Economic Perspectives—Volume 1, Number 2—Fall 1987—Pages 55–72

The Arbitrage Principle in Financial Economics

Hal R. Varian

An economics professor and a Yankee farmer were waiting for a bus in New Hampshire. To pass the time, the farmer suggested that they play a game. "What kind of game would you like to play?" responded the professor. "Well," said the farmer, "how about this: I'll ask a question, and if you can't answer my question, you give me a dollar. Then you ask me a question and if I can't answer your question, I'll give you a dollar."

"That sounds attractive," said the professor, "but I do have to warn you of something: I'm not just an ordinary person. I'm a professor of economics."

"Oh," replied the farmer, "In that case we should change the rules. Tell you what: if you can't answer my question you still give me a dollar, but if I can't answer yours, I only have to give you fifty cents."

"Yes," said the professor, "that sounds like a fair arrangement."

"Okay," said the farmer, "Here's my question: what goes up the hill on seven legs and down the hill on three legs?"

The professor pondered this riddle for a while and finally replied. "Gosh, I don't know... what does go up the hill on seven legs and down the hill on three legs?"

"Well," said the farmer, "I don't know either. But if you give me your dollar, I'll give you my fifty cents!"

The above story is an illustration of *arbitrage*: arranging a transaction involving no cash outlay that results in a sure profit. As this story shows, opportunities for arbitrage do occasionally arise. But in a well-developed market with rational, profit-seeking individuals such opportunities should be very rare indeed, since profit-maxi-

■ *Hal R. Varian is Reuben Kempf Professor of Economics and Professor of Finance, the University of Michigan, Ann Arbor, Michigan.*

mizing agents will attempt to exploit arbitrage opportunities as soon as they arise. It is generally felt that part of the definition of equilibrium in a perfect market is that no opportunities for pure arbitrage exist.

The importance of arbitrage conditions in financial economics has been recognized since Modigliani and Miller's classic work on the financial structure of the firm. They showed that if a firm could change its market value by purely financial operations such as adjusting its debt-equity ratio, then individual shareholders and bondholders could engage in analogous portfolio transactions that would yield pure arbitrage profits. If the market was efficient enough to eliminate arbitrage profits for the individual shareholders, then it would eliminate arbitrage profits for the firm as well.

Modigliani and Miller's proof of this proposition used an ingenious arbitrage argument. Subsequently, financial economists have used arbitrage arguments to examine a variety of other issues involving asset pricing. One of the major advances in financial economics in the past two decades has been to clarify and formalize the exact meaning of "no arbitrage" and to apply this idea systematically to uncover hidden relationships in asset prices. Many important results of financial economics are based squarely on the hypothesis of no arbitrage, and it serves as one of the most basic unifying principles of the study of financial markets. In this essay we will examine some of these results. To avoid cluttering up the exposition with citations, a discussion of the original sources and a guide to further reading will be found in the final section of this article.

General Principles of Asset Pricing

Consider a market for assets that pay off in different states of nature. These states need not be full-fledged Arrow-Debreu states of the world which describe all possible relevant circumstances; they are simply the outcomes of some random process. We assume that individuals care about their wealth in different states of nature, and prefer more wealth in any state of nature rather than less.

Let's denote the payoff of asset a in state s by R_{sa} and suppose that the number of assets is A and that the number of states is S. An asset is described by a vector giving its payoffs in each of the S states of nature. Thus the first security is described by the column vector $(R_{11} \dots R_{S1})$ and the ith security is described by the vector $(R_{1i} \dots R_{Si})$. The payoff matrix of the entire set of assets is then represented by the matrix

$$R = \begin{pmatrix} R_{11} & \cdots & R_{1A} \\ \vdots & \ddots & \vdots \\ R_{S1} & \cdots & R_{SA} \end{pmatrix}.$$

This S by A matrix gives the payoffs of each of the A assets in each of the S states:

each column of the payoff matrix represents a different security and each row gives the payoffs in a particular state of nature of each of the securities. The entire matrix summarizes the payoff characteristics offered by this particular collection of assets.

Let x_a indicate the amount held of asset a. A portfolio of assets is then a column vector $x = (x_1, \ldots, x_A)$. The components of the portfolio x can be of either sign: a positive value of x_a indicates that one has a "long" position in security a, and thus is entitled to receive the appropriate payoff if state s materializes, and a negative value of x_a indicates a "short" position in the security so that one must pay out the appropriate amount if state s occurs.

