CHAPTER 1
Overview

- Financial models in light of the great financial crisis.
- The difficulties of option valuation.
- An introduction to the volatility smile.
- Financial science and financial engineering.
- The purpose and use of models.

INTRODUCTION

Our primary aim in this book is to provide the reader with an accessible, not-too-sophisticated introduction to models of the volatility smile. Prior to the 1987 global stock market crash, the Black-Scholes-Merton (BSM) option valuation model seemed to describe option markets reasonably well. After the crash, and ever since, equity index option markets have displayed a volatility smile, an anomaly in blatant disagreement with the BSM model. Since then, quants around the world have labored to extend the model to accommodate this anomaly. Our main focus in this book will be the theory of option valuation, the study of the BSM model and its limitations, and a detailed introduction to the extensions of the BSM model that attempt to rectify its problems. Most of the book is devoted to these topics.

A secondary motivation for writing this book originates in the great financial crisis of 2007–2008, which began with the collapse of the mortgage collateralized debt obligation (CDO) market, whose structured credit products were valued using financial engineering techniques. When the crisis began, some pundits blamed the practice of financial engineering for the mortgage market's meltdown. Paul Volcker, whose grandson was a financial engineer, wrote the following paragraph as part of an otherwise sensible speech he gave in 2009:

> *A year or so ago, my daughter had seen . . . some disparaging remarks I had made about financial engineering. She sent it to my grandson,*

JWBT1927-c01 JWBT1927-Derman July 7, 2016 17:0 Printer Name: Trim: 6in × 9in

who normally didn't communicate with me very much. He sent me an email, "Grandpa, don't blame it on us! We were just following the orders we were getting from our bosses." The only thing I could do was send him back an email, "I will not accept the Nuremberg excuse."

Comparing financial modelers to Nazi war criminals seems extreme, and indeed, since then, opinions about modelers' responsibility for the financial meltdown have become more nuanced. Spain and Ireland developed housing market bubbles that, unlike those in the United States, were not inflated by complex financially engineered products. Paul Krugman has suggested that the root cause of the crisis lay in the West's rapid withdrawal of capital from Asia after the currency crisis of 1998, leading Asian countries thereafter to concentrate on exporting, saving, and hoarding, which led them to provide cheap credit that fueled speculation. Other competing explanations abound. As with all complex human events, it's impossible to pinpoint a single cause.

Nevertheless, models did play a part in the development of the crisis. In the face of very low safe yields, badly engineered financial models were indeed used to tempt investors—at times misleadingly and deceptively—into buying structured CDOs that promised optimistically high yields. Though our expertise lies in models for option valuation rather than mortgage securities, we also wanted to write a book that illustrates how to be sensible about model building.

THE BLACK-SCHOLES-MERTON MODEL AND ITS DISCONTENTS

Stephen Ross of MIT, one of the inventors of the binomial option valuation model and the theory of risk-neutral valuation, once wrote: "When judged by its ability to explain the empirical data, option pricing theory is the most successful theory not only in finance, but in all of economics" (Ross 1987). But even this most successful of models is far from being perfect.

Finance academics tend to think of option valuation as a solved problem, of little current interest. But readers of this book who end up working as practitioners—on options trading desks in equities, fixed income, currencies, or commodities, as risk managers or controllers or model auditors—will find that the valuation of options isn't really a solved problem at all. Financial markets disrespect the traditional BSM formula even while they employ its flawed language to communicate with each other. Practitioners and traders who are responsible for coming up with the prices at which they are willing to trade derivative securities, especially exotic illiquid derivatives, grapple with appropriate valuation every day. They have to figure out how to amend the BSM model to cope with an actual market that violates its assumptions, and they have to keep finding new ways of doing so as the market modifies its behavior based on its experiences.

