
The recent turmoil in financial markets has been partly caused by insufficient attention to rigorous financial modeling. Among the causes of this failing is the relative shortage of mathematically well trained professionals in the financial services industry. Shreve is a co-founder of one of the oldest and most successful masters degree programs in financial engineering, established at Carnegie-Mellon University in 1994. The lecture notes on which this book was based were tested and honed by Shreve over many years of teaching in this Computational Finance program. The result is a remarkable piece of pedagogy and a great service to all entrants to the field.

What follows is a review of Steven Shreve's masterful two-volume text, *Stochastic Calculus for Finance*, which introduces students to stochastic calculus as a tool for financial derivative pricing. I will begin with a brief outline of the nature of the subject and some of the major historical milestones, and then explain why I believe that Shreve’s text is the ideal introduction to the topic.

**THE STOCHASTIC INTEGRAL AS A MODEL OF TRADING PROFITS**

Michael Harrison, whose role in the development of the subject will come up shortly, once remarked to me that stochastic calculus has the appearance of having been expressly designed as a tool for financial analysis, so naturally does it fit the application. Stochastic calculus is now the language of pricing models and risk management at essentially every major financial firm, and it is the backbone of a large body of academic research on asset pricing, corporate finance, and investor behavior.

The typical stochastic-calculus-based financial model describes the random variation of the market price, say $X_t$ at time $t$, of some financial asset. For proper foundations, one fixes a probability space $(\Omega, \mathcal{F}, P)$ (a measure space with $P(\Omega) = 1$), as well as a filtration $\{\mathcal{F}_t : t \in [0, \infty)\}$ of sub-$\sigma$-algebras of $\mathcal{F}$ that determines the timing of the revelation of information. The “usual conditions” for a filtration are laid out, for example, by Protter [19]. One may loosely view $\mathcal{F}_t$ as the set of events (elements of $\mathcal{F}$) whose outcomes are certain to be revealed to investors as true or false by (or at) time $t$. For any event $A$, the probability assigned to $A$ by investors is $P(A)$. The price process $X = \{X_t : t \in [0, \infty)\}$ is adapted to the filtration, meaning that $X_t : \Omega \to \mathbb{R}$ is a random variable whose outcome is revealed to investors at or before time $t$.

---

2000 Mathematics Subject Classification. Primary 60-01, 60H10, 60J65, 91B28.

I am grateful for conversations with Julien Hugonnier and Philip Protter, for decades worth of interesting discussions with Mike Harrison, and also for the patient encouragement of the editor, Bob Devaney.

A process $X$ is adapted if, for all $t$, $X_t$ is $\mathcal{F}_t$-measurable.

©2008 American Mathematical Society

Reverts to public domain 28 years from publication
Occasionally, one hears that market efficiency implies that the price process \( X \) must be a martingale, meaning essentially that the current price \( X_t \) is a conditionally unbiased predictor of the price \( X_u \) at any future time \( u \). This is a misconception: Investors would generally not take the risk of owning an asset unless they are compensated with expected returns. Beyond compensation for risk, even a risk-free asset must offer a return that compensates the investor for tying up capital. Allowing for nonzero expected price changes, it is therefore natural to treat the price process \( X \) as, loosely speaking, a “martingale plus something,” or, to pick a precise and natural definition, a semimartingale. The most classical example of a semimartingale used in financial modeling, suggested in 1965 by the economist Paul Samuelson, is a geometric Brownian motion, which we will get to later.

A trading strategy \( \theta \) determines the quantity \( \theta_t(\omega) \) of the asset held in each state \( \omega \in \Omega \) and at each time \( t \). For a model of a well-functioning market, it is crucial to rule out trading strategies that are based on advance knowledge of price changes, for there would otherwise be arbitrages, meaning trading strategies with unlimited profits at no risk. The natural corresponding measurability restriction is that \( \theta \) is a predictable process. Given a price process \( X \) and a trading strategy \( \theta \) satisfying technical conditions, the total financial gain \( \int_s^t \theta_u \, dX_u \) between any times \( s \) and \( t \geq s \) is defined as a stochastic integral.

