: The father of modern option pricing theory, *Journal of Economic Education* (Spring), 165-171.

Trennepohl, Gary L. and Dukes, William P., 1979, CBOE options and stock volatility, *Review of Business and Economic Research* 14, 49-60.

Viswanath, P. V., 1989, Taxes and futures-forward price difference in the 91-day T-bill market, *Journal of Money, Credit and Banking* 21 (2), 190-205.

Wahab, M. and Lashgari, M., 1993, Price dynamics and error correction in stock index and stock index futures markets, *Journal of Futures Markets* 13, 711-742.

Whaley, Robert E., 2002. Derivatives: Markets, Valuation, and Risk Management. Duke University.

Whaley, Robert E., 1993, Derivatives on market volatility: Hedging tools long overdue, *Journal of Derivatives* 1, 71-84.

Whaley, Robert E., 1981, On the valuation of American call options on stocks with known dividends, *Journal of Financial Economics* 9 (June), 207-211.

Whaley, Robert E., 1982, Valuation of American call options on dividend-paying stocks: Empirical tests, *Journal of Financial Economics* 10, 29-58.

Whaley, Robert E., 1986, Valuation of American futures options: Theory and empirical tests, *Journal of Finance* 41 (March), 127-150.

Whaley, Robert E., 1996, Valuing spreads options, *Energy in the News* (Summer), 42-45.

Wiggins, James B., 1987, Option values under stochastic volatility: Theory and empirical estimates, *Journal of Financial Economics* 19, 351-372.

TABLE 1: Milestones in the history of derivative contract market development.

1750 BC	Options to default on interest payments are described in the *Code of Hammurabi.*
350 BC	Options to rent olive presses are described in Aristotle's *Politics.*
1600 AD	Forward and option contracts on tulip bulbs flourish in Holland. Tulip bulb prices collapse in the winter of 1637 causing contract defaults.
1848	Chicago Board of Trade (CBT) is formed to provide a centralized marketplace for cash and forward transactions in grains.
1865	CBT revamps forward markets by introducing futures contracts on agricultural commodities. Futures contracts are standardized contracts in terms of quality, quantity, and time and place of delivery. Futures trading involved a clearinghouse and a system of margining.
1870	New York Cotton Exchange (NYCE) is formed to trade cotton futures.
1874	Chicago Produce Exchange (CPE) is formed to trade futures on butter, eggs, poultry, and other perishable products.
1878	London Corn Trade Association introduces the first futures contract in the U.K.
1882	Coffee Exchange (CE) is formed by a group of coffee merchants to trade coffee futures.
1898	Butter and egg dealers withdraw from the CPE to form the Chicago Butter and Egg Board (CBEB).
1904	Winnipeg Commodity Exchange (WCE) introduces first commodity (oat) futures contracts in Canada.
1919	São Paulo Commodities Exchange (BMSP) introduces first commodity futures in Brazil. CBEB becomes the Chicago Mercantile Exchange (CME).
1933	Commodity Exchange (COMEX) is formed and introduces first futures contract on a non-agricultural commodity—silver.
1952	October: London Metal Exchange (LME) lists the first metal (lead) futures contract in the U.K.
1960	Sydney Futures Exchange (SFE), originally called the Greasy Wool Futures Exchange, is formed to trade greasy wool futures.
1961	September: CME introduces first futures contract on livestock—frozen pork bellies.
1972	February: CME introduces first futures contract on a financial instrument—foreign currencies.
1973	April 26th: CBT organizes the Chicago Board Options Exchange (CBOE) for the purpose of trading *call* options on sixteen New York Stock Exchange (NYSE) common stocks. Trading begins in a small smokers' lounge overlooking the futures exchange.
1975	CBT introduces first interest rate futures contract—Government National Mortgage Association (GNMA) futures. Montreal Exchange (ME) begins to list stock options in Canada. January 13th: American Stock Exchange (AMEX) begins to list call options on stocks. June 27th: Philadelphia Stock Exchange (PHLX) begins to list call options on stocks.

1976 Australian Options Market (AOA) is formed in Australia to list stock options.
January: CME begins to list T-bill futures contracts.
March: Toronto Stock Exchange (TSE) lists stock options in Canada.
April: Pacific Stock Exchange (PSE) begins to list stock options.
December: Midwest Stock Exchange (MSE) begins to list call options on stocks.

1977 June 3rd: Put options on common stocks are listed for the first time in the U.S. on the CBOE, AMEX, MSE, PHLX, and PSE.
August: CBT begins to list T-bond futures contracts.

