1. Introduction

Ever since Arrow's (1964) classic paper, "The Role of Securities in the Optimal Allocation of Risk-Bearing," economists have recognized that a relative handful of financial securities, traded in dynamic fashion, can give far-sighted consumers the opportunity to fit their consumption plans more closely to a complete-markets ideal than might at first be imagined. Criticisms of this perspective are well known, focusing perhaps most of all on the extreme level of foresight about future prices that is required; recall that, in Radner's (1972) "equilibria of plans, prices, and price expectations," individual consumers must have accurate expectations of concerning future equilibrium prices. These same basic ideas have been adopted as foundational to asset-pricing theory in finance, where dynamic trading in a few securities is held to determine, by arbitrage, the prices of many options and other types of contingent claims.

Perhaps the ultimate expression of this idea comes from the literature dealing with the Black–Scholes–Merton (BSM) model of securities markets (Black and Scholes, 1973; Merton, 1973a). In this literature, and in particular in Harrison and Kreps (1979) and Harrison and Pliska (1981, 1983), continuous trading in two assets, a bond (with a certain interest rate) and a stock whose probabilistic law is geometric Brownian motion, is "shown" to provide markets that are complete with regard to all well-behaved contingent claims that are written on the full history of the stock price. The scare quotes around "shown" are there because this story is more nuanced and complex than is this commonly held 30-second version of what is, in fact, shown (see Chapter 3).

To come to the amazing conclusion that markets are complete requires deep mathematics and, in particular, a model of information flow that is hard (at least, for me) to comprehend, as well as trading strategies that are hard to imagine. (To synthesize, say, a European call option requires an infinite volume of trade.) But Sharpe (1978) and Cox, Ross, and Rubinstein (1979) provide a discrete-time, discrete-state analog: There are two securities, a bond with a certain interest rate and a stock, which trade, one against the other, at a discrete list of times. As long as the stock's next trading-time relative price, given today's price, has only two possible values – the so-called binomial case – markets are complete. And if one looks at a particular sequence of these binomial economies, where

https://doi.org/10.1017/9781108626903.002 Published online by Cambridge University Press

trading is allowed (in each economy) at a discrete number of times, but more and more frequently along the sequence, the sequence of economies "converges to" the BSM model. While the limit results are not trivial, the mathematics is not that difficult. In a sense, then, one can regard the BSM model as an *idealization* of discrete-time binomial models with frequent trading, in the same sense that the model of perfect competition is an idealization of markets in which market participants have some, but very limited, market power.[1] This provides a better sense of what sort of information flow is entailed – albeit a very special flow of information – and it provides a reasonable case that modeling trading strategies with infinite trading volumes is "close to" what happens with much less outrageous trading strategies. (Both "converges to" and "close to" need formal, exact statements, which is the reason for the scare quotes. This is the topic of Chapter 4.)

However, as in the case of recent work concerning continuous-time models and their discrete-time (asymptotic) analogs,[2] there is a seeming problem: If the (discrete-time) stock-price process has not two but three (or more) possible next-time values, even if the limiting stock price converges in the standard probabilistic sense to the BSM model, markets are incomplete and many contingent claims are not priced by arbitrage. One can still employ arbitrage arguments, but the arbitrage bounds on simple contingent claims that are implied remain wide as the security price process converges to BSM. Either one must put into the economy a more securities ($n-2$ more "independent" securities, if the greatest number of next-time prices is n), or resort to pricing arguments that involve Sharpe ratios (Cochrane and Saa-Requejo, 2000), or gain-loss ratios (Bernardo and Ledoit, 2000) to show that the BSM model is a proper idealization of these discrete-time models, as trading becomes more frequent.

I say "seeming problem," because the problem is more apparent than real. Consider, for instance, the following model of a market with two securities. Both securities trade at times $t = 0, 1/400, 2/400, \ldots, 399/400$. The first security is a bond with, for convenience, zero interest rate: Its price at each time t is 1, and at time 1 it pays the bearer 1. The second security is a stock whose prices at different times t (and whose terminal value at time 1) are constructed as follows: For a sequence of four-hundred

[1] See, for instance, Novshek (1985).

[2] See, for instance, Fudenberg and Levine (2009) and Sadzik and Stacchetti (2015). The issues that arise in those papers are different from the issues addressed here; I'll briefly discuss their work in the final chapter.

independent and identically distributed random variables $\tilde{\zeta}_k$, where each $\tilde{\zeta}_k$ has distribution given by

$$\tilde{\zeta}_k = \begin{cases} 0.075, & \text{with probability } 2/9, \\ 0, & \text{with probability } 5/9, \text{ and} \\ -0.075, & \text{with probability } 2/9, \end{cases}$$

the price of the stock at time $t = k/400$ is given by

$$\tilde{S}(k/400) = \exp\left[\sum_{j=1}^{k} \tilde{\zeta}_j\right],$$

where $\tilde{S}(0) = 1$ with certainty and, at time 1, the terminal value of the stock (the dividend it pays, say), is $\tilde{S}(1)$ (that is, $\tilde{S}(400/400)$, defined as above). In words, the stock price is a geometric random walk, where each step has positive probability of going up, down, or staying the same. The theory of such things is developed in detail in Chapter 2, but I trust that most readers will be aware of the following fact about this simple model of a two-asset security market: Because there are only two financial securities and, at each time, three ways the stock price can evolve, trading in the two securities will not permit the construction of many contingent claims. In particular, if we look at archetype contingent claim for this literature – a European call option with exercise price 1 – the range of prices at time 0 for the call option that are consistent with the price processes of the stock and bond is quite wide; in fact, one can compute that range to be from 0 up to approximately 0.5954.

