THE YOUNG PERSON'S GUIDE TO PRICING AND HEDGING

An early outtake from my book *Models.Behaving.Badly*. An attempt to explain the principles of neoclassical finance from a rational point of view, and to point out the assumptions and their potential inapplicability. And no mean-variance optimization!!

Few people have the gift of grasping nature and using it directly: between knowledge and application they prefer to invent a phantom which they develop in great detail; doing so, they forget both object and purpose.

Maxims and Reflections: Goethe

OH ROSE, THOU ART SICK!

There is nothing so terrible as activity without insight.
Maxims and Reflections: Goethe

Recently I heard a visiting professor of finance give a seminar at Columbia. Later, discussing his work in my office, he justified a result in his presentation by invoking what he referred to as "the fundamental theorem of finance."

'Fundamental' means situated at the base, a foundation on which the entire field rests.

"Isn't it strange?" I thought afterwards, that a field such as finance might have a fundamental theorem (and that, despite more than twenty years in the field, I wasn't quite sure what it was).

When the visitor left my office I commenced a search. If you google the phrase "fundamental theorem" you soon come across the fundamental theorem of arithmetic:

Every natural number greater than 1 can be written as a unique product of prime numbers.

Anyone with a rudimentary knowledge of arithmetic can understand this, or have it explained to them in a few minutes. And you can see why someone might call it fundamental.

Google further and you find the fundamental theorem of algebra:

Every polynomial equation of degree n with complex number coefficients has n complex roots.

This too can be explained to someone familiar with algebra, though it's not quite as easy, since algebra is more sophisticated than arithmetic.

Next you can find the fundamental theorem of calculus, which states:

If f is continuous on the closed interval $[a, b]$ *and F is the indefinite integral of f on* $[a, b]$*, then the definite integral of* $f(x)$ *from a to b is* $F(b) - F(a)$.

Though formally expressed, the theorem says something simple: the slope of the area under a curve is the height of the curve itself. You don't need to be a rocket scientist to understand that, though calculus is much more difficult than algebra and required centuries of work to put it on a firm basis.

Arithmetic, algebra and calculus are branches of mathematics, and hence they can depend upon important theorems. But if you google "the fundamental theorem of chemistry" or the "fundamental theorem of physics" you won't find much, because everyone doing physics and chemistry understands that physics and chemistry are about the world around us. Those fields have laws, not theorems.

Finally, I googled the "fundamental theorem of finance". I didn't have to look too far down the list of URL's to discover the following statement:

Fundamental Theorem of Finance. *Security prices exclude arbitrage if and only if there exists a strictly positive value functional, under the technical restrictions that the space of portfolios and the space of contingent claims are locally convex topological vector spaces and the positive cone of the space of contingent claims is compactly generated, that is, there exists a compact set* K *of* X *(not containing the null element of* X*) such that*

$$C = \{x \in X : x \geq 0\} = \bigcup_{\lambda \geq 0} \lambda K.$$

How, in all seriousness, can anyone think that this nearly incomprehensible statement is fundamentally critical to finance, an ostensibly practical field that concerned with the management of money and assets?

THE DIFFERENCE BETWEEN THEOREMS AND LAWS

Theorems are "if-then" relations in mathematics, statements that specify the results that follow from certain assumptions. The "ifs" are more or less self-evident facts. The "thens" are the consequences that follow.

In geometry, Euclid's axioms and postulates are the unquestioned "ifs", and they state the properties of points and lines. One of Euclid's postulates, for example is that it is always possible to draw a straight line between any two points. One axiom is that things that are equal to the same thing are equal to each other. Hard, though perhaps not impossible, to argue with.

The points and lines that Euclid's *Elements* refers to are 'points' with no size and 'lines' with no width, idealizations of the scratchy dots and strokes we make with a stylus on a slate. They are the points and lines we imagine we could make with a perfect stylus on a perfect slate, pure mental constructs abstracted by intuition from the world and then detached from it. From them, Pythagoras's theorem proves a relation between the areas of the three squares that can be drawn, respectively, on the sides of any right-angled triangle.

Though Euclid's points and lines are abstractions, when you get familiar enough with them, they assume a reality that is hard to ignore. Even more esoteric abstractions – constructs like the vectors in the infinite-dimensional Hilbert space that forms the mathematical basis of quantum mechanics – can become real and visualizable to mathematicians.

Mathematics has theorems, thens followed by ifs. Mathematics deals with abstractions, points and lines, for example. The theorems it derives are relations between the abstractions, not necessarily between the realities that inspired them.

Science, in contrast, has laws. Laws are not "if-thens", not conditional. They describe the way the universe works. Newton's laws of motion and his law of universal gravitation describe how masses move under the influence of forces, and thereby allow us to guide rockets to the moon. Maxwell's equations (laws themselves) allow construction of radios and TV. The principles of quantum electrodynamics drive electronics. The 19th Century laws of thermodynamics constrain the conversion of heat into energy and constrain the construction of combustion engines. Even the law of evolution, a law of style rather than of quantity, formulates the hypothetical principles by which species develop.

Physics deals with abstractions too, but there is a unity between the abstraction and the object it represents. The idea and the object are two aspects of the same thing. The laws discovered *are* valid relations between the abstractions and the actual objects that inspired them.

Finance's objects of interest – markets, money, assets and securities – are also abstractions. The aim of finance is finding true relationships between the abstractions they represent, and hence between the realities they should be pinned to. In that sense, like physics, finance cannot be a branch of mathematics. Finance could conceivably have a fundamental law – *Money tends to flow into the hands of people who will let it reproduce most rapidly* might be one – but it surely cannot have a fundamental theorem, for the world is not conditional.

Only people who don't understand how accurate knowledge is discovered and accumulated can imagine that there is a fundamental theorem in finance. Only people who don't understand the difference between what's outside them and inside them can imagine that there are theorems in finance. Only people who can't distinguish between God's creations and man's idols can take their models for the truth.

Unfortunately, many economists are these kinds of people. If you open up the very academically oriented *Journal of Finance*, one of the small number of select journals in which professors in finance departments must publish if they want to get tenure, most of the papers resemble those published in a journal of pure mathematics. Replete with axioms, theorems and lemmas, they have a degree of rigor that is inversely proportional to their minimal usefulness.

The trouble is that economists have fallen in love with mathematics, rigor and formalism for their own sake, irrespective of their efficacy. The simple models they work with fail to reflect the complex reality of the world around them. It's not their fault that they can't find better models; economics is a social science and people are difficult to theorize about. But it is economists' fault that they take their simple models as gospel.

In physics, we are, so to speak, playing against God, and somehow, miraculously, we can intuit the theory behind his laws. In finance we are playing against God's creatures, and there are only dubious models: no theories have yet been found. Physicists, who have grown up on a diet of miraculously successful and literally unbelievably accurate theories, have the common sense to distinguish a theory from a model and a good model from a bad one. Economists for the most part, have never been exposed to a successful theory in their entire scholarly life, and so they are less capable of discrimination between good and bad.

Finding the truth about nature takes cunning and flexibility. The invisible worm of academic economics, and the subfield of finance in particular, is its dark secret love of mathematical elegance and its belief that one can replace cunning with rigor.