The wealth in state s that one receives from holding a portfolio $x = (x_1, \ldots, x_A)$ is given by the expression $w_s = \sum_{a=1}^{A} x_a R_{sa}$. Writing this out in matrix notation we have

$$\begin{pmatrix} w_1 \\ \vdots \\ w_S \end{pmatrix} = \begin{pmatrix} R_{11} & \cdots & R_{1A} \\ \vdots & \ddots & \vdots \\ R_{S1} & \cdots & R_{SA} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_A \end{pmatrix} \tag{1}$$

or, more succinctly,

$$w = Rx. \tag{2}$$

This equation illustrates the relationship between the ends and the means: the ends are the levels of wealth that the consumer can achieve in the different states of nature, which is what the consumer ultimately cares about. The means are the existing assets. By combining the existing assets into portfolios, the consumer can achieve different patterns of wealth across the states of nature. The patterns of wealth that can be achieved depend on the entire set of available assets. Thus, the market value of a single asset will typically depend on what other assets are available to combine with it.

The central case is where the number of assets matches the number of states of nature. In this case, any pattern of wealth can be achieved by some portfolio of existing assets. To achieve a particular distribution of wealth, $w = (w_1, \ldots, w_S)$, one simply solves the system of equations (1) for the portfolio $x = (x_1, \ldots, x_A)$ that achieves that distribution of wealth. Since the system of equations has exactly as many unknowns as equations, it will always be possible to solve for such a portfolio.[1]

If the assets outnumber the states of nature, there will be more unknowns than equations and several portfolios will exist that generate any particular distribution of wealth. On the other hand, if states of nature outnumber assets, then it will not be possible to solve the system of equations for all distributions of wealth—some patterns of wealth cannot be constructed using the existing set of assets.

This latter case is presumably the most realistic, since one would generally think that people care about more outcomes than they have assets to trade. But the question is controversial. If people really care about achieving a certain distribution of wealth

[1]Here we assume that no redundant assets exist; in other words, the matrix R has full rank.

across states of nature, doesn't it seem likely that the market will offer an asset that will achieve such a pattern? One surprising finding of financial economics is a way to construct portfolios that achieve arbitrary payoff patterns, even in situations where there appear to be more states of nature than assets. A later section will explore this result in more detail.

One important kind of assets are Arrow-Debreu securities. These are assets that pay off \$1 if and only if a particular state of nature occurs; otherwise they pay off zero. Thus the payoff pattern of an Arrow-Debreu security takes the form $(0, \ldots, 1, \ldots, 0)$, where the 1 occurs in location s.

Arrow-Debreu assets can be thought of as especially basic assets, in both the economic sense and the linear algebra sense. They are basic in the economic sense, since we will see shortly that any pattern of payoffs can be constructed from portfolios of Arrow-Debreu assets. They are basic in the linear algebra sense in that they form a basis for the linear space of all payoffs. It is this fundamental nature of the Arrow-Debreu assets that makes them so important in analyzing the pricing of financial assets.

Asset Prices

When investors choose portfolios they are in effect choosing a distribution of wealth across the states of nature. In making this choice they are constrained by their budget constraint: the constraint imposed by amount of wealth that they have to invest and the prices of the various assets that they face.

To express the idea of a budget constraint, let the price of asset a be denoted by p_a, and let $p = (p_1, \ldots, p_A)$ be the row vector of asset prices. The value of a portfolio $x = (x_1, \ldots, x_A)$ will then be given by $px = \sum_{a=1}^{A} p_a x_a$. This formulation is just like standard consumer theory: the value of a bundle of goods is expressed as the sum of the expenditures on the various goods. The difference is that the goods that are being chosen—the assets—are not the ultimate end of consumption. They are only the means to an end. What consumers really care about is the final distribution of wealth that different portfolios provide. Therefore, any two portfolios that provide the same pattern of wealth must be worth the same amount. Consider, for example, the case of the Arrow-Debreu securities. If we know the price of each Arrow-Debreu security, we can value any asset. Since the payoff pattern of any asset can be achieved by some portfolio of Arrow-Debreu securities, the price of the asset must be equal to the price of the portfolio of Arrow-Debreu securities that realizes that same distribution of wealth across the states of nature.

Formally, let π_s be the price of the Arrow-Debreu security that pays off \$1 if state s occurs, and let (R_{sa}) be the payoff pattern of some asset a. Then the equilibrium price of asset a must be given by

$$p_a = \sum_{s=1}^{S} \pi_s R_{sa}.$$

Why? There are two arguments. The intuition is that π_s measures the value of a dollar in state s, and that the security pays off R_{sa} dollars in state s. Summing over all the possible states gives the value of security a.

This argument is plausible, but rests on a notion of "value" that is somewhat slippery. A more compelling argument is based on arbitrage considerations. If a complete set of markets for all Arrow-Debreu securities exists, and the price of an asset ever deviated from the price of a portfolio of Arrow-Debreu securities that generates the same pattern of payoffs, then there would be a sure way of making money—just sell the security and buy the Arrow-Debreu portfolio, or vice-versa, depending on which was worth more. If no arbitrage possibilities exist and a complete set of Arrow-Debreu securities are marketed, then any asset can be valued in terms of the prices of the Arrow-Debreu assets.

