BASEMENT

MASSACHUSETTS INSTITUTE of TECHNOLOGY
LIBRARIES

WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

DOUBLING STRATEGIES, LIMITED LIABILITY, AND THE DEFINITION OF CONTINUOUS TIME

Mark Latham

WP 1131-80 June 1980

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

DOUBLING STRATEGIES, LIMITED LIABILITY, AND THE DEFINITION OF CONTINUOUS TIME

Mark Latham

WP 1131-80 June 1980

This is a revised version of "The Black-Scholes Economy, Doubling Strategies, and Risky Borrowing" (May 1980). My thanks to Robert C. Merton for suggesting this problem and the approach to solving it, and to Fischer Black, Saman Majd, and Stewart Myers for helpful suggestions.

I. Introduction

What do we mean, and what should we mean, by the standard basic assumptions of continuous-time finance models? I am referring to the following assumptions:

1. Continuous time -- Time is indexed by the real line (or some subinterval thereof), any investor can trade at any point or points in time, and markets clear at all times.

2. Unrestricted borrowing at the market rate of interest -- There is no limit on the amount that an investor can borrow, and the rate does not increase with amount borrowed.

3. Stochastic processes for stock prices that have ex ante uncertainty over every subinterval of time, no matter how small.

4. All investors are price-takers -- Each acts as if his actions have no effect on prices.

I had thought it was quite clear what those assumptions meant, but taken literally they imply a contradiction: Kreps [1979] has pointed out that arbitrage profits can be made by a 'bet-doubling' strategy using one stock and the riskless asset. Markets cannot clear when there is an arbitrage opportunity, so the standard assumptions above contradict each other -- they are not "internally consistent". This is a very serious matter, since contradictory assumptions tend to produce contradictory implications, which can not possibly describe an actual economy; internal

739564

consistency is imperative. Formal algebraic statement of assumptions and theorems can not only lessen the ambiguities of statement in English, but may also help in the difficult search for proofs of internal consistency. Such an approach is not taken in this paper; rather, I pursue the less ambitious goal of amending the assumptions to eliminate the arbitrage from doubling strategies.

The development of continuous-time techniques in finance has led to many elegant and enlightening results that could not be derived in discrete time; these include Merton [1971] and Black and Scholes [1973]. One might wonder if the claim that the basic assumptions are contradictory implies that the results of these models are not valid. I believe the results are valid; this paper amends the assumptions so as to preserve the results that seem reasonable while eliminating those that do not. In particular, the Black-Scholes (B-S) model for pricing options is shown to hold under the new assumptions.

There are several possible approaches to rectifying the doubling problem. While imposing either credit limits or discrete trading would prohibit Kreps' arbitrage strategy, I argue that these are unrealistic and would not in general preserve the classical results as desired. Instead, I replace the assumption of unlimited riskless borrowing by an institutional structure in which all loans have a limited liability feature: no-one can be required to repay more than his total assets. In this setting the Kreps doubling strategy can no longer produce arbitrage profits, but the B-S option pricing formula still holds. An alternative solution to the doubling problem lies in defining a continuous-time economy as the limit of a sequence of discrete-time economies, as the trading interval approaches zero. Merton [1975] argues that the only valid continuous-time results are those that hold in the limit of discrete time. The doubling strategy fails this test, while the B-S formula is shown to hold in the limit. Finally, the most sensible solution may be to combine this limiting definition with the limited-liability loan assumption.

Section II reviews the Black-Scholes economy and two derivations of the B-S formula. Section III outlines the Kreps doubling strategy. Section IV proves the B-S formula in a world of limited-liability demand loans, and shows the flaw of doubling in such a-world. The feasibility of doubling using non-demand (limited-liability) loans is analyzed and rejected in section V. In section VI, I look at the continuous-time limit of discrete-time economies with no limit on liability; a proof of the B-S formula in such a world is presented, and doubling is shown not to "work". Section VII shows that the B-S formula also holds in the limit of discrete-time economies with limited-liability loans. Section VII gives a summary and conclusions.

II. the Black-Scholes Economy

This section is a brief review, with slight changes, of the main arguments in Black and Scholes [1973] and Merton [1977]. I will work with the simplest economy possible for the purposes of this paper, so as to highlight the aspects of special interest; no doubt the results can be generalized to more complex economies.

Let us assume:

1. Capital markets operate (and clear) continuously.

2. Capital markets are perfect --

   (a) all investors are price-takers;

   (b) there are no taxes or transaction costs;

   (c) there are no restrictions on borrowing or selling short;

   (d) there are no indivisibilities.

3. Investors prefer more to less.

4. There is a riskless asset with a continuous yield r that is constant through time.

5. There is a stock paying no dividends, whose price S follows the geometric Brownian motion process

$$dS/S = \mu dt + \sigma dZ$$

where t is time, Z is a standard Wiener process, and μ and σ are known and constant through time.

6. There is a European call option on the stock, with exercise price E, maturity date T, and market price C.

Black and Scholes state that "under those assumptions, the value of the option will depend only on the price of the stock and time and on variables that are taken to be known constants." (p. 641) They did not have a proof for this assertion, although their derivation of the B-S formula depended on its being true. Therefore, I shall also discuss a proof of the B-S result that does not start by assuming that there is some twice-differentiable function F such that $C(t) = F(S(t),t)$ for all t.

To return to the well-known B-S derivation, an instantaneously riskless hedge is formed by buying one call and selling F_S shares of stock.