An elemental type of trading strategy is a “buy-and-hold” strategy \( \theta \), which initiates a position immediately after some stopping time \( T \) and closes it at some later stopping time \( U \). For a position size \( \theta \) that is \( \mathcal{F}_T \)-measurable, the trading strategy \( \theta \) is defined by \( \theta_t = 1_{\{T < t \leq U\}} \theta \). The total gain from trade for this buy-and-hold strategy is naturally \( \int_0^U \theta_t \, dX_t = \theta(U-X_T) \), the position size multiplied by the interim price change. The gain from trade for a general stochastic trading strategy can be defined as the total gain of an approximating portfolio of buy-and-hold strategies, in a particular limiting sense. For the cases most commonly encountered in financial applications, based on Brownian motion, Shreve gives a very clear explanation of this limit in his second volume, Continuous-Time Models, Sections 4.2–4.3.

A typical financial model allows for \( n \) different securities, with price processes \( X_1, \ldots, X_n \). An investor can choose an associated \( n \)-dimensional trading strategy \( \theta = (\theta_1, \ldots, \theta_n) \) from some allowable set \( \Theta \), determining the total gain-from-trade process

\[
\int \theta_t \, dX_t = \sum_{i=1}^n \int \theta_{it} \, dX_{it}.
\]
In addition to incorporating technical restrictions under which these stochastic integrals are well defined, the allowable set $\Theta$ can enforce budget limits, credit constraints, short-sales limitations, or various other natural investment restrictions.

Significant strands of research literature address the following two classes of problems:

- Given some “utility” functional $U$ on the space of potential gains from trade, solve the optimization problem $\sup_{\theta \in \Theta} U\left(\int_0^t \theta_s \, dX_s \right)$. The utility functional $U$ can encode preferences regarding risk, intertemporal substitution, and the timing of information about trading gains, among other properties.
- Apply the laws of supply and demand to characterize the price processes of the available financial securities. A minimal restriction on the behavior of prices is the absence of arbitrages; demand and supply could never be matched in the presence of an arbitrage.

Both volumes of Shreve’s text focus on arbitrage-free asset pricing, perhaps because most of the students for whom he has written the book are aiming for business careers in finance. The theory of optimal investment, while a significant subject area in academia, has achieved much less traction in business practice. In an earlier book, *Mathematical Finance*, Shreve and his frequent collaborator Ioannis Karatzas provide a detailed treatment of mathematical models of optimal investment.

The field of finance is replete with many other applications of stochastic calculus, such as the financial policies of corporations, the design of new securities, and risk management, which usually involves control or characterization of the left tail of the probability distribution of the gain from trade $\int_0^t \theta_s \, dX_s$.

**The Beginnings of Stochastic Calculus**

Even as early as 1900, Louis Bachelier had introduced Brownian motion as a financial price process. In 1905, Albert Einstein, unaware of Bachelier’s prior work, suggested the name “Brownian motion” and characterized its essential properties. (On nonoverlapping time intervals, the increments of a Brownian motion are independent and normally distributed with means and variances that are each in fixed proportions to the lengths of the time intervals.) Norbert Wiener showed the existence of such a process in 1923. Paul Lévy and Andrei Kolmogorov provided successively fuller developments of the Brownian model.

In 1944, Kiyoshi Itô laid the foundations for stochastic calculus with his model of a stochastic process $X$ that solves a stochastic differential equation of the form

$$X_t = X_0 + \int_0^t \mu(X_s) \, ds + \int_0^t \sigma(X_s) \, dB_s,$$

where $B$ is a standard Brownian motion\(^8\) and where $\mu$ and $\sigma$ are functions from $\mathbb{R}$ to $\mathbb{R}$ satisfying some technical condition. (A Lipschitz condition suffices.) As a

