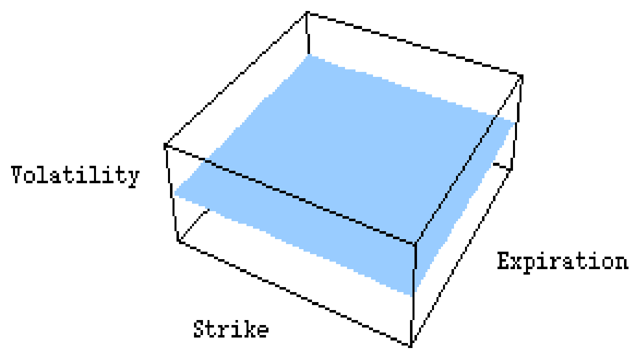
Lecture 1: Introduction to the Smile; The Principles of Valuation

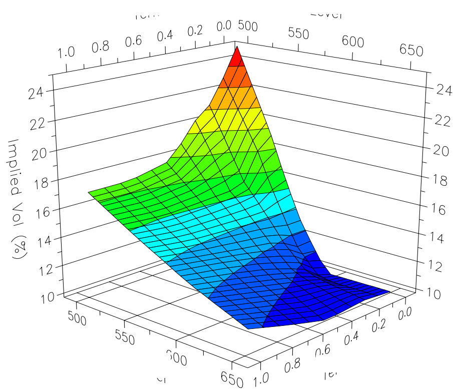
1.1 Introduction

According to classic theory, the Black-Scholes implied volatility of an option should be independent of its strike and expiration. Plotted as a surface, it should be flat, as shown at right.

The volatility surface according to Black-Scholes



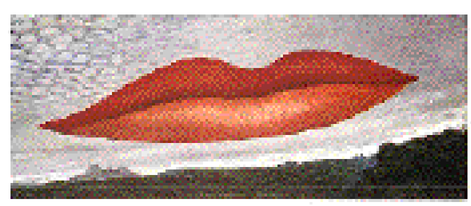
The volatility surface according to S&P options markets



Prior to the stock market crash of October 1987, the volatility surface of index options was indeed fairly flat.

Since the crash, the volatility surface of index options has become skewed. Referred to as the volatility smile, the surface changes over time. Its level at any instant is a varying function of strike and expiration, as shown at left.

The smile phenomenon has spread to stock options, interest-rate options, currency options, and almost ever other volatility market. Since the Black-Scholes model cannot account for the smile, trading desks have begun to use more complex models to value and hedge their options.



After 15 years, there is still no overwhelming consensus as to the correct model. Each market has its own favorite (or two). Despite initial optimism about finding **the** model to replace Black-Scholes, we are still in many ways searching in the dark.

This first class is very different in style from successive classes; it sets up the general methodology of financial modeling, mostly qualitatively. Subsequent classes will be predominantly quantitative (and less long-winded …).

- Aim of the Course

This isn't a course about mathematics, calculus, differential equations or stochastic calculus, though it does use all of them. Much of the time the approach is going to be mathematical, but not extremely rigorous. I want to develop intuition about models, not just methods of solution. No assumptions behind financial models are genuinely true, and no financial models are really correct, so it's very important to understand what you're doing and why.

This is a course about several themes:

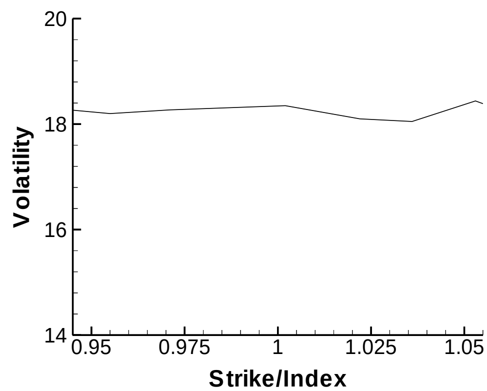
1. Understanding the practical use of the Black-Scholes-Merton model. There's more to it than just the equation and its solution.
2. The theoretical and practical limitations of the model.
3. The extensions of the model to accommodate/explain the volatility smile.
4. Understanding the consequences of these extensions. It's easy to make up new models of increasingly greater complexity, but we want to understand whether they can describe the world and what they lead to. First we'll begin with a brief look at the implied volatility smile for equity index options, a phenomenon inconsistent with the Black-Scholes model.

This course is eventually concerned with the problems in Black-Scholes, and then how to extend or replace the Black-Scholes model in order to explain or account for the smile. Therefore in this lecture we'll spend some time discussing the nature of financial modeling more generally, as a preamble to modeling the smile.

1.2 A Quick Look at the Implied Volatility Smile

The Black-Scholes model assumes that a stock's return volatility is a constant, independent of strike and time to expiration. As a consequence, if the model is correct, then when you plot the implied Black-Scholes volatilities of options of a fixed expiration over a range of strikes, you should see a flat line. This is roughly what implied volatilities looked like before the stock market crash of 1987.

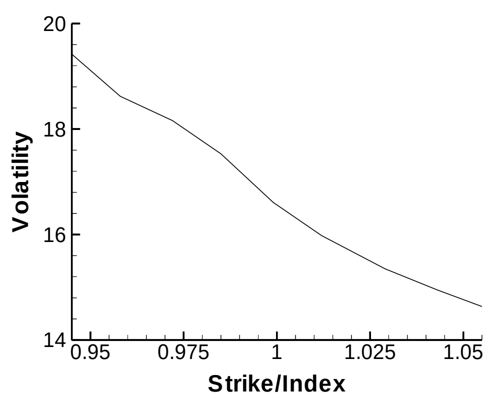
FIGURE 1.1. Representative S&P 500 implied volatilities prior to 1987. Data taken from M. Rubinstein, "Implied Binomial Trees" J. of Finance, 69 (1994) pp. 771-818



The Black-Scholes model seemed to describe the options market reasonably, at least with respect to variation with strike. Furthermore, by extending the model in an almost trivial way and allowing a stock's future volatility to vary with time, you can obtain similar shapes with different levels of volatility at different times.

Here is what volatilities looked like after the crash of 1987.

FIGURE 1.2. Representative S&P 500 implied volatilities after 1987.



The Black-Scholes model has no simple way to obtain different implied volatilities for different strikes. An option is like a photograph of the underlying stock. Each option looks at the stock from a different angle or elevation. The size of a building cannot depend on the angle from which you photograph it. Similarly, the volatility of a *stock* itself cannot depend upon the option with which you choose to view it. Therefore, if the Black-Scholes choice of volatility as an appropriate variable to quantify option value, then the stock's volatility should be independent of *option's* strike or time to expiration, because the option is a derivative that "sits above" the stock.