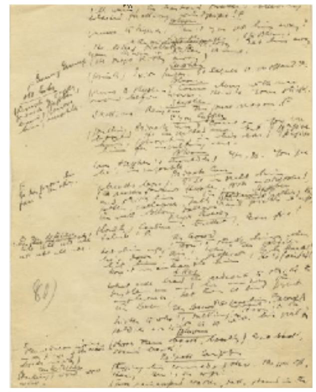
STYLE & CONTENT

> Content without method leads to fantasy; method without content to empty sophistry;
>
> Goethe: *Maxims and Reflections*

During the summer before I went to college in the Sixties, being a conscientious kid, I spent a few hours each day practicing touch-typing. It was supposed to be important to be able to type. I bought a mimeographed instruction book, and practiced the exercises on my parents' old mechanical typewriter with its symphony-hall seating orchestra of keys and its ostrich-necked typebars with letters at the head. The c*irca* 1955 one I used ten years later didn't look all that different from the picture at right. The keys jammed when you hit two of them simultaneously. Unless you practiced long and hard, letters typed with your left hand, especially with the little finger, left a distinctly lighter impression on the paper than letters typed with the right hand. It took physical strength.

In that era before word processors, spell checkers and laser printers, it took substantial effort to make a manuscript look professional, and so one didn't polish the appearance until the content was pretty much perfect. Sloppy typing mirrored sloppy thought, and even manuscripts with carefully formulated thoughts didn't appear beautiful. Take a look at museum copies of manuscripts of Joyce's *Ulysses*. Nowadays, with little care, one can make bad writing look professional. The clarity of electronic typesetting almost always exceeds the clarity of the thoughts expressed[1].

FIGURE 6.1: Aeolus, the seventh episode of Ulysses. Taken from the Rosenbach Museum at http://www.rosenbach.org/learn/collections/james-joyces-



Niels Bohr is reputed to have said about some other physicist: He writes more clearly than he thinks. Analogously, the precision of the sophisticated mathematics used in finance almost always exceeds the clarity and scope of the ideas that drive it. But ideas are what matter; mathematics is merely the language in which they are best expressed. Even the sublime Dirac equation, difficult to interpret with pictures, expresses (i) the idea that the equation for the electron should have the same form in all coordinate systems traveling less than the speed of light, and (ii) the conviction that even negative-energy solutions must carry physical significance.

Who would think that the statement of the so-called fundamental theorem of finance deals with people and money?

1. When I worked at AT&T Bell Laboratories in the early 1980s, my supervisor told me that someone high up at Bell Labs told him that the professional appearance of your presentation should vary inversely with the importance of the people you are presenting it to. If you make a presentation to people below you in the pecking order, the slides can be highly polished, indicating you expect no corrections from people below you. But, if you make a presentation to someone well above you, the slides should look unformed and casual, to indicate that your superior is therefore free to interrupt and make changes. A polished document would indicate an arrogant finality.

FIRST STEPS IN FINANCIAL THEORY

> It would be most desirable, however, to base the language for the details of a particular area on the area itself.
>
> Goethe: *Maxims and Reflections*

In this chapter I want to provide a clear account of the so-called "theory" of financial valuation.

Is the "theory" really a theory? I would say not. If the theory treated markets and their participants from observation, it might qualify. But whereas Spinoza treats the emotions as phenomena to be examined on their own terms, financial models transfer into their realm theories from more reliable fields (heat conduction and diffusion for example). Financial models are analogies. Their assumptions, though they are not completely implausible, can be seen by examination and introspection to be flawed. You cannot say the same about Dirac's theory of the electron.

My aim in what follows is to illustrate the classic foundations of financial modeling, to emphasize the assumptions being made, to comment on their plausibility, and to illustrate their limitations.

Securities

For all practical purposes, a *security* is a promise by its seller (a person, a corporation, an institution) to make *future* payments to its buyer The promise can be made in writing or electronically. A security represents *present* financial value, and financial theory aims at figuring out what the security is worth, and, if it sensible, why.

Debt Securities

The simplest kind of security is a straightforward loan: you lend money to someone who signs an IOU that obliges him to repay you. A loan is an instance of a more general debt security.

A Treasury bond is a debt security too: you lend a sum of money to a government and they promise to pay you a specified rate of interest, say 4.25% per year in two semi-annual coupon payments per year, and to return the principal (the original sum) after thirty years.

On the day you made the loan, you handed over a sum of money to the government and in return they promised to make future payments. From a different point of view, you bought their promise: they sold you the Treasury bond security that promises payments, for which you paid a price equal to the amount of the loan you made to them. Like the government, corporations can borrow money by selling their promises too, called corporate bonds. Similarly, individuals who borrow money to buy a house via a mortgage are in essence selling a promise to pay future interest and principal to the lender, with the house as collateral.

Debt securities are characterized by precise specifications: their future payoff is usually spelled out very clearly. The Treasury bond above, for example, pays exactly 2.125% of the principal semi-annually.

Equity Securities

Equity securities are characterized by less certain payoffs.

A share of stock in a public company is the most common equity security. When you buy a share of stock issued by the company, they can use the payment to run their business exactly as they please[2]. Instead of a promised rate of interest, In exchange for the purchase price you get a specified fraction of any future earnings the company generates, *if* it generates earnings. The presence and quantity of future earnings is uncertain and therefore more risky than the promised coupon payments of a bond. There is an implicit "if" to a share of stock, the contingency that earnings exist.

Most shares of stock also confer on shareholders the legal right to attend and vote at annual meetings of the corporation, to have a say in selecting boards and decisions about potential mergers. Although they are not always exercised, these rights are valuable and yet hard to quantify. This additional layer of strategic complexity is part of the reason that stock markets get more public attention than debt markets; everyone loves a good story.

---

2. To be precise, when you buy the share of stock the first time it is issued, in an IPO (initial public offering), the company receives your payment less the investment banking fee. If you buy the share in the secondary market later, it's like buying a second-hand book; you pay the owner of the book for it, not the author.

Derivatives

While shares of stock have an implicitly contingent nature, *derivative contracts* are explicitly contingent.

Futures contracts, forward contracts and stock options are securities that involve even greater uncertainty than equity. All of them are *derivative* securities, sometimes referred to more pedantically as *contingent claims* on the underlying stock. The promised future payments that the seller of any one of these securities is obliged to make depend on what happens to the price of the stock in the future; in other words, the securities *derive* their value from the value of the stock itself, since the future payments are *contingent* on the stock price. For Spinoza, *Love* is *Pleasure* associated with an external object, and hence a derivative of *Pleasure*. Analogously, a simple *Call Option* on a stock is a security whose payoff at expiration is the amount by which the stock price has risen (if any) since the purchase of the option. As *Love* derives from *Pleasure*, so the *Call Option* derives its value from the motion of the *Stock Price*.

Derivative securities are more complex than bonds or even shares of stock: the payments the seller of the security promises to the buyer depend in via a specified mathematical formula on what happens to the stock in the future. The seller of a simple call option promises to pay the buyer on some specified future date the amount that the stock price increases over its current value, if it increases; if it decreases, the seller will owe the buyer nothing. For the *Pain* of the cost of the option, its owner will get future *Pleasure* if the stock price rises, but no additional *Pain* if it falls.

A forward contract on a stock is more symmetric than an option; it has upside as well as downside risk. If the underlying stock price goes up, the buyer of the contract gets the *Pleasure* of receiving a payment from the seller equal to the gain in the stock price, but conversely, if the stock price goes down, the buyer gets the corresponding *Pain* of having to pay the decrease in the stock price to the seller.