In this book we're going to focus on the BSM model and its discontents. In one sense the BSM model is a miracle: It lets you value, in a totally rational way, securities that before its existence had no plausible or defensible theoretical value at all. In the Platonic world of BSM—a world with normally distributed returns, geometric Brownian motion for stock prices, unlimited liquidity, continuous hedging, and no transaction costs—their model provides a method of dynamically synthesizing an option. It's a masterpiece of engineering in an imaginary world that doesn't quite exist, because markets don't obey all of its assumptions. It's a miracle, but it's only a model, and not reality.

Some of the BSM assumptions are violated in minor ways, some more dramatically. The assumption that you can hedge continuously, at zero transaction cost, is an approximation we can adjust for, as we will illustrate in later chapters. Skilled traders and quants do this with a mix of estimation and intuition every day. You can, for example, heuristically allow for transaction costs by adding some dollars to your option price, or some volatility points to the BSM formula. In that sense the model is robust—you can perturb it from its Platonic view of the world to approximate the messiness of actual markets.

Other BSM assumptions are violated in more significant ways. For example, stock prices don't actually follow geometric Brownian motion. They can jump, their distributions have fat tails, and their volatility varies unpredictably. Adjusting for these more significant violations is not always easy. We will tackle many of these difficulties in this book.

In the end, the BSM model sounds so rational, and has such a strong grip on everyone's imagination, that even people who don't believe in its assumptions nevertheless use it to quote prices at which they are willing to trade.

A QUICK LOOK AT THE IMPLIED VOLATILITY SMILE

The BSM model assumes that a stock's future return volatility is constant, independent of the strike and time to expiration of any option on that stock. Were the model correct, a plot of the implied BSM volatilities for options with the same expiration over a range of strikes would be a flat line. Figure 1.1 shows what three-month equity index implied volatilities looked like before the Black Monday stock market crash of 1987.

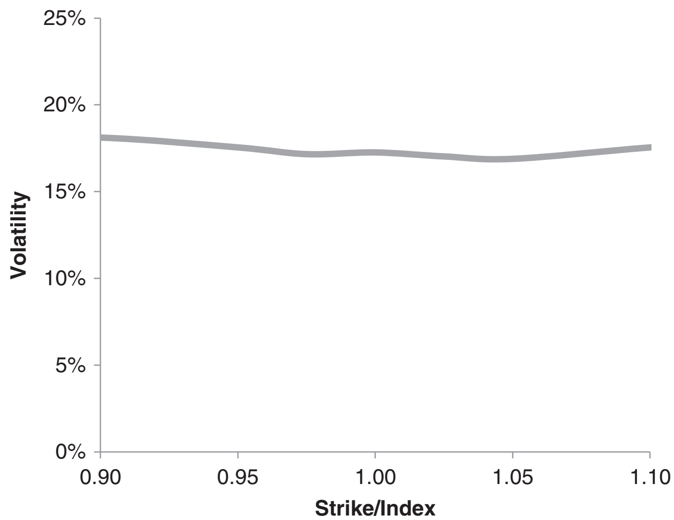

FIGURE 1.1 Representative S&P 500 Implied Volatilities prior to 1987

Prior to the crash, therefore, the BSM model seemed to describe the option market rather well, at least with respect to variation in strikes. Figure 1.2 shows typical three-month implied volatilities after the crash of 1987. Even though all the options used to generate the smile were written on the same underlier, each option had a different implied volatility. This is inconsistent with the BSM model, which assumes that implied volatility is a forecast of actual volatility, for which there can be only one value. You can think of options as metaphorical photographs of the stock's future volatility, taken from different angles or elevations. While photographs of a building taken from different points might look different, the actual size of a building remains the same. In a similar way, if the BSM model were truly reliable, the implied volatility of the stock would be the same, no matter which option you chose to view it with. The option price is *derived* from the stock price, but the stock's volatility should not depend on the option.