Using stochastic calculus, the rate of change in the hedge's value (in dollars per unit time) is found to be

$$\tfrac{1}{2}\sigma^2 F_{SS} S^2 + F_t$$

with probability one -- no uncertainty. This must equal the riskless return on investment, $(F - F_S S)r$, giving the partial differential equation:

$$rF_S S - rF + \tfrac{1}{2}\sigma^2 F_{SS} S^2 + F_t = 0 \quad .$$

Solving this subject to the boundary conditions

$$F(S,t) = \max[0, S-E]$$

$$\text{and} \quad F(0,t) = 0 \ ,$$

and a boundedness condition $F/S \leqslant 1$ as $S \to \infty$, gives the B-S call option valuation formula:

$$F(S,t) = SN(d_1) - E\exp\{-r(T-t)\}N(d_2)$$

where N(d) is the cumulative normal distribution function,

$$d_1 = \frac{\log(S/E) + (r+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \ , \text{ and}$$

$$d_2 = d_1 - \sigma\sqrt{T-t} \ .$$

For future reference, let us denote this particular solution for F (i.e. the B-S formula) by G(S,t).

Merton [1977] shows that if C ever differs from G(S,t) then there will be an arbitrage opportunity. This provides a proof of the formula without assuming that the only relevant variables are S and t. If C(0) > G(S(0),0), the arbitrage is achieved as follows: An investor could sell one call, pocket the difference C(0) - G(S(0),0), and invest G(S(0),0) in such a way as to return max[0,S-E] at maturity T with certainty, thus paying off the call. He invests G(S(0),0) by continuously borrowing enough extra money to hold $G_S(S(t),t)$ shares of stock. It can be shown that this strategy ensures that his portfolio is always worth G(S,t), even as S and t change through time. In particular, at maturity it will be worth G(S,T) = max[0,S-E], which must equal the call's market value at maturity, C(T), no matter how wildly it may have been priced in the meantime. So he will be able to pay off the call that he sold at t = 0.

The Black-Scholes option pricing methodology has proven remarkably fruitful in modern finance theory. Its applications have included the valuation of all types of corporate liabilities, the management of natural resources, measurement of market timing skills, and even the analysis of such staggeringly complex contracts as bank line commitments. With these extensive structures resting on the Black-Scholes foundation, clearly some careful attention to its cornerstone assumptions is not unwarranted.

III. Trouble in Paradise: the Problem of Doubling Strategies

The strong assumptions of the original Black-Scholes economy sheltered the analysis from many possible real-world confusions: unknown variance rates, nonexistent variance rates, jump processes, stochastic interest rates, and so forth. Then, armed with an understanding of what it takes to make the hedging argument work, subsequent research has tackled these more general problems with much success. In this era of expanding frontiers I was surprised to learn that back in the pristine garden where it all began there was still a snake in the grass: the bet-doubling strategy outlined by Kreps [1979, pp. 40-41] shows that the assumptions in section II, above, are self-contradictory. He shows a way for any investor to make limitless arbitrage profits, from which we must conclude that market-clearing is impossible. Here is a version of his argument:

Start at $t = 0$. An investor is going to make \$1 or more, with probability one, by the time $t = 1$, with no investment of his own. To do this, knowing μ, σ, and r, he first calculates a probability $q > 0$ and a rate of return $x > r$ such that over any time interval of length less than or equal to 1/2, the stock will yield a (continuously compounded) rate of at least x with probability at least q. (A proof that such q and x exist is given in Appendix. There will be infinitely many possible choices.) He then borrows and invests enough money in the stock that the chance of being up a dollar or more at $t = 1/2$ is at least q. In that case, he puts his profits into the riskless asset until $t = 1$. If he is not up a dollar or more at $t = 1/2$, he borrows and buys enough stock so that by $t = 3/4$, with probability q, he will cover any previous losses and still be up a dollar. Again, if he wins he shifts to the riskless asset, while if he doesn't win he raises the stakes enough to net \$1 by $t = 7/8$ with probability q, and so on. All he needs is to win once, so the chance of being up \$1 by $t = 1 - (1/2)^n$ is $1 - (1-q)^n$, and he is certain to be up \$1 by $t = 1$.

When I first heard of this I thought it must involve cheating somehow, but I can't see where it breaks any of the rules of the game. So I am forced to conclude that, taken literally, the assumptions in section II are not mutually consistent.

IV. One Solution: Limited Liability Demand Loans

How should we amend the B-S assumptions? Surely we do not want to allow arbitrage profits to be achievable, by doubling or any other strategy. What enables doubling to work in the B-S economy, and why would it not work in practise? "Continuous time yields the necessary opportunities to bet; unrestricted short sales yields the necessary resources to cover any losses." [Kreps, p. 41] In an earlier paper, Harrison and Kreps [1979] eliminate the doubling problem by restricting trading to discrete time points; but, as Kreps [1979, p. 43] admits, "economic motivation for it is lacking." Even if we assume that no trading can occur at night when the organized markets are closed, the continuous opportunity for trading from 10 a.m. to 4 p.m. is enough to permit arbitrage by doubling.