Because of their contingent nature, the values of derivative securities varies with the price of the security that lies beneath them in subtle, unexpected and unintuitive ways. Complex derivatives are likely to have extreme sensitivities to their underliers, as was the case with the CDOs (collateralized default obligations) in the financial crisis of 2007 – 2008. Their values depended on the value of the underlying housing market, and spread that weakness through the entire investment world to everyone that owned them.

Extreme sensitivity is not a good thing, especially when attempts to calculate it depend on an inaccurate model.

Markets

Securities are traded on markets, which are locations, real, or in recent times, increasingly virtual, where participants meet to exchange securities for money.

The most straightforward are *listed* markets, on which information about participants' offers to buy or sell a security and a record of the price at the most recent sale are listed in some central location available for all participants to see. Nowadays the central location is usually a computer, or, more precisely, a farm of computer servers distributed around the world, a location that is central in the sense of thought rather than extension. Formerly, the central location was an actual marketplace where people went from stall to stall and shouted bids and offers to each other in a crowd, or listed them on a printed sheet.

Listed markets deal with conventional standardized securities that are easily described, easily obtained, popular and available in quantity – stocks for example, or fairly simple so-called vanilla options. Less uniform securities that come in many similar but not identical variations – bonds for example, which can have an extensive variety of maturities and coupons – commonly trade in *over-the-counter* (OTC) markets where buyers and sellers negotiate prices between themselves in private, either face to face, by telephone or electronically.

On a listed market everyone can see the most recent prices, so that there is less scope for being fleeced through ignorance. Nevertheless, many securities that should by all logic trade on listed markets still trade over-the-counter because the heavyweights in a market, who prefer the advantage that the lack of transparency gives them when dealing with smaller participants, can get away with it.

News and information affect prices, and prices themselves are news and information. The information from listed markets is therefore critical in developing models or theories of prices. Many agents in financial markets try to leapfrog the information flow by legal means (using algorithms to try to predict future price movements from past ones, placing their servers closer to the source of the prices) or illegal ones (paying for *insider* information, a term whose legal scope keeps expanding to include formerly common behavior, but not fast enough to keep up with people's ability to stay one step ahead).

Value and Uncertainty

No one wants to pay more for a security than it is worth, and no one wants to sell it for less than its worth[3]. People care about value. The aim of models of financial valuation is to estimate the monetary worth of financial assets of all kinds.

Worth is a mental quality, and therefore a matter of opinion and subject to uncertainty. In contrast to the value of the electron's mass or charge, in contrast to the strength of the force of gravity, there is nothing absolute about the value of a financial asset.

In physics you can travel a very long way before you run into uncertainty. The classic triumphs of the field – Newton's laws of motion and gravitation, the explanation of orbits of the planets, Boyle's law for the relation between the pressure and volume of ideal gases, Maxwell's equations for electromagnetism, and the study of the transformation of heat into mechanical energy in thermodynamics – are all triumphs of deterministic understanding. Only when you begin to try to understand the link between the invisible microscopic atoms and the familiar visible macroscopic bodies they constitute do you finally encounter the notion of randomness. Classical randomness is a way of dealing with lack of detailed knowledge. One step deeper and smaller and you face intrinsic randomness, the quantum unknowability summarized in Heisenberg's uncertainty principle. But even there, determinism lingers: the Schrodinger equation that describes probabilities evolves predictably according to Schrodinger's equation. Even in quantum mechanics, the future can still be divined.

In finance, the thread of uncertainty runs through every square inch of the fabric, as it does through all human affairs. You can't begin to think about the discipline without recognizing the key element of unpredictability. We don't know what will befall a security, because we don't know how the future will affect the promises made by its sellers. Value is in the mind of the beholder, and uncertainty clouds it to

3. It is tempting but nevertheless a mistake to ignore non-economic value. When people buy a security for more than it seems to be "worth", it is usually because they are receiving something else with it, perhaps even something intangible, but nevertheless valuable.

the point where it cannot be accurately quantified. We don't know the laws that determine how value evolves in time. Despite Spinoza's expectation of the eventual discovery of adequate causes for everything human, I have to confess that I don't believe such laws exist.

Quoting Value

Financial values are measured in units, and the easiest unit to use is units of money. You can quote value in any currency: dollars, British pounds or Euros, for example. Money is just another financial asset, a metal coin or a piece of paper that governments promise to more or less stand behind, unless things get too bad, and they often do. What differentiates money from other securities is that, in the short run[4], it's the least risky financial asset and is therefore useful as a stable store of value and a common denominator for valuing other assets. You can keep it under your mattress or in the bank (but don't think too hard about where it actually is while it's "in the bank").

In addition to its short-term stability, money is also homogeneous and easily available, and thus serves as an accepted medium of exchange. Nevertheless, you don't have to denominate financial value in terms of dollars or Euros; you can quote a value in ounces of gold or bushels of wheat or shares of Google stock or sticks of Wrigley gum. Sometimes these denominators provide a better ruler for measuring value; if you are estimating the value of an oil company, for example, you might be better off valuing it in terms of barrels of oil rather than dollars. A good start to any financial model is thinking about the best *numeraire*, to use the aficionados term for the unit of theoretical value.

Price vs. Value

Don't get confused between price and value. Price is what it costs to buy a security, or what you receive for selling it. Value is what you think it is worth. The difference between them is what modeling and investing is about.

---

4. In the long run, of course, paper money is very risky; governments collapse, empires fall, and indestructible things that last longer than empires hold their value much better than their paper. Jim Grant has pointed out that an ounce of gold has always more or less been the price of a good suit.

JIU-JITSU FINANCE: THE EFFICIENT MARKET HYPOTHESIS

The Great Financial Crisis of 2007-2008 triggered an intense interest in the nature of quantitative financial models and their apparent inability to predict the disasters that occurred. The initial temptation to blame complex mortgage models for the market's near dissolution was overwhelming. More recently, observers have become more measured; Paul Krugman and Robin Wells [5] have pointed out that identical disasters occurred even in markets where only simple mortgages (that required no sophisticated models) were traded, Spain, for example.

I will argue later in this chapter that it is only naive users of financial models who expect them to predict the future. Models in finance rarely work well as tools of prediction; their major purpose is to translate your (good or bad) qualitative intuition into appropriate quantitative values for securities. If your model is bad or your intuition is wrong, you will get the wrong answer.

At the core of modern financial models is the Efficient Market Hypothesis (EMH). Despite its problems, it is worth a very careful look, not only because of the attention it has received as models have failed, but also because examining it in detail it makes clear the nature of the assumptions made in financial modeling. The Efficient Market Hypothesis is a simple, powerful and simplistic model of human behavior, better referred to as the Efficient Markets Model. There is nothing mysterious about it, but it is not a theory.

A Share of Stock Only Looks Simple

Because stocks are the among the most common and widely traded securities, I will use a share of stock as a prototypical security and focus on explaining models for stock valuation. Similar techniques apply to other securities too.

A share of stock is an investment in a company that grants you a stake in their business and consequently entitles you to a proportion of its profits. The invention of publicly owned companies and the limited liability of their owners for losses was one of the great facilitators of the spread of capitalism.