Though the smile appeared most dramatically in equity index option markets after the 1987 crash, there had always been a slight smile in currency option markets, a smile in the literal sense that the implied volatilities as a function of strike resembled one: $\smile$. As depicted in Figure 1.2, the equity "smile" is really more a skew or a smirk, but practitioners have persisted in using the word *smile* to describe the relationship between implied volatilities and strikes, irrespective of the actual shape. The smile's appearance after the 1987 crash was clearly connected with the visceral shock upon discovering,

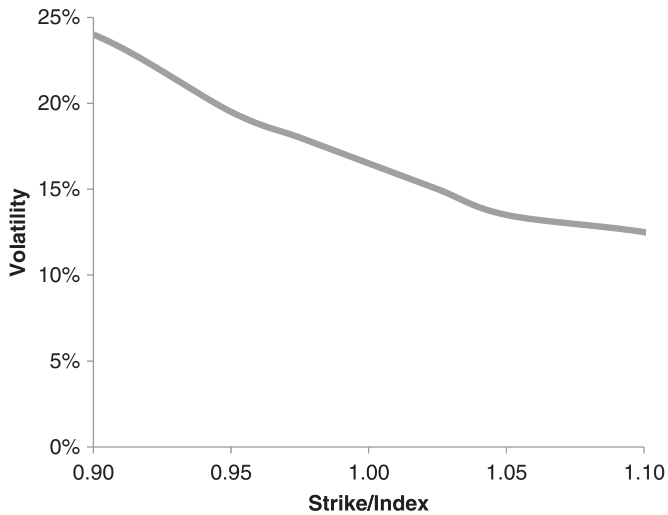

FIGURE 1.2 Representative S&P 500 Implied Volatilities after 1987

for the first time since 1929, that a giant market could suddenly drop by 20% or more in a day. Market participants immediately drew the conclusion that an investor should pay more for low-strike puts than for high-strike calls.

Since the crash of 1987, the volatility smile has spread to most other option markets (currencies, fixed income, commodities, etc.), but in each market it has taken its own characteristic form and shape. Traders and quants in every product area have had to model the smile in their own market. At many firms, not only does each front-office trading desk have its own particular smile models, but the firm-wide risk management group is likely to have its own models as well. The modeling of the volatility smile is likely one of the largest sources of model risk within finance.

NO-NONSENSE FINANCIAL MODELING

During the past 20 years there has been a tendency for quantitative finance and asset pricing to become increasingly formal and axiomatic. Many textbooks postulate mathematical axioms for finance and then derive the consequences. In this book, though, we're studying financial engineering, not mathematical finance. The ideas and the models are at least as important as the mathematics. The more math you know, the better, but math is the syntax, not the semantics. Paul Dirac, the discoverer of the Dirac equation who first predicted the existence of antiparticles, had a good point when he said:

> *I am not interested in proofs, but only in what nature does.*
>
> —Paul Dirac

About Theorems and Laws

Mathematics requires axioms and postulates, from which mathematicians then derive the logical consequences. In geometry, for example, Euclid's axioms are meant to describe self-evident relationships of parts of things to the whole, and his postulates further describe supposedly self-evident properties of points and lines. One Euclidean axiom is that things that are equal to the same thing are equal to each other. One Euclidean postulate, for example, is that it is always possible to draw a straight line between any two points.

Euclid's points and lines are abstracted from those of nature. When you get familiar enough with the abstractions, they seem almost tangible. Even more esoteric abstractions—infinite-dimensional Hilbert spaces that form the mathematical basis of quantum mechanics, for example—seem real and visualizable to mathematicians. Nevertheless, the theorems of mathematics are relations between abstractions, not between the realities that inspired them.

Science, in contradistinction to mathematics, formulates laws. Laws are about observable behavior. They describe the way the universe works. Newton's laws allow us to guide rockets to the moon. Maxwell's equations enable the construction of radios and TV sets. The laws of thermodynamics make possible the construction of combustion engines that convert heat into mechanical energy.

Finance is concerned with the relations between the values of securities and their risk, and with the behavior of those values. It aspires to be a practical field, like physics or chemistry or electrical engineering. As John Maynard Keynes once remarked about economics, "If economists could manage to get themselves thought of as humble, competent people on a level with dentists, that would be splendid." Dentists rely on science, engineering, empirical knowledge, and heuristics, and there are no theorems in dentistry. Similarly, one would hope that finance would be concerned with laws rather than theorems, with behavior rather than assumptions. One doesn't seriously describe the behavior of a market with theorems.