1.2.1 Development of the Smile

There was always a bit of a smile in currency options markets, literally a smile in the sense that the implied volatilities as a function of strike resembled a smile:  . As indicated above, the equity "smile" is really more of a skew or a smirk, but the world *smile* has come to persist in describing the surface of implied volatilities and the phenomenon itself. The smile's appearance after the 1987 crash and was clearly connected in some way with the visceral shock of discovering, for the first time since 1929, that a giant market could drop by 20% or more in a day or two. Clearly, as a consequence, an investor who doesn't hedge should pay more for low-strike puts than for high-strike or at-the-money call. For a trader who hedges, it's not so clear.

Since then the volatility smile has spread to most other options markets – currencies, fixed income, commodities, etc. – but in each market it has taken its own idiosyncratic form and shape. Slowly, and then more rapidly, traders and quants in every product area have had to model the smile. At many firms, not only does each front-office trading desk have quantitative strategists working on their particular smile model, but the Firmwide Risk Management group likely has a group of quants to value independently many of the structured illiquid deals that depend on the details of the volatility smile. I would think it's safe to say that there is no area where model risk is more of an issue than in the modeling of the volatility smile.

1.2.2 A Brief Look at the Equity Index Smile

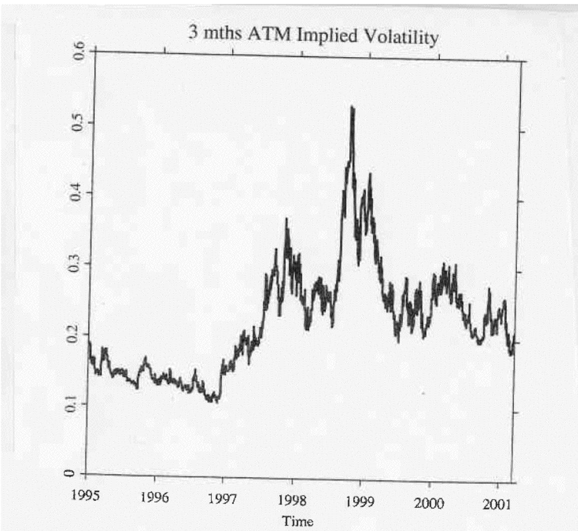


Fig. 2.13. The three-months ATM IV levels of DAX index options

short-term implieds
move more
than long-term

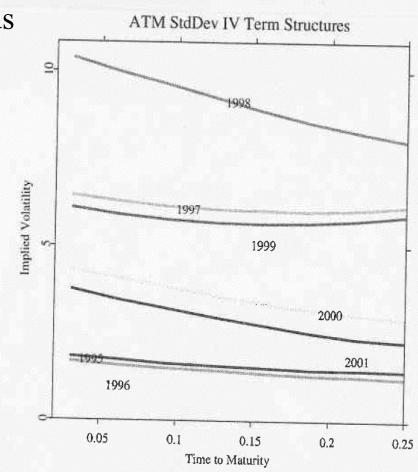


Fig. 2.12. ATM standard deviation of IV term structures between 1995 and 2001, computed from smoothed surfaces

negative correlation
during crisis

From Fengler's book

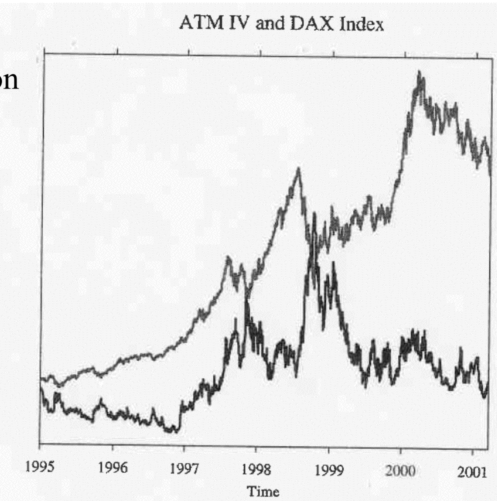
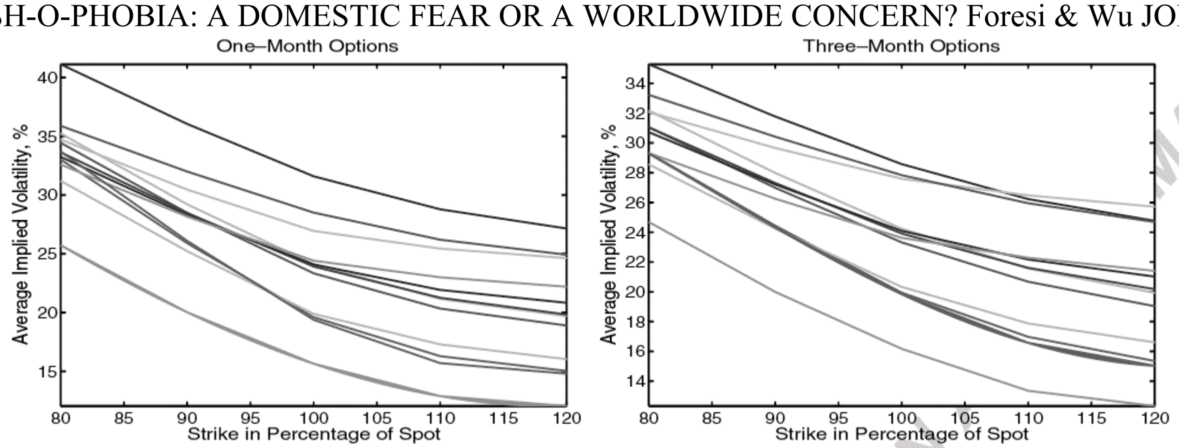


Fig. 2.14. German DAX $\times 10^{-4}$ (*upper line*) and three-months ATM IV levels (*lower line*), also given in Fig. 2.12

Implied Volatility Smirk on Major Equity Indexes

CRASH-O-PHOBIA: A DOMESTIC FEAR OR A WORLDWIDE CONCERN? Foresi & Wu JOD Winter 05



Some characteristics of the equity implied volatility smile

1. Volatilities are steepest for small expirations as a function of strike, shallower for longer expirations.

2. The minimum volatility as a function of strike occurs near atm strikes o strikes corresponding to slightly otm call options.

3. Low strike volatilities are usually higher than high-strike volatilities, but high strike volatilities can also

4. The term structure is usually increasing but can change depending on views of the future. After large sudden market declines, the implied volatility out-of-the-money calls may be greater than for atm calls, reflecting an expectation that the market may rebound.

4. The volatility of implied volatility is greatest for short maturities, as with Treasury rates.

5. There is a negative correlation between changes in implied atm volatility and changes in the underlying asset itself. [Fengler: $\rho = -0.32$ for the DAX in the late 90s, for three-month expirations.]