---

\(^8\)A standard Brownian motion $B$ is a Brownian motion whose initial condition $B_0$ is zero with probability one, and whose increments have zero expectation and have a variance equal to the length of the time interval.
generalization of (1), Itô then characterized a class of processes of the form

\[ X_t = X_0 + \int_0^t H_s \, ds + \int_0^t V_s \, dB_s, \]

for adapted processes \( H \) and \( V \) satisfying suitable technical conditions. Such a process is now called an “Itô process,” a special case of what later became known as a semimartingale. Itô’s next important achievement, in 1951, was to establish that, for any smooth function \( f : \mathbb{R} \to \mathbb{R} \),

\[ f(X_t) = f(X_0) + \int_0^t \left( f'(X_s) H_s + \frac{1}{2} f''(X_s) V_s^2 \right) \, ds + \int_0^t f'(X_s) V_s \, dB_s. \]

Itô’s formula (3), now named for him, has important generalizations. In his notes to Chapter 4 of Volume II, Shreve relates that Vincent Doeblin, a French soldier in World War II, independently discovered essentially the same result in 1940. This only came to light in May 2000. As a result, throughout his book Shreve has called (3) the “Itô–Doeblin formula.”

Jarrow and Protter in [14] offer a colorful and much more complete history of these developments up to the time of modern financial theory.

THE MODERN ERA OF FINANCIAL ASSET PRICING

In 1969, Robert C. Merton introduced stochastic calculus to finance, indeed to the broader field of economics, beginning an amazing decade of developments, most famously the formula of Black and Scholes in [2] for the price of an option that conveys the right to buy an asset at some future time \( T \) at a fixed price of \( K \). If the underlying asset has the price process \( X \), the effective payoff of the option at time \( T \) is \( \max(X_T - K, 0) \).

Black and Scholes chose a setting of constant interest rates, and followed Samuelson’s lead by taking \( X \) to be geometric Brownian motion, the solution of a stochastic differential equation of the simple form

\[ dX_t = \mu X_t \, dt + \sigma X_t \, dB_t, \]

for constants \( \mu \) and \( \sigma \). The constant \( \mu \) represents the expected rate of return on the asset, while the constant \( \sigma \) is known as the “volatility.”

Black and Scholes originally based their option pricing formula on a restrictive theory of demand and supply in market equilibrium. Robert Merton provided a revolutionary alternative derivation of the formula by pointing out that the option is actually redundant: The option payoff \( \max(X_T - K, 0) \) can be replicated by trading the underlying asset, borrowing and lending over time as needed. Merton’s replicating trading strategy has an initial market value that must be the price of the option. (If not, for example, if the option price is higher, one could arbitrage by selling the option, investing in the option replication strategy at a lower cost than the price of the option, and using the proceeds of the replication strategy at time \( T \) to cover the payoff required on the short option position.) The initial cost of the replication strategy is indeed given by the same Black–Scholes formula! In their path-breaking journal article, Black and Scholes gave both their original market-equilibrium-based justification, as well as Merton’s more general arbitrage-based argument.

9Shreve refers to “W. Doeblin,” but as pointed out in [14], Wolfgang Doeblin changed his name to Vincent Doeblin.
Although Black and Scholes initially had difficulty getting their paper published, the option pricing model was eventually the basis of a Nobel prize awarded to Robert Merton and Myron Scholes in 1997. Sadly, Fischer Black had not survived to receive the prize.

Once “the formula” was digested and researchers recognized the power of stochastic calculus for analyzing business and theoretical problems, the range and depth of applications expanded rapidly. Among the most important early applications, beyond option pricing, were dynamic models of the term structure of interest rates, beginning with those of Vasicek in [24] and Cox, Ingersoll, and Ross in [3].

As these developments unfolded, Cox and Ross in [4] noticed that, without loss of generality for the purpose of pricing a large class of derivative securities such as options, one could invariably get the correct result by pretending that all of the securities have an expected rate of return that is equal to the current risk-free interest rate. This was initially surprising, for the actual mean rate of return of a financial asset is sensitive to its risk; assets that impose adverse risk on investors demand higher mean rates of return.