1978 London Traded Options Market (LTOM) is formed and begins to list stock options.
European Options Exchange (EOE), formed in November 1977, begins to list stock options in The Netherlands.
November: NYMEX introduces first energy futures—heating oil.

1980 International Petroleum Exchange (IPE) is formed in the U.K. to list futures on petroleum and petroleum products.
First over-the-counter (OTC) Treasury bond option takes place.
September: Toronto Futures Exchange (TFE) is formed to list futures contracts on financial assets in Canada.

1981 First over-the-counter (OTC) interest rate swap transaction takes place.
December: CME introduces the first *cash settlement* futures contract—the Eurodollar futures

1982 London International Financial Futures Exchange (LIFFE) is formed in the U.K. to trade futures on financial instruments.
February: Kansas City Board of Trade (KCBT) introduces first futures on a stock index (i.e., the Value Line stock index)
April: CME begins to list S&P 500 index futures.
October: First options listed on instruments other than common stocks:

- CBOE and AMEX begin to list options on Treasury bonds, notes, and bills.
- CBT begins to list options on T-bond futures.
- Coffee, Sugar, and Cocoa (CSCE) begins to list options on sugar futures.
- COMEX begins to list options on gold futures.

December: PHLX begin to list options on currencies.

1983 January: CME and New York Futures Exchange (NYFE) begin to list options on stock index futures.
February: SFE begins to list futures on the All Ordinaries Share Price Index in Australia.
March: CBOE begins to list options on stock indexes.

1984 Singapore International Monetary Exchange (SIMEX) is inaugurated as the first financial futures exchange in Asia.
May: LIFFE begins to list futures on the FT-SE 100 index in the U.K.

1985 June 3rd: Options on NASDAQ stocks begin trading.
June 3rd: New York Stock Exchange (NYSE) begins trading stock options.

1986 May: Hong Kong Futures Exchange begins to list futures on the Hang Seng Index.
September: SIMEX begins to list futures on the Nikkei 225 Stock Average.

1991 Notional amount of OTC derivatives trading surpasses exchange-traded derivatives.

Figure 1: Average implied volatility by moneyness category for S&P 500 index options and twenty stock options traded on the Chicago Board Options Exchange during the period January 1995 through December 2000. Stock option classes are the twenty most active that traded continuously on the CBOE throughout the sample period. Implied volatilities are computed daily based on the midpoint of the bid/ask quotes as of 3PM (CST). The moneyness categories are based on the option deltas:

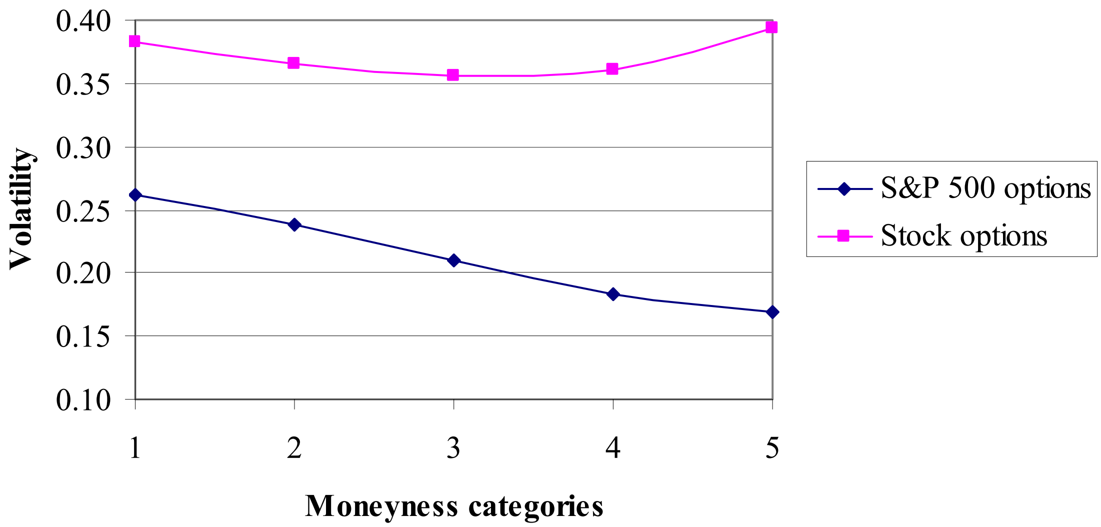
Category 1: deep out-of-the-money puts $(\Delta_p > -.125)$ /deep in-the-money calls $(\Delta_c > .875)$