This discussion shows that any asset can be valued in terms of a particular set of assets in the case where the number of assets matches the number of states of nature. However, it turns out that a similar result holds even if there are fewer assets than states of nature. But this demonstration requires a more substantive assumption about arbitrage possibilities, a point to which we now turn.

A Formalization of the No Arbitrage Condition

The purpose of this paper is to examine the concept of arbitrage, or more precisely, the absence of arbitrage. What should this mean? Basically, the no arbitrage condition must rule out "free lunches"—configurations of prices such that an individual can get something for nothing. Any portfolio that pays off nonnegative amounts in every state of nature must be valuable to individuals, so if no free lunches exist, portfolios which are guaranteed to have nonnegative payoffs must have a nonnegative cost. Recall that a portfolio's pattern of returns can be represented by Rx and the cost of the portfolio by px. Using this notation we can state the

No Arbitrage Condition

If $Rx \geq 0$ *then we must have* $px \geq 0$.

The no arbitrage condition loosely described above is now an explicit algebraic condition. This condition imposes some restrictions on the equilibrium prices p. Given a particular set of assets, as described by the payoff matrix R, only certain asset prices p are consistent with the absence of arbitrage. What characterizes such prices? What restrictions does the assumption of no arbitrage impose on the asset prices?

The Appendix shows that the No Arbitrage Condition implies that there must exist a vector of nonnegative "state prices" $\pi = (\pi_1, \ldots, \pi_S)$ such that the price of any existing asset a is given by

$$p_a = \sum_{s=1}^{S} \pi_s R_{sa}. \tag{3}$$

What is the economic interpretation of the state prices π_s? Consider an Arrow-Debreu asset that pays off \$1 if state s occurs and zero otherwise. Then according to equation (3), the value of this asset must be given by π_s. In other words, π_s measures the value of a dollar in state s. Seen in this light, equation (3) is very natural: it says that to determine the value of any asset, examine how much it pays off in a given state, multiply that by the value of a dollar delivered in that state, and then sum over all the states.

The existence of the state prices π is a necessary and sufficient condition for the absence of arbitrage. If no arbitrage possibilities exist, then there must be state prices, and if the state prices exist, there can be no arbitrage possibilities. Hence the existence of the state prices is equivalent to the absence of arbitrage: any argument that follows from the absence of arbitrage must follow from the existence of the state prices and vice versa.

The state prices (π_s) emerge as a consequence of the No Arbitrage Condition—whenever the market works efficiently enough to eliminate the possibilities of arbitrage, there must be state prices that can be used to value the existing assets. Of course the state price vector π will not, in general, be unique. With S states and A assets, the equations in (3) provide A equations in S unknowns. Thus there will in general be an $S - A$ dimension set of solutions to this system. Only if the number of (independent) assets equals the number of states of nature will a unique set of state prices exist.

Value Additivity

The No Arbitrage Condition is very simple, and yet surprisingly powerful. Let's use this condition to prove an important result in the theory of financial markets, the Value Additivity Theorem.

Consider two securities, a and b, which have payoffs (R_{sa}) and (R_{sb}). Suppose that the prices of these securities, p_a and p_b, are known. Now consider the security which has a random payoff given by some linear combination of the payoffs of the two securities; i.e., one that has the payoffs $R_{sc} = AR_{sa} + BR_{sb}$, where A and B are arbitrary constants. The No Arbitrage Condition implies that if no arbitrage opportunities exist in the market, then the price of an existing asset with a payoff vector R_{sc} must be given by $p_c = \Sigma_{s=1}^{S}\pi_s R_{sc}$. Using the fact that $R_{sc} = AR_{sa} + BR_{sb}$, then

$$\begin{aligned} p_c &= \sum_{s=1}^{S} \pi_s R_{sc} \\ &= \sum_{s=1}^{S} \pi_s (AR_{sa} + BR_{sb}) \\ &= \sum_{s=1}^{S} A\pi_s R_{sa} + \sum_{s=1}^{S} B\pi_s R_{sb} \\ &= Ap_a + Bp_b. \end{aligned}$$

These calculations prove the

Value Additivity Theorem

> *Assume no arbitrage possibilities exist. Then the price of a security whose payoffs are a linear combination of other assets must be given by the same linear combination of the prices of the other assets.*

At first glance this result appears obvious—it says that the value of the whole is equal to the sum of the values of its parts. But it has important and surprising implications. To take one simple example, note that the payoffs of the assets a and b are arbitrary; the payoffs could be highly correlated, either positively or negatively. It is well known that if two assets with negative correlation are combined, the riskiness of the resulting portfolio will be less than the riskiness of either asset held alone. Diversification has a natural value in a market with risky assets. Since a linear combination of two negatively correlated assets is less risky than either of the two assets held alone, it might seem that the combination portfolio would be worth more than the sum of the values of the two assets. Nevertheless, the Value Additivity Theorem states that the price of the portfolio consisting of both assets will just be the sum of the prices of the two individual portfolios.