Think of a gambler playing a doubling strategy at roulette (betting $1 on red and doubling if he loses). Aside from discrete time, one problem he will face is that the house will have a limit on the size of his bet. I guess the parallel to this in capital markets might be that if a doubler doubles too many times, his position would be so large that the "price-taker" assumption would be questionable. This I will not pursue here. But now suppose our gambler wants to finance his bets by borrowing. In practise, prospective lenders will ask what the money is to be used for, and will turn him down if he tells them. Doubling looks fine, even for lenders, except for that one chance of a string of losses long enough that the money to double again is not forthcoming.

This suggests another possible amendment: putting an upper limit on borrowing. Doubling would no longer 'work', but choosing some level for the credit limit would be somewhat arbitrary, and it may affect equilibrium option prices. If the constraint is binding, investors may use options for leverage. Finally, such a limit would not be realistic: we do not find that investors can borrow at r up to a ceiling on borrowing, but rather that the interest rate charged depends on the term of the loan, the investor's wealth, and what he plans to do with the money. Although we often proceed with unrealistic assumptions, I shall propose an alternative to the credit limit that is both more realistic and, I think, more appealing in its analysis and results.

Append the following to the assumptions of section II:

7. Each economic agent's wealth is always nonnegative and finite, so that no contract can require that an agent pay more than his total assets under any possible circumstances.

8. All loans, short sales, and options are negotiated on a demand basis -- either party may terminate the contract at any time without penalty. In the case of options (American or European), this is meant to mean that the option buyer can give the option back to the seller in return for a sum equal to the market price of a similar option that has no default risk (for example, a covered call would have no default risk).

You may well ask how a contradictory set of assumptions can be made consistent by adding more assumptions. Well, I believe that #7 in fact replaces an implicit assumption that negative wealth was allowed. You may also ask whether a riskless asset can exist in this limited-liability world -- wouldn't every loan have a risk of default? If all else fails, the riskless asset can be defined as a contract for future delivery of $B cash that is secured by $B of cash sitting in a vault. (A similarity to government debt should be apparent.) But there are other sources of riskless assets. Consider a demand loan owed by an investor with a portfolio of stock and option positions, where the portfolio has a market value greater than the loan owed. Since stock prices (and therefore put and call prices) have continuous sample paths with probability one, there is no risk of default on the loan, because if the portfolio value drops close to the amount owed, the loan can be called in. There is no chance of the value instantaneously dropping below the amount owed before the loan can be called; so the riskless rate will be charged. Also notice that no-one would lend to an investor with zero wealth (unless he agreed to repay all the proceeds of his investment), because such a proposition would be dominated by investing directly in whatever the borrower was planning to invest in.

I claim that in the world as here constructed, a doubling strategy will not work but the B-S option formula will still hold. First I will prove the Black-Scholes formula. In this amended economy, it will make a greater difference to the force and form of the proof whether or not one starts by assuming that the call price is a known function of only S and t.

Given this assumption, the original B-S proof carries over almost unchanged. Unless the Black-Scholes partial differential equation holds, any investor with positive wealth can make excess returns on the standard hedge portfolio of stock, call option, and riskless asset. More specifically, denoting the call price function by F(S,t), if

$$rF_S S - rF + \tfrac{1}{2}\sigma^2 F_{SS} S^2 + F_t \equiv b > 0 \quad ,$$

then the investor can buy n calls, sell short nF_S shares of stock, put the difference in the riskless asset, and make \$nb per unit time (n can be chosen as large as he likes). This is in addition to any returns on his own (unborrowed) wealth, which he needed as collateral in order to short-sell with no default risk. In other words, this is an arbitrage opportunity.

The opposite investment strategy is followed if $b<0$. So the B-S p.d.e. must hold. Clearly the boundary conditions are unchanged, so the B-S formula is valid in this economy.

If we imagine that the call price might not be a function only of S and t, the proof is more difficult. Denote the B-S formula by $G(S(t),t)$, the call price by $C(t)$, and an investor's (market-valued) wealth by $W(t)$. Suppose $C(0) < G(S(0),0)$, and $W(0) > 0$. The call is undervalued, so the investor buys n calls, sells short $nF_S(S(0),0)$ shares of stock, and puts the remainder (including $W(0)$) in the riskless asset. He adjusts his short position continuously so that he is always short $nF_S(S(t),t)$ shares of stock. As in Merton [1977], this is designed to ensure that the investor can pay off the calls and still have money left over, but here there is a hitch: If a call can be undervalued, what is to prevent its becoming even more undervalued? This is not a problem for Merton, because his investor can hold out until maturity, when he is sure he will be ahead. But in my case, if $W(t) \leqslant 0$ for some t before maturity, my investor will no longer be able to borrow (or sell short or sell calls), and thus could not maintain his position to maturity -- he would be out of the game.

This puts a limit on n, the size of the position that he may safely take. I need to show that this limit is positive, i.e. that there is a positive position that the investor could take and make excess profits for certain.

Lemma: $-E \leqslant G(S(t),t) - C(t) \leqslant E$ for all t until maturity.

Proof:

(a) $S(t)-E \leqslant C(t) \leqslant S(t)$ for all t. If C(t) broke the upper limit, investors could sell the call, buy the stock, and put up the stock as collateral for the call, making arbitrage gains. If it broke the lower limit, investors could sell the stock short, buy the call, and put up the call and \$E as collateral for the short sale. Perhaps a more convincing argument for the lower limit goes as follows. Given our assumptions, there is no benefit to having possession of a stock earlier rather than later, as long as you are certain to get it (i.e. no dividends, among other things). Under these conditions, an American call is equal in value to a European call. Thinking of C(t) as the price of an American call, if it breaks the lower limit investors could make arbitrage gains by buying the call and exercising it immediately.