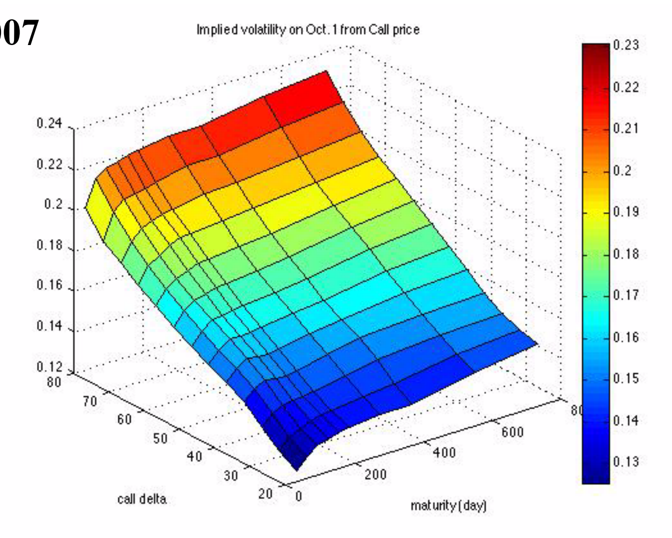
6. Implied volatility appears to be mean reverting with a life of about 60 days.

7. Implied volatility tends to rise fast and decline slowly.

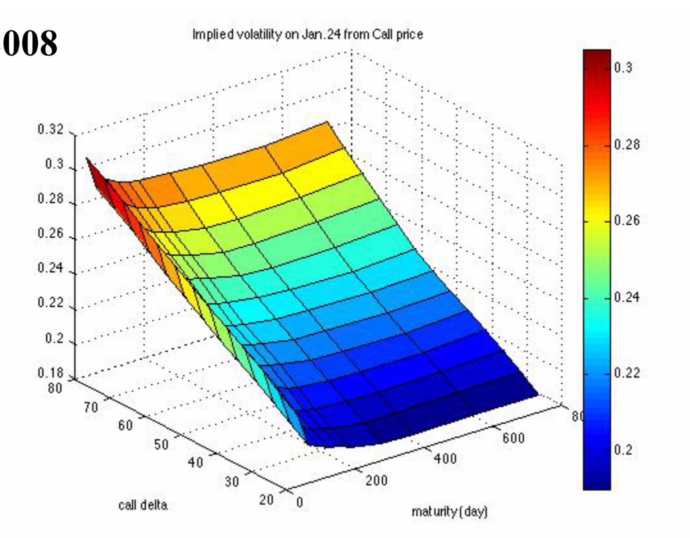
7. Shocks across the surface are highly correlated. There are a small number of principal components or driving factors. We'll study these effects more closely later in the course.

1.2.3 Recent S&P Smiles

Oct 1 2007



Jan 24 2008



1.3 The Principles of Financial Valuation

There has been a tendency for quantitative finance and asset pricing to become more and more formal and axiomatic. Many textbooks write down mathematical axioms for finance and then derive the consequences. But here we're studying quantitative finance or financial engineering, not mathematical finance. The ideas and the models are at least as important as the mathematics. This isn't Euclid, it's finance and markets to which we're *applying* math. The more math you know, the better, but math is the tool, the syntax, not the semantics…

So, what is financial engineering? 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 (friction, for example) that are too difficult to derive from first principles. Electrical engineering is the study of how to create useful electrical devices based on Maxwell's equations. Bio-engineering is the art of building prosthetics and other biologically active devices based on the principles of biochemistry, physiology and molecular biology.

Science – mechanics, electrodynamics, molecular biology, etc. – seeks to discover the fundamental principles that describe the world, and is usually reductive. Engineering is about using those principles, constructively, for a purpose.

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 functional financial devices – convertible bonds, warrants, default swaps, etc. – that perform in desired ways, not just at expiration, but throughout their lifetime. Which 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 atomic constituents. Here, unfortunately, lie dragons.

Financial engineering rests upon the mathematical fields of calculus, probability theory, stochastic processes, simulation, Brownian motion. But while these self-consistent fields can be axiomatized and can capture some of the essential features of the uncertainty we deal with in markets, how accurately do they describe the characteristic behavior of financial objects? Markets are plagued with anomalies that violate standard theories (or, more accurately, theories are plagued by their inability to systematically account for true asset distributions). For example, the magnitude of the negative return on a single day during the crash of 1987 is so many standard deviations away from a normal distribution with the market's historical standard deviation to have had a finite chance of occurring in our lifetime. Stock evolution, to take just one of many examples, isn't Brownian (see the work of Eugene Stanley and his collaborators on the scaling of stock price distributions). So, while financial engineers are rich in techniques we don't (yet?, ever?) have the right laws of science to exploit.

Therefore, I don't like the axiomatic approach to finance very much. Axiomatization is very useful in a field like geometry, where the axioms hold true to a high degree of accuracy, or even in Newtonian mechanics, where there are scientific laws that hold with great precision. But in finance, as all practitioners know and as Paul Wilmott wrote, "every financial axiom... ever seen is demonstrably wrong. The real question is *how* wrong[1]..." Teaching by axiomatization is therefore even less appropriate in finance than it is in real science. If economics is about anything it is about the messy world we inhabit. It's best to learn axioms after you've learned intuition.

Mathematics is important, and the more mathematics you know the better off you're going to be. But don't fall too much in love with it. In agreement with this, the last sentence of the introduction to John Cochrane's book on asset pricing[2] says that "The hurdles in asset pricing are really conceptual rather than mathematical." In this course I want to first concentrate on the concepts and their understanding and implementation, and then use mathematics as its tool. I'm less interested here in great numerical accuracy than in making ideas clear. And I want be hard-nosed about what works and what doesn't, from a practitioner's point of view.

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 proceed from there.

So, I'd like to start this course with a brief overview of the principles of financial valuation. There aren't too many …

1.3.1 Price & Value

First, you must distinguish between *price* and *value*. Price is simply what you have to pay to acquire a security; value is what it is worth. The price is fair when it is equal to the value.

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

1.3.2 The Purpose of Models

Here's a simple but prototypical financial model. How do you estimate the price of a 7 room apartment on Park Ave if someone tells you the price of a 2 room apartment in Battery Park City? First, you figure out the price per square foot of the BPC apartment. Then you multiply by the square footage of the Park Ave apt. Then you make some corrections for location and Park views and light and facilities. And that's how you'd get the price.