A company - take Apple, for example, who make the admirable laptop I'm writing on now – is a tremendously complex and structured endeavor. Apple owns

5. http://www.nybooks.com/articles/archives/2010/sep/30/slump-goes-why/

or leases buildings in many countries, has tens of thousands of employees, designs products ranging from desktop and laptop computers through iPhones and iPads to power plugs and cables, manufactures some of them on its own, farms out the manufacturing of others to China, distributes its products by mail order through Apple's own and other stores, and sells music and videos over the internet. They advertise, provides product support, runs websites and, enabling all of this activity, carry out research and development.

Apple is a large city. The economic value of its organization is reflected, for better or for worse, in one number, the listed price of a share of their stock, which represents the market's consensus on the value of the company. This price is the amount of money it took to buy or sell just *one* single share of the company. It is not necessarily equal to the price per share for acquiring the entire company, for which, you might expect a bulk discount, or, conversely, you might perhaps expect to pay a premium once you disclose your urgent desire to get all their stockholders to part with their shares and hand over control to you.

The price of one share of stock! Financial modeling is an attempt to project the value of the entire enterprise and all its facets, from management skills to R&D efforts, into that one number situated on a one-dimensional scale. That single number is supposed to tell you what you should pay today for a share of its future performance.

Your body is a city too. Imagine that doctors attempted to project your health onto a one-dimensional health-quotient scale that ran from zero to 100. It would be a useful simplification, but it wouldn't necessarily give a rounded picture of your future performance; health is too intricate to be encapsulated accurately in just one number. Doctors use a continually increasing set of numbers – weight, body fat percentage, cholesterol levels, bone density, white blood count, and so on – to report on your health, and no one of them, or even all of them together, suffices to determine your future health worth. Both financial and physical health are subject to error and uncertainty, as well as to the unexpected arrival of new diseases, cures, surgical techniques and inventions that may make sick people or companies better, or well ones sicker. Time and chance affect all. As a result, a stock price may not be an accurate indicator of a company's future.

If you think an asset is worth more than its most recent market price, you can buy it and hope to make a profit. For centuries, therefore, men and women have tried to predict the magnitude and direction of changes in the price of a bushel of wheat, an ounce of gold, or a share of stock, from moment to moment, day to day, month to month or year to year. They haven't had any overwhelming success. Some predictors work from fundamentals, dissecting an entire company, its management, products, pipeline, budget and style, as well as the state of the economy in which it operates. Others rely on technical analysis, a combination of rational and magical thinking that involves spotting the repetition of patterns in the trajectory of stock prices. Some patterns involve the quasi-scientific use of Fibonacci series and more esoteric mathematical structures. Perhaps these patterns, despite their cabalistic appearance, do reflect underlying psychological traits of collective human behavior. Value *is* mental. Investment and speculation are acts by people who are naturally subject to whim, superstition, foolish optimism and unnecessary panic and, especially, the influence of others' behavior. Crowds give rise to patterns.

In detail though, stock analysts don't do consistently well. When a physicist discovers a theory or model that correctly predicts the future behavior of a physical system, other physicists can tell whether it makes sense. When a stock analyst does the same in the financial domain, you can never be sure whether he was lucky or smart. A physicist can't influence nature, but a stock analyst can influence the market.

In the long run, fundamentals – the state of the economy and the state of the company – count for most, and even fundamentals will change. Meanwhile, opinions and passions count for much. Furthermore, the short run influences the long run: dips in stock price affect the confidence of the managers of the company and customers. Perception can become reality.

If you are honest and introspective, you realize that building a model of stock price movements is a vastly difficult problem that involves the interactions of investors and the company over many different time scales, where the imagined future can affect the present, and hence the actual future too. This makes the quantitative approach to the field much more daunting than electromagnetic theory, which is local in time and space.

Despite Spinoza, you can't hope for the success of physics in modeling human behavior.

The Efficient Market Model

It's a fact, then, that no one is very good at stock price prediction, whether they make use of magical thinking or examine deep fundamentals. Being right 55% or 60% of the time, consistently, over many trades, is remarkable and provides great profit. Faced with this failure, a school of academics associated with Eugene Fama at the University of Chicago in the 1960s developed what has become known as the Efficient Market Hypothesis (EMH), which I prefer to call the Efficient Market Model, since it's a model of a hypothetical world rather than the one we inhabit.

I was a persevering student of physics when the EMH became popular, though I didn't know about it then. Over the years many formulations have evolved, some more formal and rigorous, some less so. But if you take care not to get carried away, as many *faux*-precision-worshipping economists do, if you can avoid the temptation to define *strong*, *weak* and other kinds of "efficiency" as though you were dealing with a mathematical system rather than the world of humans and markets, then you will recognize that the EMH begins by acknowledging the following more or less true fact of life:

> It is difficult or well-nigh impossible to successfully and consistently predict what's going to happen to the stock market tomorrow based on all the information you have today.

The EMH formalizes this experience by stating that it is impossible to beat the market, because current prices of stocks reflect all that the information we have about the economy and the market. Only new information will change prices.

If this is true, then no one knows more than anyone else about what stock prices will do next. An important consequence of this modest acceptance of human limitations means that it may not make sense to pay people to pick stocks for you when they have no superior knowledge. About half of them will do worse than the market, and half better.

Converting their failed attempts at systematic stock price prediction into a fundamental postulate of their field was a fiendishly clever jiu-jitsu response on the part of economists. It was an attempt to turn weakness into strength: "I can't figure out how things work, so I'll make the inability to do that a principle on which to base a theory."

The efficient market model asserts that human beings cannot predict the future behavior of stock prices (which, don't forget, have no life of their own but are set by the behavior of the human beings who buy and sell them). The EMH beneath its formal cloak, is simply an assumption about human behavior. You could call it, instead, the Stupid Humans Hypothesis.

In Efficient Markets, Price Equals Value

The EMH therefore claims that at any instant, price is likely to be our best estimate of value. Stated that way, it sounds more dubious. Anyone with hindsight can see that the market is sometimes wrong.

Fischer Black, one of the discoverers of the Black-Scholes model for valuing options that grew out of the EMH, was a brilliant and original financial theorist, but also a great realist. In a widely read paper entitled *Noise*, he acknowledged, as few academics do, the vagueness with which value can be determined:

> All estimates of value are noisy, so we can never know how far away price is from value.
>
> However, we might define an efficient market as one in which price is within a factor of 2 of value, i.e., the price is more than half of value and less than twice value. The factor of 2 is arbitrary, of course. Intuitively, though, it seems reasonable to me, in the light of sources of uncertainty about value and the strength of the forces tending to cause price to return to value. By this definition, I think almost all markets are efficient almost all of the time. "Almost all" means at least 90%.

Even this estimate of the discrepancy between price and value may be optimistic. No one knows how to truly determine value or even exactly what value is, except via a model, and so the discrepancy between price and value can never be determined.

The scientific development of financial modeling is bound to be shaky when it starts from an inability to define value.

AN ASIDE: JIU-JITSU PHYSICS AND THE ANTHROPIC PRINCIPLE

Just as economists have invented the Efficient Market Hypothesis to account for the failure of their models to predict the future of stock prices, so physicists have invented The Anthropic Principle to account for their failure to derive the values of certain fundamental constants of nature.