(b) $S(t)-E \leqslant G(S(t),t) \leqslant S(t)$ for all t. This is a well-known property of the B-S formula.

(c) The lemma follows easily from (a) and (b). F and C are trapped in an interval of length E, and so can be no more than E apart.

Theorem: If $n < W(0)/\{E-[G(S(0),0)-C(0)]\}$, then $W(t) > 0$ for all t until maturity.

Proof: The constraint on n implies that

$0 < W(0) - En + n[G(S(0),0)-C(0)]$

$\leqslant W(0) + n[G(S(0),0)-C(0)] - n[G(S(t),t)-C(t)]$ (by the Lemma)

$\leqslant \{W(0) + n[G(S(0),0)-C(0)]\}e^{rt} - n[G(S(t),t)-C(t)]$

$= W(t)$

(For the last line of the proof, recall that the investor starts with $W(0)$, skims off $G(S(0),0)-C(0)$ immediately for every call he buys, is long n calls thenceforth, and ensures by continuous hedging that the remainder of his portfolio, short in stock and long the riskless asset, always has value $-nG(S(t),t)$.)

The proof is similar for the case of an overvalued call, in which the limit on n would be $W(0)/\{E-[C(0)-G(S(0),0)]\}$.

Therefore any deviation from the B-S formula allows excess profits to be made, and so the B-S formula must hold. Because of the limit on the size of position an investor can take, I think the above is more in the spirit of a dominance argument than an arbitrage argument. If $C(t) > G(S(t),t)$ then all investors with positive wealth would be wanting to write the call; if $C(t) < G(S(t),t)$ they would all want to buy the call. So the call market could only clear if $C(t) = G(S(t),t)$ for all t.

There remains the question of whether the Kreps doubling strategy can produce sure gains in a world of limited liability. Suppose a would-be doubler starts with wealth $W(0) > 0$. He may borrow as much as he wants, but if ever his net market position $W(t)$ becomes 0 the loan will be called, he will not be able to borrow any more, and so he will have lost permanently. The trouble is that no matter how small an amount he borrows at $t = 0$ to put in the stock until $t = 1/2$, there is a positive probability that the stock will drop enough to make $W(t) = 0$ for some $t < 1/2$. And even if that disaster doesn't strike him, if he has to 'double' several times in a row the stake will have risen so as to increase the risk of hitting zero wealth before the next double. So arbitrage profits can not be made this way.

But that does not prove that there is no way to make excess profits by some other type of doubling strategy. The next section looks at some of these.

V. Doubling With Non-Demand Loans

The doubler who is only allowed to borrow using demand loans may complain that this is an unfair handicap -- that the nature of his strategy is such that if only he weren't interrupted when his wealth drops temporarily below zero, he would in the end pay off all his debts with some money left over. The introduction of such real-world contracts as term loans and lines of credit may yet permit successful doubling. In this section I outline an institutional structure with these features, but conclude that doubling still fails. All effort is not wasted, though, because the rules drawn up for this economy will be useful in sections VI and VII.

Replace assumption 8 of section IV by the following:

8. All kinds of financial contracts are allowed, provided that

   (a) they satisfy the limited-liability assumption, #7;

   (b) they specify how each borrower will invest his assets (loans plus own wealth) until the contract expires, so as to avoid moral hazard problems in pricing contracts; and

(c) all loans and lines of credit are of finite size.

The idea is that if there is a chance that the borrower (or short-seller or option writer) will not have enough wealth to make the prescribed payment, it is understood at the outset that he will only pay as much as he has; and the contract is priced with that in mind. Furthermore, contracts may be written in which the borrower pledges an amount less than his total wealth. This institutional setting seems like one that could evolve in a relatively unregulated marketplace. In fact, the limited liability feature is explicit in many contracts that we observe (e.g. borrowing through corporations), and is implicit in all others because bankruptcy is allowed and slavery is not. Default risk must be the major reason for the difference in interest rates between margin loans and consumer loans.

Notice that the proofs of the Black-Scholes formula given in section IV apply here also, a fortiori. They required the existence of 'demand' securities, not the nonexistence of nondemand securities.

Suppose a doubler financed his stock positions by term loans of maturity equal to the planned holding period for the stock. That is, at $t = 0$ he borrows on a promise to repay at $t = 1/2$. If he is not up \$1 at $t = 1/2$, he borrows on a promise to pay at $t = 3/4$; and so on. This way he can not be put out of the game by hitting zero wealth during a holding period, but there is still a chance that at the end of any holding period he will have to pay all his assets to his creditor. In such a case he will have zero wealth, and be unable to recoup his losses since no-one would then lend to him. Therefore riskless excess profits can not be made this way.

Suppose instead he negotiates a single line of credit of $B for the whole period from t = 0 to t = 1, paying a fee of $f up front, and paying the riskless rate on any amounts borrowed. He then arranges his 'doubling' with the aim of making $1 plus $f plus any interest paid. Here the catch is that he may suffer enough losses that the amount he needs to invest would involve borrowing more than $B. After that he can not increase his stock holdings for more 'doubles', and so can not guarantee to be ahead by t = 1.

If a doubler could arrange an infinite line of credit, he could make arbitrage profits with Kreps' strategy while only borrowing finite amounts, but he can set no sure upper limit in advance on his borrowing, and I have eliminated infinite credit lines by assumption. It would be interesting to analyze what happens in the limit as B approaches infinity.