Why is that? The answer is that the equilibrium prices of assets must *already reflect* the value of any kind of linear portfolio manipulation. If the value of the two assets in combination exceeded the sum of the values of the assets alone, for example, then arbitragers could just buy the two assets and sell the "mutual fund" consisting of the combination. Similarly, if the values of the individual assets were less than the combination, then arbitragers could "unbundle" the combination and make a pure profit by doing so. Since this repackaging offers a sure profit, it cannot exist in equilibrium.

The Value Additivity Theorem is also the basic principle underlying the Modigliani-Miller theorem. The Modigliani-Miller theorem states that the value of a firm is independent of its financial structure; that is, independent of the fraction of the firm financed by stock and the fraction financed by bonds. Again, the argument is based on the absence of arbitrage. If firms could change their value by changing the proportion of stocks and bonds they issue, then individual arbitrages could also repackage the existing stocks and bonds and make a sure profit. Hence, the value of the firm should depend only on the sum of the values of its stocks and bonds, not on whether the firm is weighted more heavily to debt or equity. However, the Modigliani-Miller example also illustrates the restrictions of Value Additivity Theorem.

First, the Value Additivity Theorem and the No Arbitrage Condition only apply to combinations of existing assets. If a firm issues new bonds with a different pattern of returns from the old bonds—perhaps because they have higher default risk—then the Value Additivity Theorem does not necessarily hold. If a financial operation creates assets that have a new payoff pattern not currently available through linear combinations of existing assets, then the Value Additivity Theorem, or the No Arbitrage

Condition, need not apply. If an economic shift changes the matrix of payoffs (R_{sa}) so that new possibilities of consumption across states of nature are created, the vector of state prices π will typically change. In general, determining how these state prices change will require specifying and solving an entire general equilibrium model for the asset prices, which would go far beyond the scope of this paper. Luckily, it turns out that several operations that appear to create new assets really don't, so that the Value Additivity Theorem and related results can be used to value such assets.

Second, the No Arbitrage Condition only applies directly to linear operations. An asset whose payoffs are the sum of the payoffs of two different assets must have a price which is the sum of the prices of the two different assets. Similarly, an asset whose payoff is some multiple of some other asset must have a price which is a multiple of the other asset's price. If one asset is a nonlinear function of another asset's price, the No Arbitrage Condition cannot be directly applied, at least in the static framework examined here. However, it will become apparent later that under certain circumstances, operations that appear to be nonlinear can be broken down into linear operations, so that the Value Additivity Theorem can be applied.

Using Arbitrage to Bound Option Prices

The No Arbitrage Condition has several elegant applications in the theory of option pricing. First we need some definitions. A (*call*) *option* on an asset a is a security that gives one the right to purchase the asset a at some fixed price K (*the exercise price*) within some fixed time period T. An *American option* gives one the right to exercise the option to buy the asset at any time within this period while a *European option* gives one the right to purchase the underlying asset only at the expiration date.

Clearly, an American option will be worth at least as much as a European option, other things being equal. However, if the no arbitrage condition holds, then in fact an American call option and a European call option must be worth exactly the same amount!

To establish this result, let K be the exercise price of the option, let t be the time left until the option expires and let S_t be the current market price of the asset on which the option is based. Time is being measured "backwards" in this formulation—"now" is time t and the date when the option expires is time 0.

Now suppose a riskless bond can be purchased which pays one dollar at the expiration of the option, and denote the price of such a bond t periods before the expiration of the option by B_t. Since the bond is riskless, it pays off \$1 regardless of the state of nature. Using the state prices π_s to value this payoff pattern, we have $B_t = \Sigma_{s=1}^{S} \pi_s$. This simply says that the sum of the state prices gives us the value of a pure discount bond. If the interest rate is positive, the value of one dollar to be delivered in the future must be less than one; that is, $B_t < 1$.

The value of the stock at time 0, the time when the option expires, is denoted by S_0. Note that at time t the value of the stock at expiration is a random variable. Similarly, the value of the option at expiration is random. But nonetheless, there is a simple relationship between the value of the option at time 0 and the value of the stock at time 0. Two possibilities must be considered. Either the exercise price exceeds the stock price ($K > S_t$), in which case the option will not be exercised and so is worthless, or the exercise price is less than the stock price, in which case the value of the option is just the difference between the market price and the exercise price. (If the stock is selling at $100, and an option offers the right to purchase it for $75, the value of the option is just $25.) More formally, the value of the option at expiration is $\max[0, S - K]$.

These definitions allow a proof of the result stated earlier: in the absence of arbitrage, an American and a European call option must have the same price. The proof requires that two simpler results be established.

Lemma 1. *Let C_t be the value of a European call option with exercise price K, and current stock price S_t at t periods before expiration. Then,* $C_t > \max[0, S_t - KB_t]$.