1. Paul Wilmott, in the *prolog* to *Derivatives*, John Wiley & Sons, 1998.
2. John Cochrane, *Asset Pricing*, Princeton U. Press, 2001

The model's critical parameter is the implied price per square foot. You calibrate the model to Battery Park City. Then you use it to interpolate or extrapolate to Park Ave. The price per square foot is implied from the price; it's not the actual price per square foot because there are other variables – exposure, quality of construction, neighborhood – that are subsumed inside that one number.

It's the implied number, like a yield to maturity in bond pricing, that is extracted from a bond market price of one kind and used, roughly, to calculate the fair price of a similar but not identical bond. Yield to maturity takes bonds of all coupons and all maturities and reduces their value to a single number on a one-dimensional scale, and lets you compare them. That's the point of a model – comparison, not prediction. Price/Earnings does the same for stocks. It's not perfect, it's not even strictly accurate, but it provides a helpful way of thinking about things, a helpful way to compare different prices, and to move from one stock with a known P/E to the price of another new one whose earnings are known.

So, in finance, all models are used to interpolate or extrapolate from quantities you know to quantities you don't know – in this case from BPC price to Park Ave price. This is the case with options models and statistical arbitrage models. Models in finance don't really predict the future. Models tell you how to compare different prices in the present. They relate present quantities that are known to present quantities that aren't known. In Black-Scholes the model is used to go from stock price and riskless bond price to option price, by interpolation. And the volatility in the model, like the price per square foot, is implied, because all sorts of other estimates, like trading costs and hedging errors and the cost of doing business, are subsumed inside that one number.

1.3.3 Styles of Modeling

The above example is a model of *relative valuation.* One could also hope to develop models of *absolute valuation*. In physics, Newton's laws are absolute laws. They models specify a law of motion, $F = ma$, and a particular force law, the gravitational inverse-square law, of attraction, and then let you calculate any planetary trajectory. Geometric Brownian motion and other hypotheses for the movement of primitive assets – stocks, commodities, etc. – tend to be models of absolute valuation, but they aren't that accurate.

We're not going to deal much with absolute valuation in this course, because we're dealing with derivatives, and derivatives are like molecules made out of simpler atoms, and so we're dealing with their behavior relative to their constituents. Relative valuation is less ambitious, and that's what we're going to focus on. In relative valuation you try to find the value of an unknown illiquid security relative to the known prices of liquid securities, using as little detailed information as possible. That's what Black-Scholes is all about: it tells you the price of an option in terms of the price of a stock and a bond.

Why do practitioners concentrate on relative valuation for derivatives pricing? Because most of the firms you'll end up working at are not interested in simple speculation or consumption of financial products. You *can* make money by using derivatives like lottery tickets, to speculate. Then you're interested only in the price vs. the probability of payoff. But many financial firms are more interested in more complex bets, by trying to be arbitrageurs. An arbitrageur tries to make money by identifying securities that are mispriced relative to another security, and then exchanging them. More generally, an arbitrageur is a kind of manufacturer that manufactures one security out of others more cheaply than it can be bought. The price of the ingredients and the labor have to be less than the price of the finished product, if not with 100% confidence then at least with a high probability. Therefore, we're interested in relative value.

In this course on derivatives we are going to take the viewpoint of a manufacturer, because the great insight of Black-Scholes is that derivatives can be manufacture our of stock and bonds.

Options trading desks are in many ways like manufacturers, or inverse manufacturers. They acquire simple ingredients – stocks and Treasury bonds, for example – and manufacture options out of them by replication or hedging. The more sophisticated ones do the reverse: they acquire relatively simple options and construct exotic ones out of them, or acquire exotic ones and deconstruct into their simpler parts and sell them off. From their point of view, relative value is very important, because you are buying ingredients and selling a finished product, or vice versa. Similarly, exotic options valuation is often a question of replicating exotics out of more liquid vanilla options. We'll cover this later in the course.

In this course we'll mostly take the viewpoint of a trading desk or a market-maker who buys what people want to sell and sells what people want to buy, willing to go either way, always seeking to make a profit. For desks like that, valuation is always a relative concept.

Much of relative value modeling is then more or less sophisticated versions of fruit salad problems: 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?

1.3.4 The One Commandment of Quantitative Finance

According to a legend, Hillel, a famous sage, was asked to recite the essence of God's laws while standing on one leg. "Do not do unto others as you would not have them do unto you," he is supposed to have said. "All the rest is commentary. Go and learn."

You can summarize the essence of quantitative finance on one leg too: "If you want to know the value of a security, use the price of another security that's as similar to it as possible. All the rest is modeling. Go and build."

The wonderful thing about this law, when compared with almost everything else in economics, is that it dispenses with utility functions, the unobservable hidden variable whose ghostly presence permeates economic theory. But don't think that you can escape all sentiment in financial theory; the models of quantitative finance involve expectations and estimates of future behavior, and those estimates and expectations are *people's* estimates.

Financial economists refer to their essential principle as the *law of one price*, or *the principle of no riskless arbitrage*, which states that

> *Any two securities with identical future payoffs, no matter how the future turns out, should have identical current prices.*

The law of one price is not a law of nature. It's a general reflection on the practices of human beings, who, when they have enough time and enough information, will grab a bargain when they see one. The law usually holds in the long run, in well-oiled markets with enough savvy participants, but there are always short- or even longer-term exceptions that persist.

1.3.5 Valuation by Replication

How do you use the law of one price to determine value? If you want to estimate the value of a *target security*, you must find some other *replicating portfolio*, a collection of more *liquid* securities that, collectively, has the same future payoffs as the target, *no matter how the future turns out*. The target's value is then simply that value of the replicating portfolio.

Where do models enter? It takes a model to show that the target and the replicating portfolio have identical future payoffs under all circumstances. To demonstrate payoff identity, (1) you must specify what you mean by "all circumstances," for each security, and (2) you must find a strategy for creating a replicating portfolio that, in each future scenario or circumstance, will have identical payoffs to those of the target. This is mostly engineering, and *constructive* or *synthetic*.

Most of the mathematical complexity in finance involves the description of the range of future behavior of each security's price. This is mostly science, and *reductive*.

1.3.6 Styles of Replication

There are two kinds of replication, *static* and *dynamic*. Static replication is old hat. A *static replicating portfolio* is a collection of securities that, once defined and never altered, reproduces the payoff of a target security under all future scenarios. Static replication the simplest and most comprehensive method of valuation, but is feasible only in the rare cases when the target security closely resembles the liquid securities available.

For more complex or non-linear securities, like vanilla stock options, static replicating portfolios are unavailable, but sometimes (the great discovery of Black, Scholes and Merton) you can find a *dynamic replicating portfolio*, a mixture of liquid securities whose proportions, continually adjusted by switching from one security into another, will replicate the payoff of the target. Static or dynamic, the initial price of the replicating portfolio becomes the estimated value of the target. But you can try to replicate more exotic options out of vanilla stock options, as is does for variance swaps, which is another topic we'll cover later in the course.