VI. Limit of Discrete Time, No Limit on Liability

Quite a different approach to resolving the question of doubling is to only allow as valid continuous-time economies those that have the property of being the limit of a sequence of discrete-time economies, as the trading interval h approaches zero. See Merton [1975] for a discussion along these lines. This turns out to eliminate the benefits of doubling while preserving the Black-Scholes result, thus identifying the doubling phenomenon as a singularity of the continuous-time extreme. Here we are not assuming that wealth is nonnegative. Although it is hard to imagine a meaning for negative wealth, we go ahead and assume that investors have derived utility of wealth functions defined on the entire real line.

We will set up a sequence of discrete-time economies with the aim of finding the price of a European call in the continuous-time limit. Suppose without loss of generality that we want to price it at $t = 0$ and it matures at $t = T > 0$. Where n is drawn from the set $\{1,2,3,\ldots\}$, define economy n by the assumptions:

1. Capital markets operate (and clear) only at the dicrete time points $t = 0, h, 2h, 3h, \ldots$ where $h \equiv T/n$.

2. Capital markets are perfect.

3. Investors prefer more to less, and are risk-averse, throughout the range of their utility functions.

4. There is a riskless asset that yields $\$(1+rh)$ on an investment of \$1 for time interval h.

5. There is a stock paying no dividends, whose stock price S follows a stochastic process such that given $S(t)$, the mean of $S(t+h)$ is $S(t)[1+\mu h]$, and the variance of $S(t+h)$ is $\sigma^2[S(t)]^2 h$.

6. There is a European call option on the stock, with exercise price E, maturity date T, and market price $C(t)$.

In addition, we make assumptions about how the sequence of economies behaves for large n. These ensure that the limiting economy has the properties we are interested in:

C1. The parameters r, σ, E, and T are the same in all economies.

C2. I must now refer the reader to Merton [forthcoming], which develops some tools essential for the following analysis. My assumption 5 above is sufficient to guarantee Merton's assumptions E.1, E.2, and E.3. In addition, I assume E.4 which has little or no economic content but makes the analysis easier, and E.5 which ensures that the stock price will have continuous sample paths in the limit as $h \to 0$, with probability one.

With these conditions on the stock price, in the continuous-time limit it will follow the geometric Brownian motion process

$$dS/S = \mu dt + \sigma dZ \ .$$

I shall not attempt to derive a formula for the price of the call option, $C(0)$, in any of the discrete-time economies. Rather, I shall argue that whatever this sequence of prices is, it must converge to the Black-Scholes formula value, $G(S(0),0)$, as $n \to \infty$ and $h \to 0$. To do this, I will not start by assuming that it converges to some function of S and t, but I will assume that it converges.

To simplify notation, let

$o(h)$ stand for any function with the property $\lim_{h\to 0} f(h)/h = 0$;

$S(k) \equiv$ stock price at $t = kh$, for $k = 0,1,2,\ldots,n$;

$C(k) \equiv$ option price at $t = kh$;

$N(k) \equiv G_S(S(k),k) =$ B-S hedge ratio at $t = kh$;

$G(k) \equiv G(S(k),k) =$ B-S formula call value at $t = kh$; and

$H(k) \equiv$ value of hedge position at $t = kh$.

The following is based on proofs in Merton [1977] and [1978].

Suppose $C(0)$ converges to a value greater than $G(S(0),0)$. Then I claim that in an economy with large enough n, arbitrage profits can be made by the following strategy:

At $t = 0$ sell one call for $C(0)$, pocket the difference $C(0) - G(0)$, and invest $G(0)$ in a levered stock portfolio composed of $N(0)$ shares of stock and riskless borrowing of $\$[N(0)S(0)-G(0)]$ to make up the difference. So the portfolio's initial value $H(0)$ will equal the B-S formula call value $G(0)$, but will be less than the call's market price $C(0)$.

The change in the hedge value by $t = h$ will be

$$H(1) - H(0) = N(0)[S(1)-S(0)] - rh[N(0)S(0)-H(0)]$$

and the expected change (viewed from $t=0$) is

$$E_0\{H(1)-H(0)\} = N(0)\mu h - rh[N(0)S(0)-H(0)] + o(h).$$

At each stage ($t = kh$ for $k = 1,2,\ldots,n-1$), the hedge is revised so as to hold $N(k)$ shares of stock, financed by borrowing whenever necessary. The expected change in the hedge value over the next time interval is

$$E_k\{H(k+1)-H(k)\} = N(k)\mu h - rh[N(k)S(k)-H(k)] + o(h).$$

Let $D(t) \equiv H(t) - G(t)$, the difference between the hedge value and the B-S formula option value. Then

$$E_k\{D(k+1)-D(k)\} = E_k\{H(k+1)-H(k)\} - E_k\{G(k+1)-G(k)\}$$

$$= N(k)\mu h - rh[N(k)S(k)-H(k)] - [G_S(k)u+G_t(k)+\tfrac{1}{2}\sigma^2S^2(k)G_{SS}(k)]h + o(h)$$

by Ito's lemma.