Proof. Consider the following transactions: sell K bonds that each obligate you to pay one dollar at the time the option terminates and buy one share of the stock at price S_t. The cost of this portfolio today is $S_t - KB_t$.

When the option expires at time 0 the stock will be worth S_0 and the K bonds will be worth $\$K$. Therefore the value of the portfolio at time 0 will be $S_0 - K$. The value of the option when it expires at time 0 will be $\max[S_0 - K, 0]$. This means that at expiration, the value of the option will be at least as large as the value of the portfolio whatever happens. Hence the value of the option today must be worth at least as much as this portfolio, which proves the result.

Lemma 2. *An American call option will never be exercised prior to maturity.*

Proof. Since an American call option must be worth at least as much as a European call, the inequality given in Lemma 1 holds for it also. Exercising an American option at intermediate time t gives you $S_t - K$, while the market value of the option is always at least as great as $S_t - KB_t$. Hence it is always better to sell the option at its market value than to exercise it.

Theorem

An American call option has the same value as a European call option.

This result follows from Lemmas 1 and 2. Clearly, giving an investor an additional opportunity that he will never rationally take advantage of cannot affect the value of the asset. Since an American option will always be exercised at the termination date, just like a European option, the values must be the same.

The crucial step in the argument is the "no arbitrage" result of Lemma 1. But the no arbitrage part of Lemma 1 is certainly well disguised! How was that particular portfolio chosen to construct the proof? Fortunately, more systematic ways are available to uncover the implications of option pricing that are implied by the absence of arbitrage. These methods make use of the state prices that we developed earlier.

Earlier we demonstrated that the No Arbitrage Condition implies the existence of state prices π such that the price of an asset with random payoffs R_{sa} is given by $p_a = \Sigma_{s=1}^{S} \pi_s R_{sa}$. The existence of those state prices (combined with Lemmas 1 and 2) can now help to provide an alternative proof that the current price of the option C_t must be at least as great as $S_t - KB_t$.

Let S_t be the value of the stock t periods before the option expires, and let S_{s0} be the value of the stock in state s at the time of the expiration of the option. Let K be the exercise price of the option, and let C_t be the value of the option at time t. We have already seen that the value of an option at termination in each state s is given by $\max[S_{0s} - K, 0]$.

According to the discussion about state prices, if no arbitrage opportunities exist, the value of the option t periods before expiration must be given by

$$C_t = \sum_{s=1}^{S} \pi_s \max[S_{s0} - K, 0]. \tag{4}$$

Using this expression, we calculate the following:

$$\begin{aligned} C_t &= \sum_{s=1}^{S} \pi_s \max[S_{s0} - K, 0] \\ &\geq \sum_{s=1}^{S} \pi_s (S_{s0} - K) \quad \text{by definition of max,} \\ &= \sum_{s=1}^{S} \pi_s S_{s0} - K \sum_{s=1}^{S} \pi_s \quad \text{by linearity,} \\ &= \sum_{s=1}^{S} \pi_s S_{s0} - KB_t \quad \text{by the definition of } B_t, \\ &= S_t - KB_t \quad \text{since } \sum_{s=1}^{S} \pi_s S_{s0} \quad \text{is the current price of the stock, } S_t \\ &\geq S_t - K \quad \text{since } B_t < 1. \end{aligned}$$

These calculations tell us that the current price of the option, C_t, must be at least as great as $S_t - K$. The remainder of the proof proceeds as before: since an option is always worth more alive than dead, it will never be exercised prior to maturity.

Option Values and State Prices

Not only can state prices be used to value options, but option prices can be used to value state prices. It turns out that a complete set of options at all exercise prices is

Table 1
Constructing pure securities from options

S	$C1$	$C2$	$C3$	$C12$ $C1-C2$	$C23$ $C2-C3$	$C12-C23$
1	0	0	0	0	0	0
2	1	0	0	1	0	1
3	2	1	0	1	1	0
4	3	2	1	1	1	0

equivalent to a complete set of Arrow-Debreu markets. This section will show how to derive state prices from a set of option prices.

Table 1 depicts the payoff from three different options with difference striking prices and some portfolios constructed from these options. Option $C1$ has an exercise price of \$1, $C2$ has an exercise price of \$2, and $C3$ has an exercise price of \$3. The entries in the table depict the value of the option depending on the different outcomes of the stock price, $S = 1,\dots,4$. The terminal payoffs to the option at time 0 are of the form $\max[S_{s0} - K, 0]$, as indicated above.

The objective now is to construct a portfolio that pays off in only one state. By looking at the cost of that portfolio, it will be easy to calculate the state price.