Dynamic replication is very elegant, and all the advances in the field of derivatives over the past 30 years have been connected with extending the fundamental insight that you can sometimes replicate complex securities dynamically. But, and practitioners who work with trading desks know this from their own experience, it's not a walk in the park.

Neftci's *Principles of Financial Engineering*, pp. 188-89, is one of the few books I've seen that points out the actual problems that arise when you try to replicate dynamically:

> "Real life complications make dynamic replication a much more fragile exercise than static replication. The problems that are encountered in static replication are well known. There are operational problems, counterparty risk, and so the theoretically exact synthetics may not be identical to the original asset. There are liquidity problems and other transactions costs. But all these are relatively minor and in the end, static replicating portfolios used in practice generally provide good synthetics.
>
> With dynamic replication, these problems are magnified, because the underlying positions needs to be readjusted many times. For example, the effect of transaction costs is much more serious if dynamic adjustments are required frequently. Similarly, the implications of liquidity problems will also be more serious. But more importantly, the real-life use of dynamic replication methods brings forth *new* problems that would not exist with static synthetics."
>
> We have to worry not just about current liquidity and bid-ask spreads, but about how they vary in the future.
>
> Dynamic replication is imperfect; it depends upon models, which imply assumptions and the approximations involved in working in discrete time steps.

> Even if the theory is easy, “the strategy needs to be implemented using appropriate position-keeping and risk-management tools. The necessary software and human skills required for these tasks may lead to significant new costs, but also to many jobs producing and taking care of these tools.
>
> Finally, dynamic replication is often used to replicate securities with non-linear payoffs. This leads to exposure to the level of volatility, and who knows what the future level of volatility will be. Managing exposure to volatility can be much more difficult than managing exposure to interest rates or currencies, because there are (almost) no underlyers to trade.”

So, in this course, wherever we can, we will *first try use static replication* for valuing new securities. If we cannot, *then we will use dynamic replication*. And finally, if we can’t replicate exactly, then you have to speculate on risk and return, which is less mathematical but more difficult.

1.3.7 The Limits of Models

Models are only models, toy-like descriptions of idealized worlds. No mathematical model is likely capture the intricacies of human psychology and the strategy changes it leads to.

For that reason, because models are unreliable guides to the world of finance, and because you don’t know which is the right one, it’s best to use as little modeling as possible. And, if you have to use a model, it’s always good to use more than one so you understand the model-dependence of your result.

1.3.8 Implieds and Realizeds

Physics models start from values today and infer the future. Only the future values are relevant.

Financial models have a kind of double vision. They calibrate the future to current known prices to produce implied variables for the future that match known prices today. One then has to compare these implied values to the future values that are actually realized as time passes. For example, one can compare today’s one-month implied volatility to realized volatilities over the next month, or implied interest rates to realized ones. An interesting question that follows is: how seriously do you take implied variables? Should you hedge with implied volatility or realized volatility? We’ll discuss this and see the effects of the choice in the next lecture.

You don’t have these problems in physics or engineering, because current values of variables are not “implied by” and don’t depend on future behavior, but, rather, future behavior depends on current values.

1.3.9 Testing Models

Here are some counter-intuitive and interesting remarks about quantitative finance by Fischer Black.

"It's better to 'estimate' a model than to test it. I take 'calibration' to be a form of estimation, so I'm sympathetic with it, so long as we don't take seriously the structure of a model we calibrate. Best of all, though, is to 'explore' a model."

"My job, I believe, is to persuade others that my conclusions are sound. I will use an array of devices to do this: theory, stylized facts, time-series data, surveys, appeals to introspection and so on."

"In the real world of research, conventional tests of [statistical] significance seem almost worthless."

No-one could make these remarks about models or theories in physics, chemistry or engineering. Think about the differences between these fields and quantitative finance, and why that should be the case.

1.3.10 Modeling (Relatively) Riskless Bonds

How do you value the promise of a future payment? The simplest answer is: find an institution, corporation or government whose liquid bonds have an estimated risk of default similar to your debtor's and use their term structure of interest rates to discount the payoffs of the target bond.

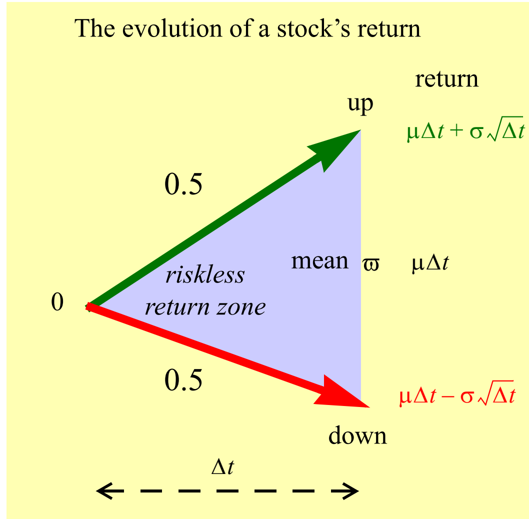
Why use interest rates? Because they allow you to compare investments. Quoting bond prices in terms of yields is a sort of model too, albeit a simple one, that projects bonds of different coupons and maturities onto a one-dimensional yield scale that's more enlightening than mere dollar price. The most easily embraceable models convert inexpressive prices into a more eloquent one- or two-dimensional scale that makes comparison more intuitive.

Of course, when you get more sophisticated, risky bonds need more complex models of default, and require options-based models.

Complex bonds, mortgages or swaptions, for example, whose payoffs are contingent on interest rates or other market parameters are best modeled using dynamically replicating portfolios, as described later.

1.3.11 5. The Risk of Stocks

A stock's most important feature is the uncertainty of its return. About the simplest model of uncertainty is the risk involved in flipping a coin. The figure at right illustrates a similarly simple model, a binary tree, for the evolution of a the price of a stock with return volatility σ over a small instant of time Δt. On an up-move the price increases by $\mu\Delta t + \sigma\sqrt{\Delta t}$, while on an equally probable down-move it increases by only $\mu\Delta t - \sigma\sqrt{\Delta t}$. The volatility σ is the measure of the stock's risk.



According to the law of one price, the riskless rate of return r must lie in the zone between the up- and down-returns. For, if both the up- and down-returns were greater than the riskless return, you could create a portfolio that is long \$100 of stock and short \$100 of a riskless bond with zero price and a paradoxically positive payoff under all future scenarios. Any model with such possibilities is in trouble before it leaves the ground. A riskless return in the riskless zone can be written as a convex combination of the up- and down- returns where the pseudo-probabilities of up and down moves will ultimately be the no-arbitrage probabilities in the risk-neutral measure used in options valuation.