But $N(k) \equiv G_S(k)$, and G satisfies the Black-Scholes p.d.e. $G_t + \frac{1}{2}\sigma^2 S^2 G_{SS} = r[G - G_S S]$, so

$$E_k\{D(k+1)-D(k)\} = rh[H(k) - G(k)] + o(h)$$

$$= rhD(k) + o(h).$$

Consider the difference D(t) for some date t between 0 and T. Now $D(t) = D(0) + \sum_{k=0}^{nt/T}[D(k+1)-D(k)]$ plus an error of order h that arises if nt/T is not an integer (so the error disappears as $h \rightarrow 0$). And $D(0) = 0$ by construction. Merton [forthcoming, esp. pp. 20-26] shows that as $h \rightarrow 0$ the error in approximating the remaining series by the series $\sum_{k=0}^{nt/T} E_k\{D(k+1)-D(k)\}$ approaches zero with probability one. So in the limit,

$$D(t) = \lim_{\substack{h \to 0 \\ n \to \infty}} \sum_{k=0}^{nt/T} rhD(k) = \int_{s=0}^{t} rD(s)ds.$$

The solution to this integral equation is $D(t) = D(0)\exp[rt]$. Since $D(0) = 0$, this means that $D(t) = 0$ for all t from 0 to T, and the hedge value H(t) will always equal the B-S formula call value G(t). The clincher is that at $t = T$ when the investor has to make good on the call he sold, he will have to pay $\max[0,S(T)-E]$; but the Black-Scholes formula has the property $G(T) = \max[0,S(T)-E]$ so $H(T) = \max[0,S(T)-E]$ also -- the hedge portfolio will be just sufficient to pay off on the call at maturity. Thus the profits that were pocketed at $t = 0$ were 'free' -- arbitrage.

By a similar argument, if the limit of C(0) as $h \rightarrow 0$ is less than G(0) then arbitrage is achieved by buying the call and shorting the hedge porfolio. So it must be that the B-S formula gives the limit of the call's market price, as the trading interval approaches zero.

As an example of why doubling fails in the limit of discrete time, imagine a doubler betting red on a fair roulette wheel (probability of win = .5), in a discrete-time world where negative wealth is allowed. He finances all his bets with riskless borrowing. To make each bet a fairly priced security, let us say that its outcome is independent of everything else, the riskless rate is zero, and bets pay off at 2 to 1. He has n opportunities to bet between t=0 and t=1, wants to be up \$1 by t=1, and so he first bets \$1 and keeps doubling until he wins once. His net payoff as of t=1 will be \$1 with probability $1-(1/2)^n$, or $\$(1-2^n)$ with probability $(1/2)^n$.

Now the parallel to Kreps' doubling argument would be to next let n be infinite, so that the payoff is \$1 with probability 1 -- the disastrous string of losses gets swept under the rug of zero probability, and demand for doubling is large. But in the sequence of discrete-time economies as n increases, the demand for such a doubling strategy is identically zero (and hence is zero in the limit). This is because the doubling payoff only adds noise to any portfolio's return, and so would not be undertaken by any risk-averse investor [Rothschild and Stiglitz, 1970].

VII. Limit of Discrete Time, With Limit on Liability

Defining continuous time in terms of the limit of discrete time is very appealing in the way it seems to preserve results that make sense while rejecting those that don't. At the same time, the other approach of requiring wealth to be nonnegative is desirable because it is hard to imagine a realistic meaning for negative wealth. This section combines both methods in what may be the most reasonable way to resolve the questions addressed here.

To the six assumptions on economy n in section VI, append the following:

7. Each economic agent's wealth is always nonnegative and finite, so that no contract can require that an agent pay more than his total assets under any possible circumstances.

8. All kinds of financial contracts are allowed, provided that
   (a) they specify how each borrower will invest his assets until the contract expires;
   (b) all loans and lines of credit are of finite size; and
   (c) transactions occur only at the time points $t = 0, h, 2h, \ldots$

The cross-economy assumptions C1 and C2 are also retained.

In this setting, a riskless asset is much harder to come by. In a discrete time interval, any loan on a portfolio with finite pledged wealth and some stock-price risk will have a positive probability of default. Note that in an option-stock hedge portfolio the stock-price risk cannot be completely hedged away in discrete time. Furthermore, any short sale of stock or writing of options will have default risk also. In all these cases, the contracts will be priced to take account of the default risk. Define the default premium to be the difference between the prices today of a given promise to pay with and without the default risk. The price without default risk can be imagined as the limit of the price as the pledged wealth (collateral) approaches infinity or, better still, no-default securities can be invented. The price for a short sale is not a problem -- it should equal the stock price if there is no default risk. A call writer can guarantee not to default by putting up the stock as collateral. And a loan can be riskless if backed by enough cash; although no private investor would want to borrow on such terms, the government might, so let us assume the existence of riskless government debt.

The pursuit of a doubling strategy in this world would suffer from all the handicaps of sections V and VI above, so it will not generate arbitrage profits here either. The remainder of this section is devoted to showing that in spite of default problems in the discrete-time economies, the market price of a call will approach the B-S formula value in the continuous-time limit. My plan is to adapt the discrete portfolio adjustment argument of section VI, and to show that in a time interval of length h, each default probability is o(h), each default premium must be o(h), so that as $h \to 0$ the cumulative effect of these probabilities and premia from $t = 0$ to $t = T$ disappears. Just as in section IV, there will be a limit on the size of position that an investor may take to make money on a mispriced call, but since all investors would be wanting to transact on the call in the same direction, dominance would prevent the market from clearing until the price is right.