How then should we think about the foundations of finance and financial engineering?

On Financial Engineering

Engineering is concerned with building machines or devices. A device is a little part of the universe, more or less isolated, that, starting from the constructed initial conditions, obeys the laws of its field and, while doing so, performs something we regard as useful.

Let's start by thinking about more familiar types of engineering. Mechanical engineering is concerned with building devices based on the principles of mechanics (i.e., Newton's laws), suitably combined with empirical rules about more complex forces that are too difficult to derive from first principles (friction, for example). Electrical engineering is the study of how to create useful electrical devices based on Maxwell's equations and quantum mechanics. Bioengineering is the art of building prosthetics and biologically active devices based on the principles of biochemistry, physiology, and molecular biology.

Science—mechanics, electrodynamics, molecular biology, and so on—seeks to discover the fundamental principles that describe the world, and is usually reductive. Engineering is about using those principles, constructively, to create functional devices.

What about financial engineering? In a logically consistent world, financial engineering, layered above a solid base of financial science, would be the study of how to create useful financial devices (convertible bonds, warrants, volatility swaps, etc.) that perform in desired ways. This brings us to financial science, the putative study of the fundamental laws of financial objects, be they stocks, interest rates, or whatever else your theory uses as constituents. Here, unfortunately, be dragons.

Financial engineering rests upon the mathematical fields of calculus, probability theory, stochastic processes, simulation, and Brownian motion. These fields can capture some of the essential features of the uncertainty we deal with in markets, but they don't accurately describe the characteristic behavior of financial objects. Markets are plagued with anomalies that violate standard financial theories (or, more accurately, theories are plagued by their inability to systematically account for the actual behavior of markets). For example, the negative return on a single day during the crash of 1987 was so many historical standard deviations away from the mean that it should never have occurred in our lifetime if returns were normally distributed. More recently, JPMorgan called the events of the "London Whale" an eight-standard-deviation event (JPMorgan Chase & Co. 2013). Stock evolution, to take just one of many examples, isn't Brownian.[1] So, while financial engineers are rich in mathematical techniques, we don't have the right laws of science to exploit—not now, and maybe not ever.

Because we don't have the right laws, the axiomatic approach to finance is problematic. Axiomatization is appropriate in a field like geometry, where one can postulate any set of axioms not internally inconsistent, or even in Newtonian mechanics, where there are scientific laws that hold with such great precision that they can be effectively regarded as axioms. But in finance, as all practitioners know, our "axioms" are not nearly as good. As Paul Wilmott wrote, "every financial axiom . . . ever seen is demonstrably wrong. The real question is how wrong . . ." (Wilmott 1998). Teaching by axiomatization is therefore even less appropriate in finance than it is in real science. If finance is about anything, it is about the messy world we inhabit. It's best to learn axioms only after you've acquired intuition.

Mathematics is important, and the more mathematics you know the better off you're going to be. But don't fall too in love with mathematics. The problems of financial modeling are less mathematical than they are conceptual. In this book, we want to first concentrate on understanding concepts and their implementation, and then use mathematics as a tool. We're less interested here in great numerical accuracy or computational efficiency than in making the ideas we're using clear.

We know so little that is absolutely right about the fundamental behavior of assets. Are there really strict laws they satisfy? Are those laws stationary? It's best to assume as little as possible and rely on models as little as possible. And when we do rely on models, simpler is better. With that in mind, we proceed to a brief overview of the principles of financial modeling.

THE PURPOSE OF MODELS

Before examining the notion of modeling, we must distinguish between price and value. Price is simply what you have to pay to acquire a security, or what you get when you sell it; value is what a security is worth (or, more accurately, what you believe it is worth). Not everyone will agree on value. A price is considered fair when it is equal to the value.

But what is the fair value? How do you estimate it? Judging value, in even the simplest way, involves the construction of a model or theory.