The column labeled $C12 = C1 - C2$ shows the payoffs to a portfolio consisting of a long position in the option with exercise price 1, and short position in an option with exercise price 2. The column labeled $C23 = C2 - C3$ has a similar definition, while the column labeled $C12 - C23$ involves holding the $C12$ portfolio long and the $C23$ portfolio short. Note that the payoff to this latter portfolio is just the payoff to the pure security that pays off \$1 if and only if $S = 2$. This shows that portfolio of options exists that can generate the same pattern of wealth as a pure Arrow-Debreu security.

We can further explore the relationship between option values and state prices through the use of some elementary calculus. Up until now the assets under discussion have had a finite number of payoffs. Let's relax this assumption and allow the stock to have a continuum of payoffs at time 0, which we index by s. In effect, we are simply indexing the states by the values that the stock takes on. The state prices will now be a function which we denote by $\pi(\cdot)$. The state price $\pi(s)$ measures the value now of a dollar to be delivered at time 0 in the event that the value of the stock happens to be s at that time.

The value of the stock at time t is then given by $\int_0^\infty \pi(s) s\, ds$. (The lower limit of the integral is zero due to limited liability—a stock can never be worth less than zero.) If the stock takes on value s at time 0, the option will be worth $\max[s - K, 0]$, so that the value of the option at time t will be given by

$$C_t = \int_0^\infty \max[s - K, 0]\pi(s)\, ds.$$

This is the natural analog of the expression for the option value given above. This formula can be rewritten as:

$$C_t = \int_K^\infty (s - K)\pi(s)\,ds. \tag{5}$$

Now differentiate (5) with respect to K to get

$$\frac{dC_t}{dK} = -\int_K^\infty \pi(s)\,ds.$$

Note the use of the Fundamental Theorem of Calculus and the Chain Rule in this calculation. Differentiate the expression once more to find

$$\frac{d^2C_t}{dK^2} = \pi(K).$$

This equation says that the second derivative of the option price with respect to the exercise price is the state price, a remarkable result that follows entirely from the No Arbitrage Condition. Knowing the values of options at many different exercise prices makes it possible to calculate the state prices.

By the way, the discrete calculation given earlier and the derivative calculation given here are perfectly consistent. Consider the discrete calculation once again. Instead of thinking of the stock price as changing by 1, consider it changing by ΔS. In this case, the analog of the price of portfolio $C12 - C23$ becomes

$$\frac{\dfrac{C(K-\Delta S) - C(K)}{\Delta S} - \dfrac{C(K) - C(K+\Delta S)}{\Delta S}}{\Delta S}.$$

As ΔS goes to zero, this expression approaches the second derivative of $C(K)$. Hence, the second derivative of the option price equals the state price, as established earlier.

Pure Security Prices for Dynamic Stochastic Processes

A security whose value is a known function of another securities value at some point in time is known as a *contingent security*, or a *derivative asset*. For example, a call option on a stock S has value of $\max[0, S_0 - K]$ at the expiration of the option. But how much would such an option be worth at some time t before maturity?

An earlier section derived bounds on the value of such an option that were independent of the stochastic process followed by the stock. But if the stochastic process followed by the stock is known, much tighter bounds can be derived. In fact, in most interesting cases, it is possible actually to derive the explicit value of the option —or any other derivative security whose value depends on the value of the stock—using only the arbitrage principle described above.

This method was historically used to value options, but I will illustrate by calculating the case involving a pure Arrow-Debreu security, since any other derived security based on a given asset can be valued by using the Arrow-Debreu state prices.

Assume that we have a common stock, for which the price will either increase to Su with probability q, or fall to Sd with probability $(1 - q)$, as illustrated by

$$S \begin{cases} Su & q \\ Sd & 1 - q. \end{cases}$$

The problem is to value a pure Arrow-Debreu security that has payoff pattern given by

$$\pi_u \begin{cases} 1 & q \\ 0 & 1 - q. \end{cases}$$

How can this be done? Consider a portfolio that has x shares of the stock and B bonds that pay off Br in period two. (Here r is simply one plus the rate of interest.) Then this portfolio has a return pattern of

$$xS + B \begin{cases} xSu + Br & q \\ xSd + Br & 1 - q. \end{cases}$$

Let us choose x and B so as to create the same return pattern as the pure security. Thus we want

$$xSu + Br = 1$$

$$xSd + Br = 0.$$

Some algebra shows that the solution to these equations, (x^*, B^*), is given by

$$x^* = \frac{1}{(Su - Sd)}$$

$$B^* = -\frac{d}{(u - d)r}.$$

Since this portfolio (x^*, B^*) has the same returns as the pure security, it must have the same price. Thus, the value of an Arrow-Debreu security that pays off \$1 in the up state and \$0 in the down state must be given by

$$\pi_u = x^*S + B^* = \frac{1}{(u - d)} - \frac{d}{(u - d)r} = \frac{(r - d)}{(u - d)r}.$$

A similar argument shows that the pure security that pays off \$1 only in the down state must be worth

$$\pi_d = \frac{u - r}{(u - d)r}.$$

So far this is simply an example of the general argument given in Section 2: if there are only two states of nature, (u, d), and there are two assets with payoffs (Su, Sd) and (rB, rB), we can construct the Arrow-Debreu securities with payoffs $(1, 0)$ and $(0, 1)$ and value those pure securities using the prices of the other two assets.