This apparently naive either-up-or-down model captures much of the inherent risk of owning a stock and many other securities. Repeated over and over again for small time steps, it mimics the more-or-less continuous motion of prices in a reasonable though imperfect way, much as movies produce the illusion of real motion by changing images at the rate of 24 frames per second.

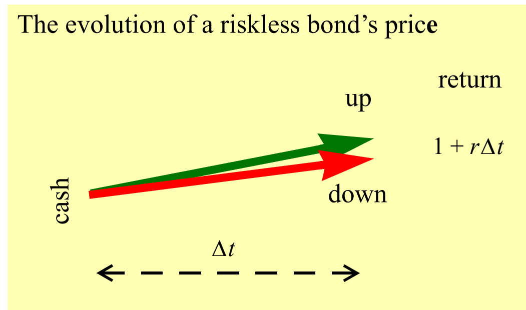
1.3.12 Risk Reduction With Riskless Bonds: More Risk, More Excess Return

What future expected reward justifies a particular present risk? This is the paramount question of life and finance. If you know the future volatility σ of a stock, what rate of return μ should you expect?

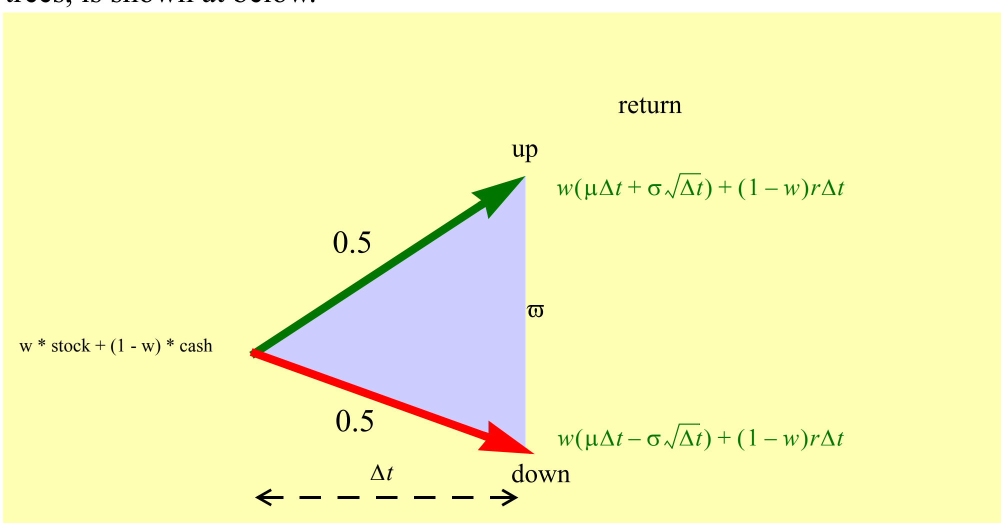
The law of one price tells you how to value securities you can replicate, but some payoffs cannot be replicated. Certain risk factors are intrinsic and unavoidable. We can extend the law of one price (identical payoffs have identical prices) to demand that *identical expected risks have identical expected returns.* But we have to be careful; some risks can be mitigated. Therefore, to be more precise, we demand that *identical unavoidable expected risks have* *identical expected returns.* We can use this less restrictive version of the law of one price to determine the relation between risk and return, illustrated in what follows in our binomial world. The key ideas is that there are three ways to modify risk: (i) by adding a riskless bond to the portfolio of stocks, (ii) by diversifying, and (iii) by hedging away a common factor.

We begin by examining the effect of adding a riskless short-term bond to any risky position.

The binary tree describing a riskless bond, shown at right, is degenerate: whether the stock moves up or down, the bond produces a guaranteed return ($1 + r\Delta t$), as shown at right. The returns of the riskless bond have zero volatility. By adding a riskless bond with zero volatility to the stock of volatility σ and expected return μ, you commensurately reduce both the risk and return of your investment.



Consider a mixture of w% risky stock with volatility σ and (1 - w)% riskless bond. The binary tree for this mixture, obtained by combining the two previous trees, is shown at below.



The expected return for this mixture is $w\mu\Delta t + (1-w)r\Delta t$. The volatility of returns is $w\sigma$.

Suppose we chose $w = \sigma'/\sigma$. Then the volatility of the mixture is σ' and the expected return is $\mu' = r + (\sigma'/\sigma)(\mu - r)$. Rearranging terms, we see that

$$\frac{\mu' - r}{\sigma'} = \frac{\mu - r}{\sigma} = \lambda \qquad \text{Eq.1.1}$$

All stocks or portfolios of stocks must have the same Sharpe ratio λ, and so

$$\mu - r = \lambda\sigma \qquad \text{Eq.1.2}$$

Excess return is proportional to risk.

In the next class we'll show that the a similar result holds for options values, and is equivalent to the Black-Scholes equation.

1.3.13 Risk Reduction by Diversification

You can reduce risk by combining a stock with a riskless bond. But you can also diminish risk by diversifying. If you can accumulate a portfolio of so many uncorrelated unavoidable risks that their risks cancel in the limit as the number of stocks become large, so that the portfolio's net volatility approaches zero, then, by the law of one price, it must produce the riskless rate of return. The Sharpe ratio in the equations above for this portfolio must be zero. But the Sharpe ratio is the same for all portfolios of stocks, so that $\lambda = 0$ in general. Thus,

$$\mu = r$$

, each stock in the portfolio must earn the riskless rate too, so that $\mu = r$ and the Sharpe ratio λ is zero If all risk factors can be cancelled by diversification, investors should expect only the riskless return on any single stock.

1.3.14 Risk Reduction by Hedging

Stock markets aren't truly amenable to the total diversification we assumed above. That's because large financial markets are more than a collection of individual uncorrelated risk factors. Experienced investors are always trying to detect patterns in the universe of stock returns. They perceive stocks as belonging to groups that have in common their sensitivity to a particular asset, factor or set of factors.

Let ρ be the correlation of the returns between any stock S with volatility σ and some tradable risky asset M with volatility σ_M. Because of the correlation between each stock's return and that of M, you can *hedge away* the M-related risk of any stock. If you have \$100 invested in a stock S, you can short β times as many dollars of the factor M against it, where $\beta = \rho(\sigma/\sigma_M)$ is the number of percentage points the stock S is expected to rise when M rises by 1%. This *factor-hedged* portfolio, consisting of a \$100 long position in the stock combined with β times as much of a short position in M, is expected to carry no net exposure to the price of M, because any increase in the price of the stock will, on average, be cancelled by a correlated decrease in value of the short position in M.