Let us take some care in defining just what sort of contracts our discrete-hedging investor is going to invest in. When he buys stock, calls, or bonds, let him do so no a no-default-risk basis: by buying covered calls or government debt (the stock is not a problem). When he sells short, writes calls, or borrows, there will be no way to avoid some default risk on his part, but let the risk be confined to one time unit per contract by means of the following arrangement. Each of his liabilities is to be resettled after each time unit h (in the style of futures contracts). There is nothing unusual in this for borrowing and short sales (borrowings of stock) -- these are to be repaid next period. But when I have him write an option with some distant maturity date T, I in fact will mean for him to contract to pay, at time h from now, the market price that that option will then command if it has no default risk from then on (e.g. the market price of a covered call). This is the next best thing in discrete time to the continuous-time demand-basis assumption #3 on page 9, section IV.

Define the notation $f(h) \sim h^a$ to mean that $\lim_{h \to 0} f(h)/h^a$ is a nonzero real number.

Consider any portfolio of short and/or long positions in a stock, options on that stock, and bonds, with positive pledged wealth (positive market value), where all short positions are to be resettled each period. In order for default to occur on any of the short positions, the stock price must move by an amount $\sim h^0$ in a time interval of length h. With the assumptions we have made on the stock price process, the probability of such an outcome is $o(h)$. [Merton, forthcoming, esp. pp. 12-13]

Now the default premium can be thought of as the market value of a loan guarantee; the guarantor will only have to pay off with probability $o(h)$. Suppose the market value of such a liability were $\sim h^1$ or larger. Then rolling over a sequence of such claims would give, as $h \to 0$, a summed value $\sim h^0$ or greater for a payoff of probability zero -- arbitrage. So the default premium must be $o(h)$.

The hedge strategy to be followed is the same as that in section VI, except that there is a limit on the number of calls that can be bought or sold, for the same reason as in section IV. Here, the expected change in the hedge value from time k to time k+1, conditional on default not occurring, is

$$E_k\{H(k+1)-H(k)\} = N(k)\mu h - rh[N(k)S(k)-H(k)] + o(h)$$

as before, but here the o(h) term includes any default premium. With D(t) defined as before, we again get

$$D(t) = \lim_{\substack{h \to 0 \\ n \to \infty}} \sum_{k=0}^{nt/T} rhD(k) = \int_{s=0}^{t} rD(s)\,ds$$

conditional on there being no default between time 0 and time t. But the probability of such a default is bounded above by something of the form $\sum_{k=0}^{nt/T} o(h)$, which approaches 0 as $h \to 0$. So we have $D(t) = D(0)\exp[rt]$ with probability one, the call position can be closed at maturity using only the proceeds from the hedge portfolio, and riskless gains are made. This completes the dominance argument for the Black-Scholes formula in this setting.

VIII. Conclusions

The Kreps doubling strategy raised doubts about the internal consistency of the standard assumptions in continuous-time finance models. Two ways of resolving the doubling problem were presented: requiring wealth to be always nonnegative, and defining a continuous-time economy as the limit of a sequence of discrete-time economies, as the trading interval approaches zero. These two methods were combined in the last section, to form what seems to be a more reasonable model of the world. In all three cases, the doubling strategy fails to produce arbitrage gains, while the Black-Scholes option pricing argument still obtains. It is hoped that the amended assumptions (any of the three sets) will provide a general partition of results that make sense from those that do not.

Appendix

Theorem: Given the stock price process

$$dS/S = \mu dt + \sigma dZ$$

and the (constant) riskless rate of return r, there exist a probability $q > 0$ and a rate of return $x > r$ such that over any time interval of length $T \leqslant 1/2$, the stock will yield a rate of at least x with probability at least q. (All rates of return are meant as continuously compounded.)

Proof: Let y denote the (random variable) rate of return on the stock in time T. Then $S(T) = S(0)\exp[yT]$, and $y = \frac{1}{T}\log S(T)/S(0)$. For given S(0), the lognormal distribution of S(T) implies that log S(T) is distributed normally with mean = $\log S(0) + aT$ and variance = $\sigma^2 T$, where $a \equiv \mu - \sigma^2/2$. Therefore y is distributed normally with mean = a and variance = σ^2/T.

Case 1 -- $a > r$

Choose $x = a$ and $q = 1/2$. Then clearly for any T (including $T \leqslant 1/2$) the probability that $y \geqslant x$ is 1/2 ($\geqslant q$), because x is the mean of y's normal distribution.

Case 2 -- $a \leqslant r$

Choose $x = r + 0.001$ and $q = \text{Prob}\{y \geqslant x \text{ for } T=1/2\}$. For $T=1/2$ the condition on q holds by the definition of q. (Clearly $q > 0$.) For smaller T, it is intuitively obvious that since

(a) the mean of y is unchanged,

(b) the variance of y is increased,

and (c) x is greater than the mean of y, it must be true that $\Pr\{y \geqslant x\}$ increases, and so is $> q$. Here is a formal proof:

$$q = \Pr\{y(1/2) \geqslant x\}$$

$$= \Pr\left\{\frac{y(1/2)-a}{\sigma\sqrt{2}} \geqslant \frac{x-a}{\sigma\sqrt{2}}\right\}$$

Note that both have the standard normal distribution.

$$= \Pr\left\{\frac{y(T)-a}{\sigma/\sqrt{T}} \geqslant \frac{x-a}{\sigma\sqrt{2}}\right\} \quad \text{for all } T, \text{ including } T<1/2.$$

$$< \Pr\left\{\frac{y(T)-a}{\sigma/\sqrt{T}} \geqslant \frac{x-a}{\sigma/\sqrt{T}}\right\} \quad \text{for any } T < 1/2 \text{ , since } \frac{x-a}{\sigma\sqrt{2}} > \frac{x-a}{\sigma/\sqrt{T}} .$$

$$= \Pr\{\, y(T) \geqslant x\} \ .$$

Notes:

1. To calculate the amount that our "bet-doubler" should borrow and invest in the stock each time, let

$n \equiv$ the number of consecutive losses he has suffered so far,

$W \equiv$ his net winnings up to that time (may be negative), and

$T \equiv (1/2)^{n+1}$ = the length of his next holding period.