The Rydberg Constant

The clinching triumph of Bohr's 1912 planetary model of the atom was the formula he derived to explain and predict the series of discrete wavelengths of light emitted by heated hydrogen gas, the so-called spectral lines of hydrogen.

At that time, physicists were aware of the so-called *Rydberg formula*, the following elegant empirical equation discovered years earlier that fit the wavelengths λ of all the spectral lines of hydrogen:

$$\frac{1}{\lambda} = R\left(\frac{1}{n^2} - \frac{1}{m^2}\right) \quad \text{The Rydberg Formula}$$

In this formula, λ is one of the many wavelengths of the light emitted by hydrogen, n is an integer greater than or equal to 1, and m is another integer greater than n. Each possible set of values of n and m correspond to one possible wavelength of emitted light. The coefficient R is the *Rydberg constant* whose value, approximately 1.1×10^7 inverse meters, was measured experimentally.

Once measured and inserted into the simple formula above, R could account for all the lines. So, for example, for $n = 1$ and $m = 2$, the formula predicted that hydrogen should emit light of a wavelength λ given by the formula for its inverse:

$$\frac{1}{\lambda} = 1.1 \times 10^7 \times \left(\frac{1}{1^2} - \frac{1}{2^2}\right) = 1.1 \times 10^7 \times \left(1 - \frac{1}{4}\right) = 1.1 \times 10^7 \times \frac{3}{4}$$

This corresponds to a wavelength of approximately 1.212×10^{-7} meters.

Bohr's Derivation of the Rydberg Constant

Physicists before Bohr knew the Rydberg formula for the spectral lines, but had no idea why it was right. In particular, they assumed that the Rydberg constant was a fundamental quantity whose value, 1.1×10^7, could be obtained only by measurement, much like the speed of light or the charge of an electron. Bohr's model treated the hydrogen atom as composite made out of an electron and a proton and calculated the quantized energy that corresponded to each orbit, as described in Chapter 5. He then derived the Rydberg formula from the differences in energies between planetary electron orbits in the model, identifying n and m in the formula with the numbers labelling discrete sequence of orbits. His model produced a formula for R itself that was a simple product of powers of the electron mass m, its charge e, the speed of light in a vacuum c, and Planck's constant h, given, to be specific, by

$$R = \frac{2me^4\pi^2}{h^3c} \quad \text{Bohr's Formula for the Rydberg}$$

The values of all the fundamental constants on the right hand side were known. The computed value of R on the left hand side agreed with its experimentally measured value. Bohr had shown that the mysterious Rydberg constant was not at all fundamental, but depended on other quantities already known; it was, in a sense, a derivative, a contingent quantity.

This discovery provoked a kind of Spinozan (anti-)Wonder. I write *anti* because Wonder is the feeling you get when contemplating something magnificently unconnected to everything else you understand; the calculation of the Rydberg constant provokes Wonder at the magnificent connection.

The Fine Structure Constant

Ever since, buoyed by this triumph, physicists deep in their hearts hope to find a theory that will explain the values of other apparently fundamental constants too. If R is a derivative quantity that can be explained in terms of m, e, c, h, then why shouldn't m, e, c, h and the gravitational coupling constant G, be explained in terms of other more primitive quantities?

Occasionally the community of physicists has become excited about a new formula, often discovered by chance or by mathematical tinkering, that seems to derive a value for some of the dimensionless constants of physics. Dimensionless constants are pure numbers that have no units, are not measurable in feet or seconds or kilograms. They are numbers with no scale. Perhaps the most famous is the *fine structure constant* α, called so by Sommerfeld because of the way it appears in an extension of Bohr's planetary atomic model that he developed in order to take account of the fact that an electron cannot move faster than the speed of light. Sommerfeld's formula was able to explain some of the finer details of the structure of atomic spectral lines.

The fine structure constant is defined as $\alpha = 2\pi\frac{e^2}{hc}$, and is dimensionless. Its observed value is about $\frac{1}{137.035999084}$, measured nowadays to an astonishing accuracy of more than ten significant figures.

Although it was first used by Sommerfeld to derive small corrections to the formulas for spectral lines, the fine structure constant is more important than that: it represents the strength of the electromagnetic force, just like G represents the strength of gravity. Since α is a pure number with no scale, it reminds physicists of the purely mathematical number π, and so they dream of finding a formula for its value from pure mathematics, as has been done with π

Here is what Feynman* had to say in 1985 about α:

> "There is a most profound and beautiful question associated with the observed coupling constant, *e*, the amplitude for a real electron to emit or absorb a real photon. It is a simple number that has been experimentally determined to be close to –0.08542455. (My physicist friends won't recognize this number, because they like to remember it as the inverse of its square: about 137.03597 with about an uncertainty of about 2 in the last decimal place. It has been a mystery ever since it was discovered more than fifty years ago, and all good theoretical physicists put this number up on their wall and worry about it.) Immediately you would like to know where this number for a coupling comes from: is it related to π or perhaps to the base of natural logarithms? Nobody knows. It's one of the greatest damn mysteries of physics: a magic number that comes to us with no understanding by man. You might say the "hand of God" wrote that number, and "we don't know how He pushed his pencil." We know what kind of a dance to do experimentally to measure this number very accurately, but we don't know what kind of dance to do on the computer to make this number come out, without putting it in secretly!"

* QED: The Strange Theory of Light and Matter, Princeton University Press, p. 129, ISBN 0691083886

Respectable physicists with mystical leanings (Wolfgang Pauli and Arthur Eddington among others) as well as psychologists (Carl Jung, who treated Pauli) have speculated on the significance of the value of the fine structure constant, and tried to find formulas for calculating α or $1/\alpha$. One simple and elegant formula, accurate to several decimal places is

$$\frac{1}{\alpha} \approx 4\pi^3 + \pi^2 + \pi \approx 137.0363037...$$

There are others that do even better.

When you then see a simple formula like this for α, your romantic heart leaps. Though many formulas of this kind have been derived, some more accurate and some less so, none of them have had proofs that physicists find convincing or acceptable. I remember several exciting mini-sensations in the physics community during the time I was a graduate student about near incomprehensible formal "derivations" of the value of α.

Faced with these failures, physicists over the past thirty years have adopted their own jiu-jitsu approach to "explain" the values of fundamental constants such as α: they invented the anthropic principle. The anthropic principle states that fundamental constants are what they are because if they were different we wouldn't be here to measure them. For example, some physicists argue that if α and/or various other fundamental constants had even slightly different values, our carbon-molecule-based life would not have been possible. The value of α we find in the universe, they claim, is what it is. If it weren't, we wouldn't be here to measure it. In some versions of string theory there is in fact room for many universes (multiverses, as they are fashionably called by physicists progressively excising any residual notion that we, our earth, our solar system, or our universe are central), each with their own values for the fundamental constants. As Nabokov wrote in *Lolita*, "You can always count on a murderer for a fancy prose style."

The anthropic principle disguises the failure of physicists to find a theory for the value of α. As with the EMH, there are "weak" and "strong" formulations of the anthropic principle too, but these are nitpicking distinctions.

This seems to me more metaphysics than physics.

Very recently, some astronomical experiments have very tentatively indicated that α varies ever so slightly across the visible universe. If that were so – if the strength of the electromagnetic interaction is different in different places – the fundamental constant wouldn't be constant, and ambitious idealistic physicists would try to find a formula for its variation.