The net expected return on this M-neutral stock is proportional to the return of the stock S less β times the return of M, namely $\mu - \beta\mu_M$. This portfolio's return is uncorrelated with M, and, assuming there are no other factors influencing the stock, its residual risk is unavoidable. Only the residual risk/volatility of each stock remains.

Now, having hedged away the risk of M, we can combine the resultant M-neutral stock with a riskless bond to obtain any residual volatility. The excess return of the M-neutral stock must then be proportional to its residual volatility, as in Equations 1.2. Furthermore, by diversifying over many stocks whose M-risk has been hedged away, we can show that the Sharpe ratio here is zero too. As a result[1],

$$(\mu - r) = \beta(\mu_M - r) \qquad \text{Eq.1.3}$$

This is the result of the Capital Asset Pricing Model or Arbitrage Pricing Theory: in a world of rational investors, the excess return you can expect from buying a stock is its β times the expected return of its hedgable factor. Put differently, you can only expect to be rewarded for the unavoidable factor risk of each stock, since all other risk can be eliminated by diversification.

The rationality of investors is still hotly debated. As with most questions in finance, a social science overlaid with a useful veneer of quantitative analysis, the exact truth is hard to determine.

1. Spelled out in more detail in Section 2 of *The Perception of Time, Risk and Return During Periods of Speculation*, Quantitative Finance Vol 2 (2002) 282–296, or http://www.ederman.com/new/docs/qf-market_bubbles.pdf

1.3.15 Derivatives Are Not Independent Securities

A derivative is a contract whose value is determined by a specified relation between its payoff and that of a simpler security called its *underlyer*. Often, the relation is nonlinear. The most relevant characteristic of a derivative is the *curvature* of its payoff *C(S)*, as illustrated for a simple call option at right.

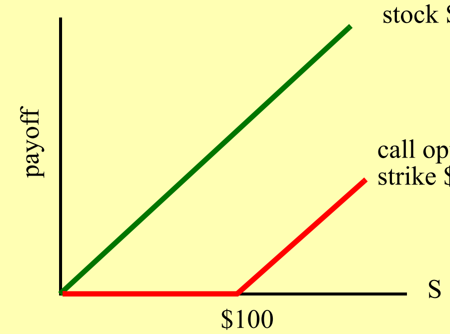
If you know the price of a simple stock, whose payoff is linear, what is the value of the nonlinear option? What is the value of curvature?

In the binary tree model, there are only two possible states (*up* and *down*) for the stock or the option after a time Δt. These states are spanned by two independent securities, the stock itself and a riskless bond. You can use linear algebra to decompose the stock and bond into a basis of two more elemental securities π_u and π_d, each respectively paying off in only one of the final states, as shown at right. By creating a portfolio that is a suitable mixture of these two one-state securities π_u and π_d, you can replicate the payoff of any non-linear function *C(S)* over the next instant of time Δt, no matter into which state the stock evolves. Note that the portfolio consisting of both π_u and π_d is riskless and pays out $\$(1+r\Delta t)$ and therefore the sum of their present values is 1.

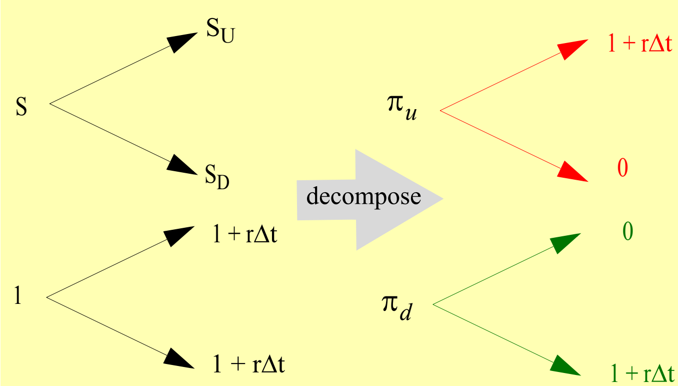
The value of the option is therefore the price of the mixture of stock and bond that replicates it. In the limit as Δt becomes infinitesimally small, you can repeat this replication strategy instant after instant, so that, if the final payoff C(S) is known, its replication at all earlier times is determined. The value of the replicating portfolio depends on the spread between the up-return and the down-return at each node in the binary tree, that is, on the stock's volatility σ. If you know the future volatility, you can synthesize an option out of stock and bonds.

The same strategy of dynamic replication can be extended to more complex and realistic models of stock evolution, as well as to the replication of derivatives on all sorts of other underlyers.

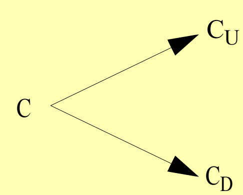
A stock is straight, an option curved.



Binary price trees for a stock S and a $1 investment in a riskless bond. The stock and bond can be decomposed into a security $\pi_{\mathbf{u}}$ that pays off only in the up state and a security $\pi_{\mathbf{d}}$ that pays off only in the down state.



You can replicate an option's nonlinear payoff over each instant by suitable investments in the elemental securities π_u and π_d.



$$(1+r\Delta t)C = (C_U).\pi_{\mathbf{u}} + (C_D).\,\pi_{\mathbf{d}}$$

1.3.16 The choice-of-currency/numeraire trick

It seems obvious that value of a security in dollars should be independent of the currency (US dollar, Japanese yen, share of IBM, etc.) you choose for modeling its evolution. A little thought often suggests a natural choice of currency that can greatly simplify thinking about a problem, and sometimes reduce its dimensionality. Just as stock analysts find it convenient to use the P/E ratio to measure a stock's price in units of projected earnings rather than dollars, thereby making stock comparison easier, so financial modelers can sometimes cleverly choose a more meaningful currency than the dollar when valuing a complex security. Convertible bonds, for example, which involve an option to exchange a bond for stock, can sometimes be fruitfully modeled by choosing a bond itself as the natural valuation currency.

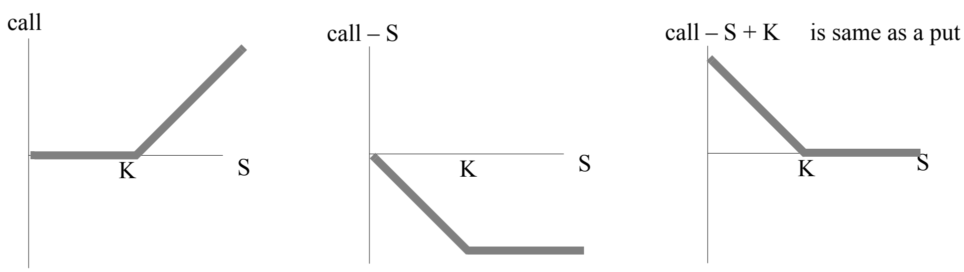
1.4 Methods of Replication

1.4.1 Static Replication

If you can create a static hedge or replicating portfolio for your payoff, you have very little trading or model risk.