2: out-of-the-money puts $(-.125 \geq \Delta_p > -.375)$ /in-the-money calls $(.625 < \Delta_c \leq .875)$

3: at-the-money puts $(-.325 \geq \Delta_p > -.625)$ /at-the-money calls $(.375 < \Delta_c \leq .625)$

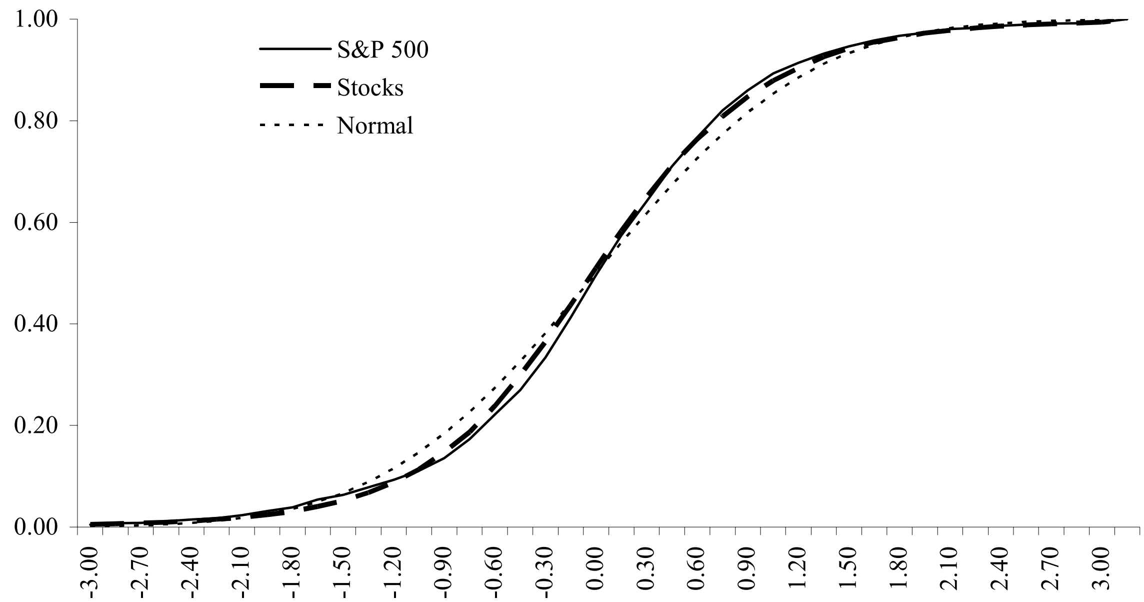
4: in-of-the-money puts $(-.625 \geq \Delta_p > -.875)$ /out-of-the-money calls $(.125 < \Delta_c \leq .375)$

5: deep in-the-money puts $(-.875 \geq \Delta_p)$ /deep out-of-the-money calls $(\Delta_c \leq .125)$



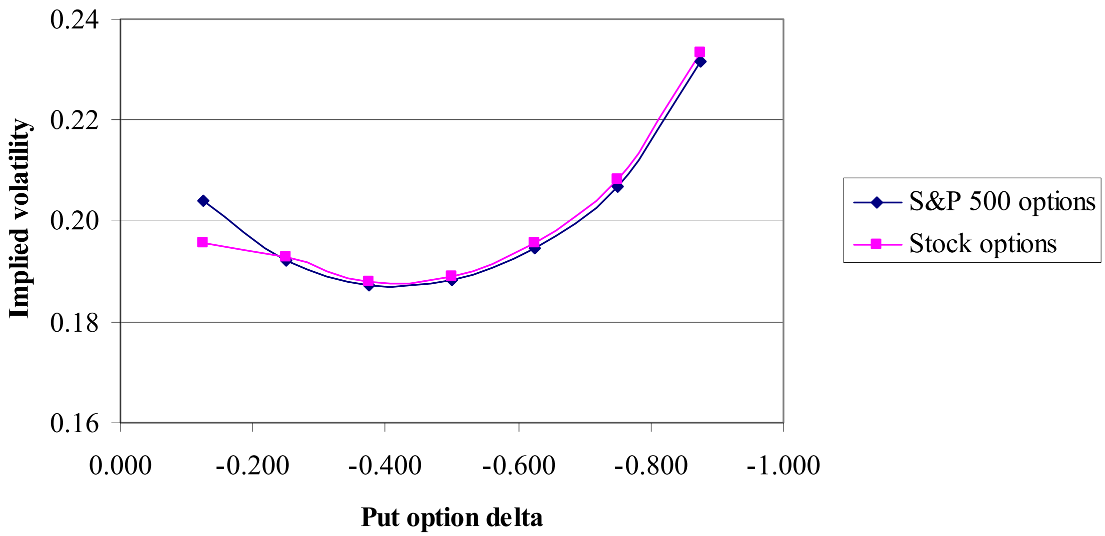
Source: Bollen and Whaley (2002).