Note that $\pi_u + \pi_d = 1/r$. This says that a portfolio that pays off \$1 in each state must be worth $1/r$ dollars today. This makes perfectly good sense: a portfolio that pays off \$1 in each state gives a certain return of \$1 no matter what, and the value today of a dollar for sure one period from now is simply the present value $1/r$. (That is, the value of a pure discount bond with one period left to maturity is simply the reciprocal of one plus the rate of interest.)

We have determined the Arrow-Debreu state prices by constructing a portfolio that has the same payoffs. Another way to determine the state prices is to use the arbitrage relations directly. The No Arbitrage Condition implies that the value of the stock S is given by

$$S = \pi_u Su + \pi_d Sd,$$

which reduces to

$$\pi_u u + \pi_d d = 1. \tag{6}$$

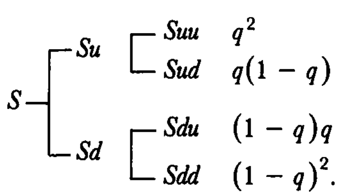
Similarly, the value of the bond today is $1/r$, so

$$\pi_u + \pi_d = 1/r. \tag{7}$$

Solving the two equations (6) and (7) for the two unknowns, π_u and π_d will give us the state prices.

Note that the state prices π_u and π_d do not depend on the probability that the stock goes up or down. They only depend on *how much* the stock can go up or down—that is, they only depend on the parameters of the stochastic process governing the behavior of the stock. Once the stochastic process is known, the state prices can be calculated.

Let us now generalize this method to a problem involving three time periods: period 0 (today), period 1 (when the stock goes up or down once) and period 2 (where the stock can go up or down once again). Schematically we have



Note that there are only three events in the last period, despite appearances. The states (ud) and (du) have exactly the same payoff and probability of occurring.

We would like to value a pure security that pays off \$1 in state (uu). Consider the situation from the vantage point of period 2. We know that in the second period, a security that pays off \$1 in state (uu) will be worth $(r - d)/(u - d)r$ by the argument given above. We also know that in the second period a security that pays off 0 in state ud and 0 in state (dd) will be worth 0 today. (If not, there is an obvious possibility for arbitrage.)

Now consider the situation in the first period, where we now need to value a security that has payoff $(r - d)/(u - d)r$ in the up state and 0 in the down state. Using the pure security prices π_u and π_d derived above, we have

$$\pi_{uu} = \pi_u\left(\frac{r - d}{(u - d)r}\right) + \pi_d 0 = \pi_u^2.$$

Continuing in a like manner we find that

$$\pi_{ud} = 2\pi_u\pi_d = 2\pi_u(1 - \pi_u)$$

$$\pi_{dd} = \pi_d^2.$$

It is not hard to see the pattern emerging. If we want to value a pure security that pays off \$1 n periods from now if and only if there are u up jumps and $(n - u)$ down jumps, then we have

$$\pi_{u,\,n-u} = \binom{n}{u}\pi_u^u\pi_d^{n-u}.$$

The resemblance to the binomial distribution is striking; in fact this is the binomial distribution with π_u and π_d playing the roles of the binomial probabilities. But note that in this problem the actual probability q plays *no role at all* in the formula for valuing the pure security. The value of a security that pays off \$1 if a given state occurs will be independent of the probability of that state occurring.

Of course, the price of the stock itself will depend on the probability of that state occurring, or more generally, it will depend on peoples' beliefs about the probability of that state occurring. But any collection of people who agree on the current value of the stock and that the stock price follows a binomial distribution must agree on the values of all pure securities whose payoffs are conditional on realizations of that stock's price in future states of nature, as long as no arbitrage opportunities are available.

Now that we have the state prices for each state $(u, n - u)$ we can calculate the value of any contingent security now, once we know its value in each of the states. For example we could calculate the value of a call option on the given stock since its payoff at expiration is a known function of the value of the stock at expiration, namely $\max[0, S_0 - K]$.

Applying the above formula, we find that the value of an option on a stock that follows a binomial process is given by

$$C_n = \sum_{u=1}^{n} \frac{n!}{u!(n-u)!} \pi_u^u \pi_d^{n-u} \max[0, u^u d^{n-u} S - K].$$

Thus in the case of a binomial stochastic process we can calculate the state prices for any possible realization of the stochastic process at any date in the future using only the knowledge of the current value of the stock and the bond. Even though there are only two assets available we can "span" the arbitrarily large number of states generated by the binomial process. The trick is simply that the way the states evolve from each other is known a priori; since we can span the evolution of the process from every period to the following period, we can in effect value any asset whose value evolves according to this process.