However, suppose that a consumer living and trading in this two-security world wishes to synthesize the call option. She starts at time 0 by buying 0.69145 shares of the stock and selling 0.30906 bonds, for a net cost to her of 0.38239; these are the opening positions that she would take were she living in the world of BSM (where the bond has interest rate 0 and the stock is geometric Brownian motion for a standard zero-drift, unit-infinitesimal-variance Brownian motion). Subsequent to taking this opening position, at each time $t = k/400$ for $k = 1, \ldots, 399$, she rebalances her portfolio, changing her stock holding to the level she would hold in the BSM world as a function of the price of the stock, financing any required purchases of stock by selling bonds, and using the proceeds of any sale of stock to buy bonds. How does she fare?

This depends on the realized sequence of stock prices. Using simple Monte Carlo methods, I simulated sequences of stock prices and this

https://doi.org/10.1017/9781108626903.002 Published online by Cambridge University Press

trading strategy, with remarkable results, at least for anyone who believes that, in this trinomial world, a consumer can't get close to synthesizing the call option: Figure 1.1 shows the results of 500 simulations, where the final value of the consumer's portfolio is graphed against the final stock price. This is very, very close to synthesizing the call option.

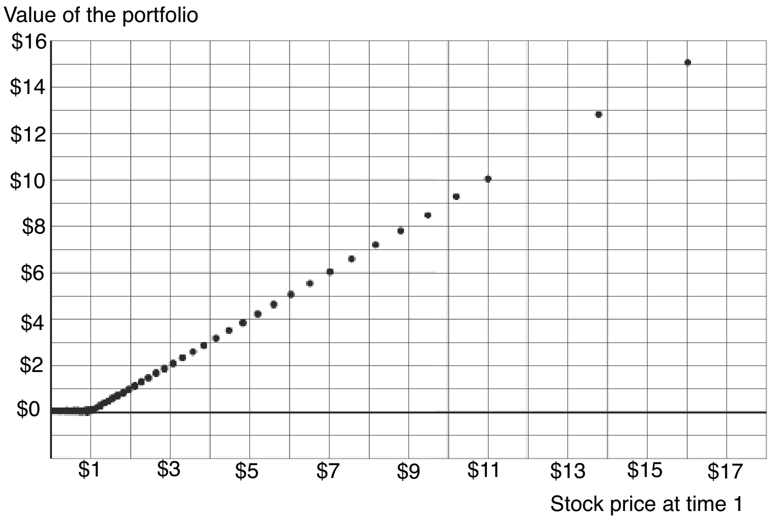

Figure 1.1. Scatter plot of stock price versus value of a portfolio that imperfectly synthesizes a European call option with exercise price 1, using the trinomial (tilde) model for $n = 400$. This shows the results of 500 rounds of simulation.

This is not perfect synthesis of the call option. But when we return to this example, we'll see that a more realistic set of "arbitrage bounds" on the price of the call option, generated by constructing dynamic hedges from the stock and bond, is more on the order of 0.38239 plus or minus 0.02. Because we have two financial securities and, at each time, three possible next positions, markets are incomplete. But they come close to being complete. Going back to Arrow's original insight that frequent trading in a few financial securities can substantially complete markets, perhaps, in this trinomial world, this insight still is substantially valid.

This is not my discovery. The community of scholars (primarily probabilists) who work in the field of financial mathematics have shown that, in a mathematical sense, the incompleteness of markets for this discrete-time trinomial model (and many more models besides) asymptotically vanishes as trading becomes more frequent; Figure 1.1 comes as no surprise to them.

https://doi.org/10.1017/9781108626903.002 Published online by Cambridge University Press

This stream of literature is largely under-appreciated by mainstream scholars in finance and economic theory, perhaps because this literature has significant entry barriers in terms of mathematical sophistication. In this monograph, I try to lower those barriers enough so that mainstream economists are not surprised by Figure 1.1, but, instead, understand what can (and cannot) be said along these lines in terms of mathematics. And I try to put the mathematical results on a sound footing in terms of what they say about economics.

I first review the basic foundations of this literature. The story I wish to tell – how, how well, and when the BSM model idealizes discrete-time trading models with frequent trading – doesn't begin until Chapter 4, more than 50 pages in. Many readers will already know the material covered in Chapters 2 and 3 (concerning discrete-time and continuous-time models, respectively). But I think it helps to present a unified version of the full story, beginning with the theory in discrete and continuous time, separately.

Proofs

Throughout, I try to provide detailed and solid proofs. I do not prove everything; in particular, for deep mathematical results (e.g., the uniqueness of an equivalent martingale measure for BSM, Donsker's Theorem, the Skorohod Representation Theorem), I cite sources where the results can be found. But, with one exception, when it comes to results that are specific to the story I want to tell, I provide details. (The exception is the very complex proof of Proposition 5.1*b*, for which I provide a sketch and then a reference.) For the most part, a reader who is well versed in real and convex analysis should be able to follow what goes on. But the proofs – especially those that invoke Taylor's Theorem – are tedious. Readers who want to get the gist of this theory can safely skip the details, although I hope they will skim the proofs to get a sense of the logical flow being employed.

Website

A public website at the URL *discrete2continuous.stanford.edu* contains a variety of supplementary materials, including a list of errata (as they are discovered), further pertinent references (as they are suggested), and (I hope) notes on the resolution of some of the questions left open in this monograph. I suggest that readers visit the website to see what is there.