FIGURE 2: Empirical cumulative distribution functions (CDF) of standardized daily returns for the S&P 500 index, average empirical CDF of the standardized returns of twenty stocks, and the analytical CDF of a standard normal. Return data are from January 1995 through December 2000.



Source: Bollen and Whaley (2002).

FIGURE 3: Hypothetical prices of one-month European-style put options based on an asset price of 100, a volatility rate of 20%, and a risk-free rate of interest of 5%. The underlying empirical distributions for the S&P 500 index and the individual stocks are tabulated using daily returns over the period January 1995 through December 2000.



Source: Bollen and Whaley (2002).

Figure 4: Average difference between implied volatility and historical volatility over the most recent sixty trading days by moneyness category for S&P 500 index options and twenty stock options traded on the Chicago Board Options Exchange during the period January 1995 through December 2000. Stock option classes are the twenty most active that traded continuously on the CBOE throughout the sample period. Implied volatilities are computed daily based on the midpoint of the bid/ask quotes as of 3PM (CST). The moneyness categories are based on the option deltas:

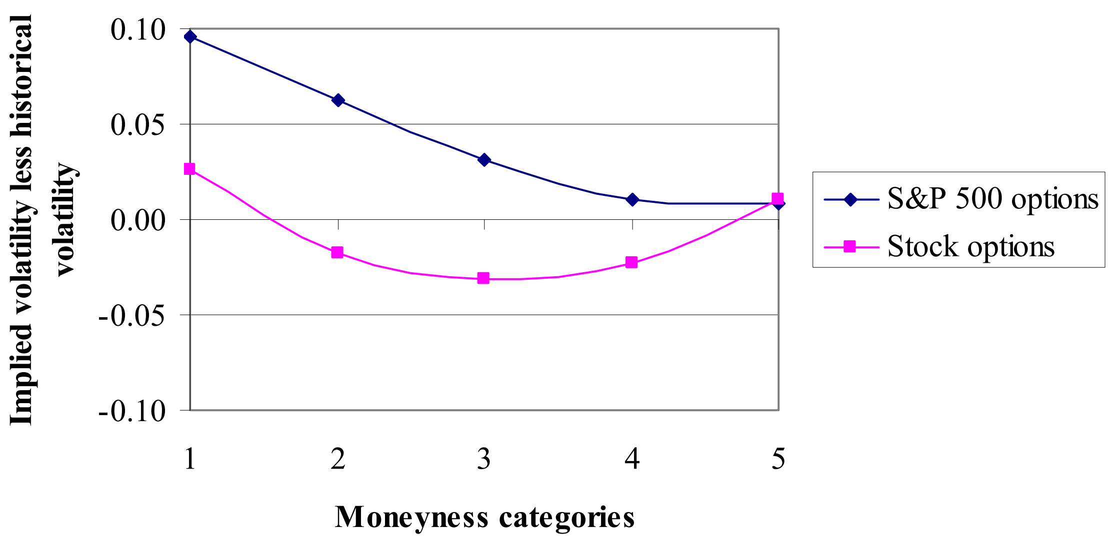
Category 1: deep out-of-the-money puts $(\Delta_p > -.125)$ /deep in-the-money calls $(\Delta_c > .875)$

2: out-of-the-money puts $(-.125 \geq \Delta_p > -.375)$ /in-the-money calls $(.625 < \Delta_c \leq .875)$

3: at-the-money puts $(-.325 \geq \Delta_p > -.625)$ /at-the-money calls $(.375 < \Delta_c \leq .625)$

4: in-of-the-money puts $(-.625 \geq \Delta_p > -.875)$ /out-of-the-money calls $(.125 < \Delta_c \leq .375)$

5: deep in-the-money puts $(-.875 \geq \Delta_p)$ /deep out-of-the-money calls $(\Delta_c \leq .125)$



Source: Bollen and Whaley (2002).