[1] See, for example, Mandelbrot (2004) and Gabaix et al. (2003).

A Simple but Prototypical Financial Model

Suppose a financial crisis has just occurred. Wall Street is laying off people, apartments in nearby Battery Park are changing hands daily, but large luxurious apartments are still illiquid. How would you estimate the value of a seven-room apartment on Park Avenue, whose price is unknown, if someone tells you the price of a two-room apartment in Battery Park? This would be a reasonable model: First, figure out the price per square foot of the Battery Park apartment; second, multiply by the square footage of the Park Avenue apartment; third, make some adjustments for location, views, light, staff, facilities, and so forth.

For example, suppose the two-room Battery Park apartment cost $1.5 million and was 1,000 square feet in size. That comes to $1,500 per square foot. Now suppose the seven-room Park Avenue apartment occupies 5,000 square feet. According to our model, the price of the Park Avenue apartment should be roughly $7.5 million. But Park Avenue is a very desirable location, and so we understand that there is about a 33% premium over Battery Park, which raises our estimate to $10 million. Furthermore, large apartments are scarce and carry their own premium, raising our estimate further to $13 million. Suppose further that the Park Avenue apartment is on a high floor with great views and its own elevator, so we bump up our estimate to $15 million. On the other hand, say the same Park Avenue apartment is being sold by the family of a recently deceased parent who hasn't renovated it for 40 years. It will need a lot of work, which causes us to lower our estimate to $12 million.

Our model's one initial parameter is the implied price per square foot. You calibrate the model to Battery Park and then use it to estimate the value of the Park Avenue apartment. The price per square foot is truly *implied* from the price; $1,500 is not the price of one square foot of the apartment, because there are other variables—views, quality of construction, neighborhood—that are subsumed into that one number.

With financial securities, too, as in the apartment example, models are used to interpolate or extrapolate from prices you know to values you don't—in our example, from Battery Park prices to Park Avenue prices. Models are mostly used to value relatively illiquid securities based on the known prices of more liquid securities. This is true both for structural option models and purely statistical arbitrage models. In that sense, and unlike models in physics, models in finance don't really predict the future. Whereas Newton's laws tell you where a rocket will go in the future given its initial position and velocity, a financial model tells you how to compare different prices in the present. The BSM model tells you how to go from the current price of a stock and a riskless bond to the current value of an option, which it views as a mixture of the stock and the bond, by means of a very sophisticated and rational kind of interpolation. Once you calibrate the model to a stock's implied volatility for one option whose price you know, it tells you how to interpolate to the value of options with different strikes. The volatility in the BSM model, like the price per square foot in the apartment pricing model, is implied, because all sorts of other variables—trading costs, hedging errors, and the cost of doing business, for example—are subsumed into that one number. The way property markets use implied price per square foot illustrates the general way in which most financial models operate.

Additional Advantages of Using a Model

Models do more than just extrapolate from liquid prices to illiquid values.

Ranking Securities A security's price doesn't tell you whether it's worth buying. If its value is more than its price, it may be. But sometimes, faced with an array of similar securities, you want to know which security is the best deal. Models are often used by investors or salespeople to rank securities in attractiveness. Implied price per square foot, for example, can be used to rank and compare similar, but not identical, apartments. Suppose, to return to our apartment example, that we are interested in purchasing a new apartment in the Financial District. The apartment lists at $3 million, but is 1,500 square feet, or $2,000 per square foot, appreciably higher than the $1,500 per square foot for the Battery Park apartment. What justifies the difference? Perhaps the Financial District apartment has better features. We might even go one level deeper and start to build a comparative model for the features themselves, or for both the features and the square footage, to see if the features are fairly priced.