To be up \$1 with probability at least q at the end of the next holding period, he borrows and invests $\$\left[\dfrac{1-W}{e^{xT}-e^{rT}}\right]$. Then if $y(T) > x$, his net position will be

$$W + \left(\frac{1-W}{e^{xT}-e^{rT}}\right)\left(e^{yT}-e^{rT}\right) > W + \left(\frac{1-W}{e^{xT}-e^{rT}}\right)\left(e^{xT}-e^{rT}\right) = 1 \ .$$

2. A simpler bet-doubling strategy, which can be defined and analyzed without reference to a specific stochastic process, is found by making use of short sales. It almost works (but not quite) to set $q = 1/2$, to sell short a large enough amount whenever the median of the stock holding-period

return is less than r, and to go long when the median exceeds r; but this fails if the median equals r. Instead, set q = 1/4, go long when the upper quartile exceeds r, and in the case when it does not exceed r go short an amount calculated with reference to the lower quartile of the stock holding-period return. This works unless both quartiles equal r, a possibility which I think can be safely ignored.

3. Kreps' statement that "there is some positive probability q that over any time interval the second security [stock] will strictly outperform the first [riskless asset]" [1979, p. 40] is not quite accurate. If $a < r$ then for any $q > 0$ there exists a time interval T large enough that the probability of the stock beating the riskless asset is less than q. But this is not a problem for his doubler, who can put an upper bound on the lengths of his holding periods.

References

Black, F., and M. Scholes [1973], "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, vol. 81, 637-659.

Harrison, J., and D. Kreps [1979], "Martingales and Arbitrage in Multiperiod Securitics Markets", *Journal of Economic Theory*, vol. 20, no. 3, pp. 381-408.

Kreps, D. [1979], "Three Essays on Capital Markets" (especially essay III, "Continuous Time Models"), Institute for Mathematical Studies in the Social Sciences (IMSSS), Technical Report No. 261, Stanford University.

Merton, R. C. [1975], "Theory of Finance from the Perspective of Continuous Time", *Journal of Financial and Quantitative Analysis*, vol. 10, pp. 659-674.

Merton, R. C. [1977], "On the Pricing of Contingent Claims and the Modigliani-Miller Theorem", *Journal of Financial Economics*, vol. 5, 241-249.

Merton, R. C. [1978], Seminar notes for "On the Mathematics and Economic Assumptions of Continuous-Time Models", unpublished manuscript.

Merton, R. C. [forthcoming], "On the Mathematics and Economic Assumptions of Continuous-Time Models", forthcoming in *Financial Economics: Essays* *in Honor of Paul Cootner*, W. F. Sharpe (ed.).

Rothschild, M., and J. E. Stiglitz [1970], "Increasing Risk: I. A Definition", *Journal of Economic Theory*, vol. 2, pp. 225-243.

Date Due

FEB 2[illegible]

SEP 4 [illegible]

SEP 12 '88

OCT 9 1986

APR 6 '87

[illegible]

SE 19 '89

NOV 15 1989

JAN 2[illegible]

FEB 28 1992

March 27 '12

JUL 24 1992

MAY [illegible]

NOV. 25 1990

MAR 18 2002

MAR 23 2002

Lib-26-67

ACME
BOOKBINDING CO., INC.

DEC 8 1983

100 CAMBRIDGE STREET
CHARLESTOWN, MASS.

MIT LIBRARIES
3 9080 004 524 226

HD28.M414 no.1129- 80A
Barley, Stephe/Taking the burdens :
739858 D*BKS 00132368
3 9080 001 992 640

HD28.M414 no.1129- 80B
Roberts, Edwar/Critical functions :
739558 D*BKS 00132650
3 9080 001 996 203

HD28.M414 no.1130- 80
Schmalensee, R/Economies of scale and
739574 D*BKS 00132648
3 9080 001 996 161

HD28.M414 no.1130- 80A
Sterman, John /The effect of energy de
739856 D*BKS 00132625
3 9080 001 995 726

HD28.M414 no.1131- 80
Latham, Mark. /Doubling strategies, li
739564 D*BKS 00153013
3 9080 002 261 649

HD28.M414 no.1131- 80 1981
Latham, Mark. /Doubling strategies, li
743102 D*BKS 00152665
3 9080 002 257 613

MIT LIBRARIES
3 9080 004 524 242

HD28.M414 no.1132- 80
Keen, Peter G./Building a decision sup
740317 D*BKS 00136494
3 9080 002 041 785

HD28.M414 no.1132- 80A
Wong, M. Antho/An empirical study of t
739570 D*BKS 00152666
3 9080 002 257 639