UNDERSTANDING UNCERTAINTY, RISK AND REWARD

Quantifiable and Unquantifiable Risk

The great stumbling block in finance is the uncertain future. Uncertainty implies risk; risk means danger; danger means the possibility of loss.

We all abhor uncertainty; hence we have turned to divination, tarot, bone throwing, tea leaves, astrology and finally, science. Despite that, a known future is the exception rather than the rule. Even in classical physics, two simple pendulums connected to each other display a chaotic behavior that leads to a proliferation of unpredictable, widely differing possible futures from almost identical initial situations. Since we can never know the initial situation with complete accuracy, long-term prediction is impossible. Nothing is certain. Everything is merely more or less uncertain. A life of certainty wouldn't be life at all.

In economics, thoughtful people have come to distinguish between *quantifiable* and *unquantifiable* uncertainty. Uncertainty that is unquantifiable is clearly, for example, the likelihood of a revolution in China, the probability of the terrorist detonation of a nuclear bomb in midtown Manhattan, the chance that one of my children will fall in love with someone from Greece. All of these events are more or less unlikely, but no one knows what their odds are, and no one knows a reliable way of estimating them. Similarly unquantifiable is the likelihood of there being intelligent life in the universe (including Earth), or the chance that an earthquake of magnitude 6.7 or greater will occur before the year 2030 in the San Francisco Bay Area[6].

All of these events are low-probability outliers. When someone says there is a one in a million chance of a terrorist attack, it's no more than a guess that can never be substantiated. Estimating the likelihood of intelligent life in the universe depends upon a model for how planets and life develop. Calculating the probability of a large earthquake requires a model of the earth's crust and its motion. There is no way of knowing of how accurate such models are, and hence no way of truly estimating the probabilities. The best you can do with unquantifiable uncertainty is to be aware of it, and aware of your inability to quantify it, and then to act accordingly. As Freedman and Stark[6] argue, common sense is best:

6. See *What Is The Chance Of An Earthquake,* D.A. Freedman and P.B. Stark, Department of Statistics University of California Berkeley, Technical Report 611, January 2003.

> Another large earthquake in the San Francisco Bay Area is inevitable, and imminent in geologic time. Probabilities are a distraction. Instead of making forecasts, the USGS could help to improve building codes and to plan the government's response to the next large earthquake. Bay Area residents should take reasonable precautions, including bracing and bolting their homes as well as securing water heaters, bookcases, and other heavy objects. They should keep first aid supplies, water, and food on hand. They should largely ignore the USGS probability forecast

Nevertheless, in some idealized cases, uncertainty is (nearly) quantifiable. It's traditional to refer to *quantifiable* uncertainty as *risk*. A classic example is the uncertainty involved in tossing an unbiased coin: will it come up heads or tails? For a unbiased coin, the probability of heads or tails is 1/2, which means that, if you toss the coin enough times, the ratio of heads to tails will approach unity ever more closely as the number of coin tosses increases. Similarly, one can determine the probability for three successive heads followed by two successive tails to be $(1/2)^5$, or 1/32. Probability defined in this way is called a frequentist probability.

For an immaculate coin, perfectly balanced and coming up heads or tails with equal probability, the uncertainty is quantifiable. There is, however, no coin like that, just as there is no real-world Euclidean point or line. It's a fiction. There isn't even a perfect computer-simulated coin, since random number generators on digital computers, necessarily produced by an algorithm, are by definition not truly random. As a result, there is no truly quantifiable uncertainty. A real coin is inevitably slightly biased. It experiences variable forces with each successive flick of the thumb, variable gusts of moving air as it spins, variable irregularities of the floor as it bounces. The chain of these multiple small variations on the trajectory of the coin, concatenated, make every detail of the environment so important that, counterintuitively and again jiu-jitsu like, the details become unimportant *over many tosses*. Consequently, the coin is effectively uncoupled from the environment. If everything affects you so much then, in a sense, nothing affects you very much. The coin's trajectory between initial flip and final rest is not worth observing, and the history of each toss has no effect on the subsequent toss. Therefore, frequentist probabilities provide a good description of the uncertainty, and a large number of tosses will produce pretty close to a 50% probability of heads. The uncertainty of a physical coin is close to quantifiable.

The cornerstone of meaningful frequentist probabilities is the ability for a system to undergo independent identical situations, over and over again. For coins, history is unimportant. In economics, politics and human affairs, history matters, because people are altered by every experience. The first occurrence alters the response to subsequent occurrences in a complex way that makes impossible the repetition of identical trials needed to determine a frequentist probability. To take a topical example, credit markets after the great financial crisis of 2007-2008 will not behave like credit markets before, despite a lowering of interest rates by the Federal Reserve Bank, because we cannot forget our losses. Only Mnemosyne or Alzheimer can restore us.

Human institutions display a metaphorical hysteresis, a physical phenomenon in which the state of a system is path-dependent: each next move depends on the entire history of the system, not merely where it is at present. The history of the world doesn't affect a coin toss, but it *does* have a bearing on the next move in a stock's price. The uncertainty in a stock price is qualitatively different from the uncertainty of a coin flip. The likelihood of a stock market crash differs in nature from the likelihood of throwing three heads and two tails in succession, because market crashes are crowd events. Crowds aren't coins. Crowds remember the last crash and fear the next. A coin doesn't fear a sequence of three heads and two tails, and isn't affected by its nearest-neighbor coins in your pocket.

Despite this, the Efficient Market Model assumes that all uncertainties about the future are quantifiable as frequentist probabilities. This is not strictly true, of course, which is why the Efficient Market Model is a model of a possible world rather than a theory about the one we live in.

In physics, effects propagate locally in time and space and the future cannot affect the present. In the social sciences, the imagined far-away future can affect the present, and hence affect the actual future too.

The Rationale for Risk is Expected Return

Some risks are unavoidable by virtue of being alive.

Exposing oneself to the risk of the stock market is strictly voluntary. The key is always: Does the (perceived) possible reward outweigh the (perceived) risk of loss? I have stressed the word *perceived*. Financial models, unlike those of physics, are primed by current perceptions about the *future* of the world rather than its *cur-*

rent state. Stock prices and interest rates are quoted perceptions about the future of aspects of the economy. And it humans who are do the perceiving.

What does it mean to say an investment is risky? I will define the efficient market model's definition of risk more precisely in the next section. Meanwhile, it is sufficient to point out that the risk of a security is measured by the range of possible returns it will provide to an investor who buys it. Which of these return will actually occur is unknown, and depends on many future events both economic and social. For lack of a better word, one says it is up to chance. But it's not up to chance in the strictly frequentist sense.

Faced with a range of uncertain rewards, people simplify and focus on the *average reward*. The average *percentage* increase in the value of an investment one can expect to get over the range of all possible returns is called the *expected return*.

The Efficient Market Model assumes that investors care about only two properties of a stock: its *expected return* and its perceived *risk*. No matter what business the company is in, no matter who runs it, no matter how the company is organized, the model assumes that all this subsidiary information is reflected and incorporated in just two quantities, the (perceived) *expected risk* and *expected return*.

AN ASIDE: BEHAVIORAL FINANCE

Behavioral finance is a relatively new area of economics that overlays psychology on finance. The field focusses on the emotional biases that affect investors' ability to make rational judgements.