Implied price per square foot provides a simple, one-dimensional scale on which to *begin* ranking apartments by value. The single number given by implied price per square foot does not truly reflect the value of the apartment; it provides a starting point, after which other factors must be taken into account. Similarly, yield to maturity for bonds allows us to compare the values of many similar but not identical bonds, each with a different coupon, maturity, and/or probability of default, by mapping their yields onto a linear scale from high (attractive) to low (less so). We can do the same thing with price-earnings (P/E) ratio for stocks or with option-adjusted spread (OAS) for mortgages or callable bonds. All these metrics project a multidimensional universe of securities onto a one-dimensional ruler. The implied volatility associated with options obtained by filtering prices through the BSM model provides a similar way to collapse instruments with many qualities (strike, expiration, underlier, etc.) onto a single value scale.

Quantifying Intuition Models provide an entry point for intuition, which the model then quantifies. A model transforms linear quantities, which you can have intuition about, into nonlinear dollar values. Our apartment model transforms price per square foot into the estimated dollar value of the apartment. It is easier to develop intuition about variation of price per square foot than it is about an apartment's dollar value.

In physics, as we stressed, a theory predicts the future. In finance, a model translates intuition into current dollar values. As a further example, equity analysts have an intuitive sense, based on experience, about what constitutes a reasonable P/E ratio. Developing intuition about yield to maturity, option-adjusted spread, default probability, or return volatility may be harder than thinking about price per square foot. Nevertheless, all of these parameters are directly related to value and easier to judge than dollar value itself. They are intuitively graspable, and the more experienced you become, the richer your intuition will be. Models advance by leapfrogging from a simple, intuitive mental concept (e.g., volatility) to the mathematics that describes it (geometric Brownian motion and the BSM model), to a richer concept (the volatility smile), to experience-based intuition (the variation in the shape of the smile), and, finally, to a model (a stochastic volatility model, for example) that incorporates an extension of the concept.

Styles of Modeling: What Works and What Doesn't

The apartment model is an example of *relative valuation*. With relative valuation, given one set of prices, one can use the model to determine the value of some other security. One could also hope to develop models that value securities *absolutely* rather than relatively. In physics, Newton's laws are absolute laws. They specify a law of motion, $F = ma$, and a particular force law, the gravitational inverse-square law of attraction, which allow one to calculate any planetary trajectory. Geometric Brownian motion and other more elaborate hypotheses for the movement of primitive assets (stocks, commodities, etc.) look like models of absolute valuation, but in fact they are based on analogy between asset prices and physical diffusion phenomena. They aren't nearly as accurate as physics theories or models. Whereas physics theories often describe the actual world—so much so that one is tempted to ignore the gap between the equations and the phenomena—financial models describe an imaginary world whose distance from the world we live in is significant.

Because absolute valuation doesn't work too well in finance, in this book we're going to concentrate predominantly on methods of relative valuation. Relative valuation is less ambitious, and that's good. Relative valuation is especially well suited to valuing derivative securities.

Why do practitioners concentrate on relative valuation for derivatives valuation? Because derivatives are a lot like molecules made out of simpler atoms, and so we're dealing with their behavior relative to their constituents. The great insight of the BSM model is that derivatives can be manufactured out of stocks and bonds. Options trading desks can then regard themselves as manufacturers. They acquire simple ingredients—stocks and Treasury bonds, for example—and manufacture options out of them. The more sophisticated trading desks acquire relatively simple options and construct exotic ones out of them. Some even do the reverse: acquiring exotic options and deconstructing them into simpler parts to be sold. In all cases, relative value is important, because the desks aim to make a profit based on the difference in price of inputs and outputs—the difference in what it costs you to buy the ingredients and the price at which you can sell the finished product.

Relative value modeling is nothing but a more sophisticated version of the fruit salad problem: Given the price of apples, oranges, and pears, what should you charge for fruit salad? Or the inverse problem: Given the price of fruit salad, apples, and oranges, what is the implied price of pears? You can think of most option valuation models as trying to answer the options' analogue of this question.

In this book we'll mostly take the viewpoint of a trading desk or a market maker who buys what others want to sell and sells what others want to buy, willing to go either way, always seeking to make a fairly safe profit by creating what its clients want out of the raw materials it acquires, or decomposing what its clients sell into raw materials it can itself sell or reuse. For trading desks that think like that, valuation is always a relative concept.