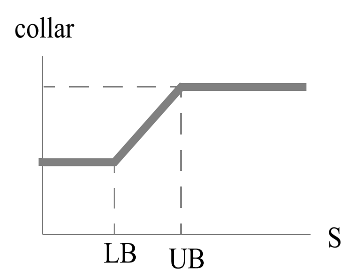
Put from a call:

Suppose you need a put but you can only buy a call. Then compare the payoffs at expiration:



Thus price of put = price of call - price of stock + PV(K).

A collar: A collar is a very popular instrument for portfolio managers who have made some gains during the year and now want to make sure they keep some upside but don't lose too much downside. The payoff is shown at right.



You can write the payoff as

$$LB + call(S, LB) - call(S, UB)$$

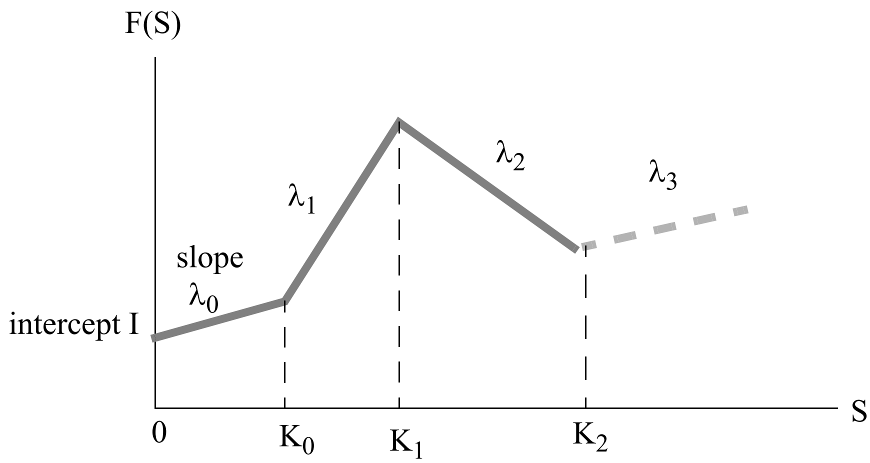
which is the payoff of a bond plus a call spread. Using put-call parity you can also write it as

$$S + put(S, LB) - call(S, UB)$$

which is long the stock, long a protective put and short a call that gives away the upside. A long position in a put and a short position in a call with a higher strike is called a collar.

Generalized payoffs:

Suppose you can approximate the payoff at a fixed single time in the future as a piecewise-linear payoff function of the terminal stock price S that is defined by its intercept and successive slopes as in the following payoff diagram:



Then the payoff is the payoff of a portfolio consisting of a zero-coupon bond ZCB(I) with face value I, plus a series of calls $C(K_i)$ with successively higher strikes K_i:

$$ZCB(I) + \lambda_0 S + (\lambda_1 - \lambda_0)C(K_0) + (\lambda_2 - \lambda_1)C(K_1) + \ldots$$

and the value of this generalized payoff is the value of the bond and the calls in the market.

You can check that, for example, between K_1 and K_2, the payoff of the portfolio above is

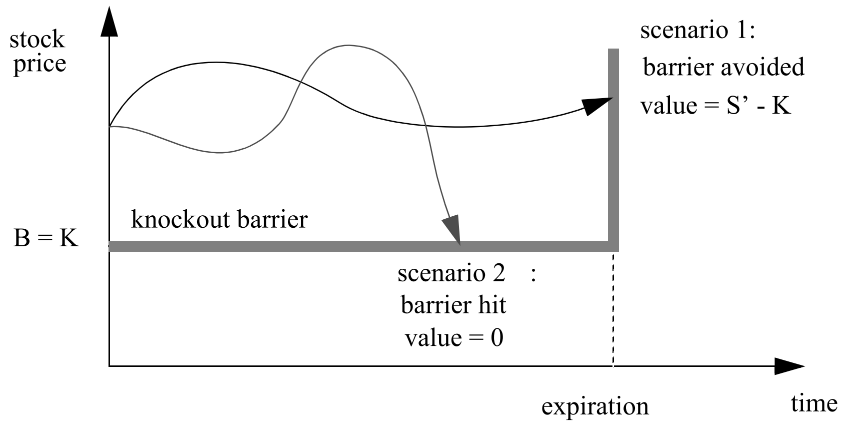
$$\begin{aligned} & I + \lambda_0 S + (\lambda_1 - \lambda_0)(S - K_0) + (\lambda_2 - \lambda_1)(S - K_1) \\ = \; & I + \lambda_0 K_0 + \lambda_1 (K_1 - K_0) + \lambda_2 (S - K_1) \end{aligned}$$

which is the correct intercept and the correct slope.

1.4.2 Static Hedge for a European Down-and-Out Call

Consider a European down-and-out call option with time t to expiration on a stock with price S and dividend yield d. We denote the strike level by K and the level of the out-barrier by B. We assume in this particular example that B and K are equal and that there is no cash rebate when the barrier is hit. There are two classes of scenarios for the stock price paths: scenario 1 in which the barrier is avoided and the option finishes in-the-money; and scenario 2 in which the barrier is hit before expiration and the option expires worthless. These are shown in Figure 1.3 below.

FIGURE 1.3. A down-and-out European call option with *B* = *K*.



In scenario 1 the call pays out $S'-K$, where S' is the unknown value of the stock price at expiration. This is the same as the payoff of a forward contract with delivery price *K*. This forward has a theoretical value $F = Se^{-dt} - Ke^{-rt}$, where *d* is the continuously paid dividend yield of the stock. You can replicate the down-and-out call under all stock price paths in scenario 1 with a long position in the forward.

For paths in scenario 2, where the stock price hits the barrier at any time t' before expiration, the down-and-out call immediately expires with zero value. In that case, the above forward *F* that replicates the barrier-avoiding scenarios of type 1 is worth $Ke^{-dt'} - Ke^{-rt'}$. This matches the option value for all barrier-striking times t' only if $r = d$. So, if the riskless interest rate equals the dividend yield (that is, the stock forward is close to spot[1]), a forward with delivery price *K* will exactly replicate a down-and-out call with barrier and strike at the same level *K*, no matter what.

When the stock hits the barrier you must sell the forward to end the trade.

1. In late 1993, for example, the S&P dividend yield was close in value to the short-term interest rate, and so this hedge might have been applicable to short-term down-and-out S&P options