You don't have to look at yourself or your friends for very long to accept many of the conclusions of behavioral finance. Investors *are* obviously poor at thinking in strictly logical terms; they *are* slow to give up preconceived notions and prejudices; they *are* often overconfident about their skills and judgements; they can reach different conclusions from the same data depending upon how the data is presented to them; they tend to take larger risks when they are losing in the hope of recouping their losses, and smaller risks when they are winning in order to safeguard their gains; they are willing to overpay for small odds of large successes, which they regards as lottery tickets. There is good empirical evidence to confirm all of these observations.

Interesting and valid though many of these points are, they do not amount to a *comprehensive* theory of valuation. Behavioral finance is less a unified theory or model than it is a collection of interesting facts and habits. It does have the advantage of being empirical, documenting how investors actually behave rather than how one imagines they should behave.

It is worth remarking that all financial models are at bottom behavioral models, either more or less naive. The Efficient Market Model makes extreme assumptions about investors' *behavior* and works out their consequences. In this extreme narrowness lies both its strength and weakness, accounting for the generality of the results that will follow from here on, and also for their disagreements with reality.

THE LAY OF THE LAND BEFORE US

I am now about to embark on the best explanation I can devise for the logic of the Efficient Market Model for stock valuation. My explanation isn't the one usually taught; I invented it myself and I think it gives a transparent idea of the assumptions and logic behind the model, so that I can pick them apart. Let me therefore give you a look ahead at the rest of this chapter so you can understand where I'm heading.

- In "How The EMH Models Risk" on page 30 below I explain the Efficient Market Model's description of the risk of a stock. The model assumes, somewhat simplistically, that the price of a risky stock resembles the path of a ship steered by a helmsman with a hand tremor. Suppose the helmsman tries to guide the ship north by north west. Because his hand shakes, sometimes he steers a little too far to the west and sometimes a little too far to the north. Though the ship never goes exactly north by north west at any instant, it does on average indeed sail in that direction.

  Analogously, the stock price grows at a steady average rate, but, instant to instant, it sometimes grows more rapidly and sometimes less so.

  The essential risk is the momentary uncertainty about the return over the next instant: will it be greater or less than the average?
- In **The Fundamental Principle of Quantitative Finance** on page 39 I therefore pose the essential question of quantitative finance: *What average return should you expect to earn in the future for exposing yourself to a given amount of risk today?*
- Answering this question demands a principle. Physics has many successful theories, among them Newton's laws and Maxwell's equations. In finance, in contrast, there is in the end only one reliable principle. One version of it can be stated as follows: *One should expect the same return from any two securities with the same risk.*
  What else can one reasonably expect?
- There is a uniquely privileged security in the financial world, the security that bears no risk at all. If you buy it, you receive a guaranteed return at a guaranteed interest rate, the *riskless rate*. A good example is (used to be?) a loan to the U.S.

Government; you buy a Treasury bill, and three months later, for example, you get your principal back and a guaranteed predetermined rate of interest. The U.S. Government will certainly[7] pay back the lender. Given that there is a security with a guaranteed return and (correspondingly) zero risk, it serves as a baseline, and all risk and all return is naturally measured relative to it.

Finance aficionados call the amount by which the return of a stock exceeds the riskless guaranteed return of a Treasury bill the stock's *excess return*.

If you believe the fundamental logical principle that all stocks with the same risk should provide the same expected return, you can prove the following result:

The expected excess return of any individual stock must be proportional to the risk to which it exposes you.

Here we have the essence of the psychology of the capitalist system: take on more risk, expect more return.

- Stocks don't exist in isolation. Usually there is a universe of stocks that are each potential sources of return. In aggregate, they comprise *the market*, a larger entity which has its own net risk. I will show that when you apply the Efficient Market Model to a market consisting of many stocks, an even stronger condition holds:

  The expected excess return of any individual stock must be proportional not to its own risk, but only to the amount of market risk it delivers.

I will demonstrate all of these results below, and discuss how likely they are to be true.

HOW THE EMH MODELS RISK

In this chapter I will use a share of a company's stock as the prototypical risky security. Other securities – bonds, currencies, commodities, mortgages, real

7. Well, almost certainly. Sometimes governments do default on their obligations, but since the U.S. can print its own *fiat* currency, this is very improbable. Still, the possibility of revolutions, wars, natural disasters and other acts of violence could in principle make even this loan slightly risky.

estate, etc. – are also risky, but a share of stock will illustrate adequately the fundamental quality of unpredictable returns.

As I described, a share is not an indivisible atomic entity. A company is comprised of people, property, physical equipment and mental ideas; it involves thought and extension, is distributed in space and time, and the outcome of its activities is uncertain. If you like physics metaphors, you can think of the it as an immensely complicated molecule; if you prefer biology, think of it as a community of organisms. How on earth is one going to model the risk of a share of its stock?

The Efficient Market Model's answer to this question is radical: ignore complexity![8]

The Invisible Hand of the Invisible Helmsman

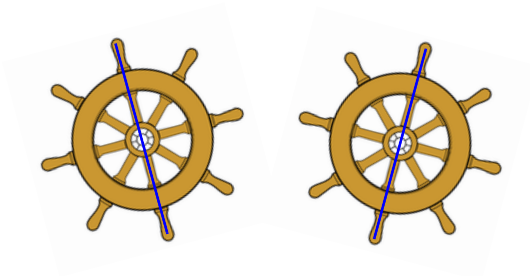
At any instant in an efficient market the current price of a share of stock is assumed to reflect all publicly available knowledge about the financial future of the company, and so the price of the share equals the expected value of the investment.

The next change in the stock price will arise from new information which is equally likely to be good or bad.

The Efficient Market Model assumes that all changes in the stock price are random, but random a very regular, controlled, limited way, described below. Nothing too wild is allowed to happen.

8. When I say ignore complexity, I do not mean that you shouldn't analyze the internals of the company in detail. I mean that, after all your analysis, you must distill your knowledge into two numbers, the company's risk and its expected return. In the Efficient Market Model, nothing else matters.

FIGURE 6.2. The helmsman with the trembling hand.



More specifically, as time passes, you can expect a stock price to appreciate at some expected rate, *but in an uncertain way*. Part of the time it will grow more rapidly, part of the time less so. Think of the stock's price as a steam ship, guided by an aged helmsman (Figure 6.2) who has a Parkinsonian tremble; he points the steering wheel in a certain direction, but since his hand shakes uncontrollably, sometimes the ship goes a little far to port, sometimes a little far to starboard.

In keeping with the maritime analogy, finance afiçionados call the average rate at which the stock price increases its *drift*. Like all physicists manqué, finance afiçionados love Greek symbols. They use the symbol μ (pronounced 'mew') to denote the drift, defined as the average return the stock is expected to earn over time. The fluctuation in the instantaneous drift is called the *volatility* of the stock, denoted by the Greek letter σ (sigma). The volatility causes the stock to grow randomly, sometimes by several percentage points more than the drift, and sometimes by less.

Use these terms to refer to stocks and you will sound very sophisticated.