The notion of “risk-neutral” pricing was given abstract foundations by Harrison and Kreps in [10], who formalized (under conditions) the near equivalence of the absence of arbitrage with the existence of some new “risk-neutral” probability measure $Q$ under which all expected rates of return are indeed equal to the current risk-free rate. In particular, this allows one to price any security as the $Q$-expected sum of its future discounted cash flows, a dramatic practical simplification that was almost immediately adopted in a wide range of applications in the financial services industry. This new “risk-neutral” pricing method augmented, and in many cases largely displaced, the partial-differential-equation (PDE) approach to derivative pricing that had been used by Black and Scholes in [2].

As an illustration, suppose that the risk-free rate is a constant $r$ and that a stock has an Itô price process $X$ of the form (2). We are interested in the price of an option that conveys the right to buy the stock at some time $T$ at a price of $K$. By the definition of a risk-neutral probability measure $Q$, the initial price of the option must be

$$E^Q \left( e^{-rT} \max(X_T - K, 0) \right),$$

where $E^Q$ denotes expectation with respect to the measure $Q$. For the geometric Brownian model (1) of $X$, Harrison and Kreps applied Girsanov’s Theorem, laid out in Chapter 5 of Shreve’s second volume, to show that $\log X_T$ is normally distributed under $Q$ with mean $(r + \sigma^2/2)T$ and variance $\sigma^2T$. The explicit calculation of (5) is then routine, and of course matches the Black–Scholes option pricing formula that was originally derived by solving a PDE.

Harrison and Pliska in [11] pushed on, placing the theory of trading gains squarely on the more general foundation of stochastic integration with respect to semimartingales. Only much later did Delbaen and Schachermayer in [6] finally establish essentially minimal conditions connecting the notions of risk-neutral probability measures with the absence of arbitrage. These theoretical developments are not covered in Shreve’s text; they are accessible in full generality only to more advanced researchers.

Meanwhile, application after application, in both academic research and industry practice, was built on the familiar template of arbitrage-free pricing and risk-neutral probability measures.
Volume I. The Binomial Asset Pricing Model

Shreve’s wonderful two-volume treatment of the topic can be viewed as two self-contained books. Throughout the first volume, The Binomial Asset Pricing Model, the price of an asset changes only at integer dates, and, conditional on its current level, the next price change has only two possible directions, “up” and “down.” In this “binomial” framework, measure-theoretic technicalities are easily avoided, making this an ideal setting for an introduction to the modeling of derivative prices.

Under the assumption that $X$ is a Markov process under the risk-neutral probability measure $Q$, the analysis is dramatically simplified, as follows. Suppose that a derivative security pays $g(X_T)$ at time $T$, for some payoff function $g$. For our previous example of an option, $g(x) = \max(x - K, 0)$. The definition of a risk-neutral probability measure implies that the initial price of the derivative security is $E^Q(e^{-rT}g(X_T))$.

In order to compute the initial derivative price in the binomial setting, suppose that at a given time $t$ and current price level $x$, the underlying price jumps up by the factor $U_{x,t}$ with risk-neutral probability $p_{x,t}$, and jumps down by the factor $D_{x,t}$ with probability $1 - p_{x,t}$. Then the associated derivative price $f(x,t)$ must satisfy the recursion

$$f(x,t) = e^{-rt} (p_{x,t} f(xU_{x,t}, t+1) + (1 - p_{x,t}) f(xD_{x,t}, t+1)).$$

Given the boundary condition $f(x,T) = g(x)$, this allows an easy calculation of the initial derivative price $E^Q(e^{-rT}g(X_T))$.

For the trivial case of a derivative that is a claim to the underlying asset itself, the price is $f(x,t) = x$ for all $x$ and $t$, so we can apply (6) in this case and calculate that

$$p_{x,t} = \frac{U_{x,t} - D_{x,t}}{e^r - D_{x,t}}.$$  \hspace{1cm} (7)

(If $U > D$, we must have $U_{x,t} > e^r > D_{x,t}$, for otherwise an arbitrage arises.) Thus, the model is completely determined by the interest rate $r$ and the jump factors $U_{x,t}$ and $D_{x,t}$.