In order to construct state prices for assets on other stochastic processes we can use a limiting argument and let the binomial process approach a Normal or a Poisson distribution. This will give us the state prices implicit in Itô process or a jump process in the limit. In fact the ideas described above were originally developed in the search for a formula to value options on a stock that followed a continuous time Itô process. Only recently has it been recognized that an elementary treatment of these topics was possible.

Summary

We have seen how the simple principle of arbitrage described in the introduction can be used to calculate the necessary equilibrium relations between the values of assets in financial markets. The power of this simple hypothesis always seems somewhat surprising. In a way the applications of the arbitrage principle seem to contradict its own statement: the arbitrage principle says that you can't get something for nothing. But the results given above, and described further in the references given below, show that you can get quite a bit in the way of theorems, for very little in the way of assumptions!

Guide to Further Reading

The field of financial economics contains a vast literature on these topics. Here I will only describe the main sources used in preparing the above exposition and briefly mention some of the major references and surveys that can be used as an introduction to the literature.

Formalization of the No Arbitrage Condition. The formulation of the arbitrage principle and the treatment given here are due to Ross (1976, 1978). These elegant treatments provide an excellent introduction to this topic.

Using Arbitrage to Bound Option Prices. The results developed here are due to Merton (1973) and Cox and Ross (1976).

Option Values and State Prices. The result that the second derivative of the option price gives the state price is due to Breeden and Litzenberger (1978), although the development using the No Arbitrage Condition seems to be new.

Pure Security Prices for Dynamic Stochastic Processes. The treatment here follows that of Cox, Ross, and Rubinstein (1979). They were interested in the particular problem of valuing options rather than constructing the state prices implicit in the binomial model, but the procedures used are nearly the same. Breeden and Litzenberger (1978) show how to use option prices to derive state prices in a general setting. The original development of option pricing (in a continuous time model) is due to Black and Scholes (1975); further treatments are available in Cox and Ross (1976), Cox and Rubinstein (1985), and Merton (1973).

Appendix
Proof of the No Arbitrage Theorem

The problem is to show that the No Arbitrage Condition given in the text implies the existence of the nonnegative state prices π. To attack this question, consider the following linear programming problem:

$$\min \quad px$$

$$\text{s.t.} \quad Rx \geq 0$$

where the components of x are unconstrained in sign. This linear program problem will identify the cheapest portfolio that gives a vector of all nonnegative returns. By construction the portfolio x can involve positive or negative positions in each asset. Certainly $x = 0$ is a feasible choice for this problem, and the No Arbitrage Condition implies that it indeed minimizes the objective function. Thus the linear programming problem has a finite solution.

The dual of this linear program is

$$\max \quad \pi 0$$

$$\text{s.t.} \quad \pi R = p,$$

where π is the S-dimensional nonnegative vector of dual variables. Note that since x is unconstrained in sign, the constraints in the dual program are all equalities. See any text on linear programming for a detailed discussion. (The objective function looks a bit odd due to the multiplication by zero, but this is the proper form for the dual problem.)

Since the primal has a finite feasible solution, so does the dual. Thus, a necessary implication of the No Arbitrage Condition is that a nonnegative S-dimensional vector π must exist such that

$$p = \pi R.$$

The sufficiency proof is even easier. Explicitly, the problem is to show that the existence of the nonnegative state prices (π_s) implies that the No Arbitrage Condition must be satisfied. Begin with a portfolio x such that

$$Rx \geq 0.$$

Multiplying each side of this inequality by π and using the fact that $p = \pi R$, we have

$$\pi Rx \geq 0$$

$$px \geq 0$$

which proves the result.

■ *Thanks to Carl Shapiro, Joseph Stiglitz, and Timothy Taylor for helpful comments.*

References

Black, Fischer, and Myron Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 1975, *81*, 637–654.

Breeden, Douglas, and Robert Litzenberger, "Prices of State Contingent Claims Implicit in Option Prices," *Journal of Business*, 1978, *51*, 621–651.

Cox, John, and Stephen Ross, "A Survey of Some New Results in Financial Option Pricing Theory," *Journal of Finance*, 1976, *31*, no. 1, 383–402.

Cox, John, Stephen Ross, and Mark Rubinstein, "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 1979, *7*, 229–263.

Cox, John, and Mark Rubinstein, *Options Markets*. Englewood Cliffs: Prentice-Hall, 1985.

Merton, Robert, "The Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, Spring, 1973.

Ross, Stephen, "A Simple Approach to the Valuation of Risky Streams," *Journal of Business*, 1978, *51*, no. 3, 453–475.

Ross, Stephen, "Risk, Return, and Arbitrage." In Friend I., and J. Bicksler, eds., *Risk and Return in Finance*. Cambridge: Ballinger, 1976.