If the jump factors $U_{x,t}$ and $D_{x,t}$ do not depend on $x$ or $t$, then by virtue of (7), neither does $p_{x,t}$, so the logarithm of the price process $X$ is, under $Q$, the sum of independent and identically distributed Bernoulli trials. This is a popular special case in Shreve’s book and in financial industry practice during the 1970s and 1980s. After the Black–Scholes formula appeared, Bill Sharpe (better known for his Nobel-prize-winning Capital Asset Pricing Model) noticed that the continuous-time option pricing model has this simple “binomial” analogue.

Cox and Ross in [5] developed the binomial model in a general multi-period setting and showed how the Black–Scholes formula arises as the limit of the option price in the binomial model as the number of trading periods per unit of time goes to infinity, and the jump factors $U$ and $D$ are normalized to give a fixed return variance per unit of time. This limit can be viewed as a consequence of Donsker’s Theorem (recorded, for example, in [9]), by which the distribution of a Brownian motion such as $\log X$ can be approximated as a normalized sum of Bernoulli trials.

After laying out the basics, Shreve provides two extensive applications of the binomial model: American options and interest-rate derivatives. The nomenclature “American,” due to Paul Samuelson, refers to the contractual freedom offered to the investor to collect the derivative payoff $g(X_\tau)$ at any stopping time $\tau$ before the
expiration date $T$. It can be shown, using the theory of optimal stopping, that the price of the American security must, in the absence of arbitrage, be what Merton in [17] called the “rational price,” given by

$$\sup_{\tau \leq T} E^Q \left( e^{-r\tau} g(X_\tau) \right).$$

Curiously, Shreve does not show precisely why this pricing formula is required by the absence of arbitrage, but he does show something tantamount to this. Using a novel pedagogical approach based on the reflection principle, Shreve goes on to treat “perpetual” American options, those with no upper bound on the exercise stopping time $\tau$.

Elementary by design, Shreve’s first volume is a lovely introduction to financial modeling, and could be taught to masters or even undergraduate students as a short, say month-long, course.

**Volume II. Continuous-Time Models**

Shreve’s much longer second volume, *Continuous-Time Models*, is a self-contained introduction to stochastic calculus and its applications to financial modeling. In my view, there is no better introductory treatment of the topic. Shreve’s facility with the subject at a deep mathematical level, combined with his pedagogical talent, allowed him to make this treatment both rigorous and easily accessible to good masters-level students. Apparently no effort has been spared to get it right and to make it understandable. Shreve uses measure theory to explain and reinforce concepts, as opposed to some introductory treatments that attempt to avoid measure theory and invariably get into difficulties in providing intuitive, not to mention correct, results.

Volume II begins with a general measure-theoretic introduction to probability theory, a routine two-chapter summary. Chapter 3 is devoted to the definition and properties of Brownian motion. Getting to the heart of the matter, Chapter 4 introduces stochastic calculus and, in a lovely stroke of exposition, uses the Black–Scholes option pricing model to explain how the theory works.

Chapter 4 ends with the Brownian bridge, a stochastic process whose paths between 0 and $t$ have the probability distribution of a Brownian motion conditional on a given ending point at time $t$. Shreve provides two alternative constructions of the Brownian bridge: as a particular Gaussian process and as the solution of a stochastic differential equation. I had never seen the Brownian bridge so nicely explained from several viewpoints. Although the Brownian bridge does have some applications in finance, none are developed in Shreve’s book. Perhaps this section was originally written for some other book?

Chapter 5 completes the underpinnings of the basic theory by covering the Harrison–Kreps theory of risk-neutral probabilities (what they call “equivalent martingale measures”), allowing an often painless route to the calculation of derivative prices.

For example, consider again the price of a derivative paying $g(X_T)$ at $T$, where the underlying asset price process $X$ is the Ito process [2]. An equivalent martingale measure is a probability measure $Q$ that is equivalent to the reference measure $P$ (in the sense that $P$ and $Q$ assign zero probability to the same set of events), and under which $e^{-rt}X_t$ is a martingale. Indeed, a defining property of $Q$ is that any asset price, discounted by $e^{-rt}$, is a $Q$-martingale. So, the initial derivative price
is $E^Q(e^{-rT}g(X_T))$. For a well-defined derivative price, it is now only a question of ensuring that such a measure $Q$ exists and deducing the probability distribution of $X_T$ under $Q$. In Shreve’s Brownian setting, following the Harrison–Kreps approach, this is accomplished with Girsanov’s Theorem\(^\text{10}\).

Chapter 6 connects the foundational theory with Markov processes and partial differential equations. If $X$ is a Markov process under $Q$ solving a stochastic differential equation of the type \(^\text{11}\), then the function $f$ defined by $f(X_t, t) = E^Q[e^{-rT}g(X_T)|X_t]$ can be viewed as the solution of a particular (backward Kolmogorov) parabolic partial differential equation. This PDE approach is now little used in mainstream financial engineering. Typically, derivative prices, as risk-neutral expectations, are computed by Monte Carlo simulation. PDE methods are used principally for initial conceptual model designs. This is in part because of the restrictiveness of the Markov setting, and partly because essentially all models in use at a bank or trading firm must be integrated for risk management and other purposes within a Monte Carlo setting.

Most of the remainder of the book is taken up with a range of applications: Exotic Options (Chapter 7), American Derivative Securities (Chapter 8), Change of Numéraire (Chapter 9), and Term-Structure Models (Chapter 10). To this point, aside from a few gems like the Brownian bridge section, the topic coverage and overall structure are quite conventional; one would expect as much in a textbook designed for future practitioners. What marks the book as unusual in comparison with many other available texts covering this topic is the care applied to detail, rigor, and exposition.

The final chapter extends from the Brownian model to allow processes with jumps, for example Poisson processes. This extension has become crucial with the advent of a deep market for credit derivatives, which allows investors to hedge against, or speculate on, financial default. As default frequently occurs with little or no warning, it would be difficult to parsimoniously model the price of a credit derivative with the continuous paths of an Itô process. Unfortunately for readers, although Shreve has worked on default risk in some of his own research, his textbook does not deal with the subject. (Perhaps he can be convinced to extend his coverage to credit derivatives in a second edition!)

Each chapter of Stochastic Calculus for Finance ends with notes to the literature and instructive exercises. I am pleased that Shreve has not succumbed to the usual plea that a textbook should include solutions to the exercises. Rather, he nicely calibrates the difficulty of his exercises and lards them with hints, so that the student can learn by solving problems. In my experience, this is likely to reinforce the concepts and techniques more deeply than merely learning by reading.

For years, I have profited from reading Steven Shreve’s papers and books and have been captivated by his remarkably lucid research presentations. It is easy to see why he is widely admired for bringing clarity to the many subject areas that he has touched over his career, ranging well beyond financial modeling. Not only do I recommend these two volumes as the place to start one’s education in asset

\(^{10}\)The Radon–Nikodým derivative $\frac{dQ}{dP}$ has a density process $\xi$, defined by $\xi_t = E(\frac{dQ}{dP} | \mathcal{F}_t)$. The martingale representation theorem, also found in Shreve’s Chapter 5, implies that $d\xi_t = -\xi_t \eta_t dB_t$, for a process $\eta$ that plays a special role in pricing models. Girsanov’s Theorem, provided by Shreve in Section 5.4, is that the process $B^*$ defined by $B^*_t = B_t + \int_0^t \eta_s ds$, is a standard Brownian motion under $Q$. 
pricing, I am convinced that even more advanced scholars would profit from pure
enjoyment of the exposition.

REFERENCES


**DARRELL DUFFIE**

GRADUATE SCHOOL OF BUSINESS, STANFORD UNIVERSITY, STANFORD, CALIFORNIA 94305-5015