DRIFT μ AND VOLATILITY σ: AN EXAMPLE

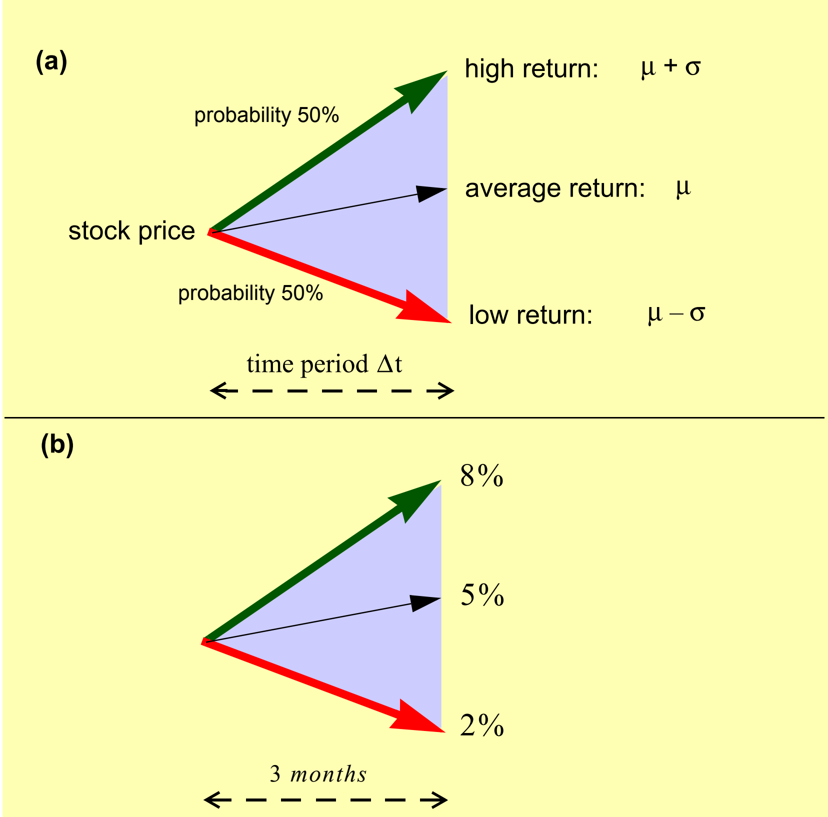
Suppose that $\mu = 5\%$ and $\sigma = 3\%$. This means that the stock price is expected to grow at an average rate of 5% per year with a volatility of 3%, meaning you can expect random up-or-down oscillations of ± 3 percentage points about the drift of 5%. For some short periods, therefore, the stock price will grow at an annual rate of 8%, and for other short periods at a rate of 2%. The risk is that you will not know which rate is about to prevail.

Another more volatile stock could have the same drift of 5% with a volatility of 10%, so that sometimes its price may grow at 15%, but at other times at -5%, so that part of the time its price will decrease rather than grow.

Figure 6.3 illustrates how the stock price is allowed to change in the model over a very short period of time[9] Δt. Since the Efficient Market Model postulates that price equals value, only significant news can change the price of stock. So, think of Δt as the number of milliseconds between the arrival of bits of (positive or negative) news about the economy, the company, or its market. If good news has just arrived during time Δt (the helmsman's hand trembles towards the right), the company's value will grow at a rate denoted by $\mu + \sigma$. If bad news has arrived (the helmsman's hand trembles towards the left), the value will grow at a rate $\mu - \sigma$. On average, it is expected to grow at the rate μ, and the size of the allowed fluctuations about μ are restricted to magnitude σ.

9. In mathematics the Greek symbol Δ (Delta) before any symbol indicates an infinitesimally small change in the quantity represented by the symbol.

FIGURE 6.3(a) A simple model for a stock's returns during an instant of time Δt. The stock has equal probabilities of returning more or less than the average μ because of its volatility σ. (b) A simplified numerical example assuming Δt = 3 months, with μ = 5% and σ = 3% during that time.

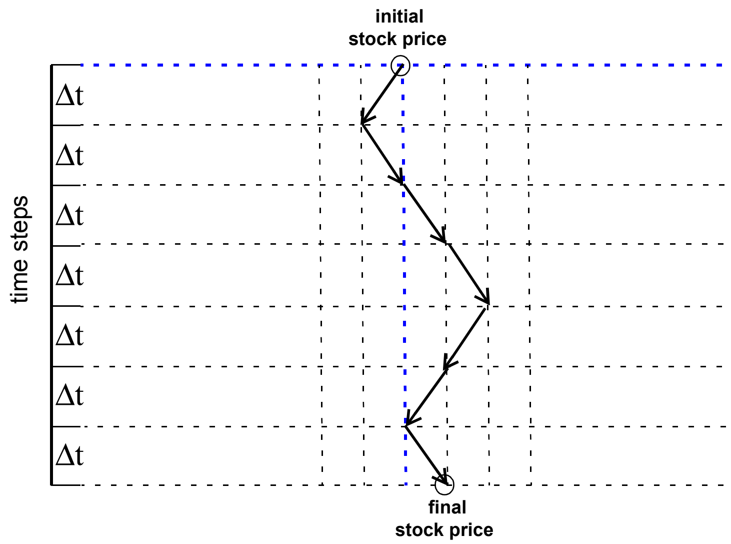


How Risk Increases With Time: The Square Root Rule

If you expect a stock to earn 8% in one year, then, all circumstances being equal, you will naturally expect it to earn 16% over two years. Expected return grows proportional to the time you hold the investment. Counterintuitively, that is not the case with risk.

To comprehend this, look at Figure 6.4, which shows a stock price that has undergone seven random up or down moves in (imagined) correspondence with the arrival of bits of good or bad news.

FIGURE 6.4An illustration of seven successive random moves in the stock price. Notice that after seven moves the stock price has moved only one net unit from its initial value.



With news arriving randomly, the stock price is unlikely to have moved in the same direction all seven times; up (higher than average) moves are as likely as down (lower-than-average) ones, and so the ups and down will mostly cancel each other, though imperfectly, because they occur at random and don't follow each other in perfect alternation.

When up and down moves occur at random, the *net* number of moves up or down after cancellations will be smaller than the total number of moves. Clearly the net risk is much less than seven times the risk at each step.

It's not hard to show that, *on average*, after n random up or down moves, the magnitude of the *net* number of moves $(\text{up} - \text{down})$ is not n but rather $\sqrt{n}$. In other words, the net number of moves to the up or down side is the *square root* of the number of random moves that have taken place, and could occur to the up or down side. The risk lies in not knowing whether the net appreciation after cancellations is up or down.

Thus, after 100 items of news arrive the stock price has most likely drifted only about $\sqrt{100}$, i.e. 10 price steps away from its original value. And after 1000 time steps, assuming that news arrives at regular intervals, the average – positive or negative – distance from the stock price origin is about 32 steps. Similarly, after a time t elapses the stock price has increased or decreased by an amount proportional

PROVING THE SQUARE-ROOT RULE

Here is a sketch of the proof that the risk grows proportional to the square root of the time the investment is held.

Suppose you buy the stock and then each successive random up or down move in the stock price is $\pm 1\%$.

The path illustrated in Figure 6.4. represents successive moves of $\{1\%, 1\%, 1\%, 1\%, -1\% -1\%, 1\%\}$.

One of many other possible paths the price might have taken is $\{-1\%, -1\%, 1\%, -1\%, -1\%, 1\%, 1\%\}$.

It's not simple to figure out the effect of random up or down moves because the number of cancellations is itself random. But mathematicians are cagey and always look for something easier to do.

One can instead calculate instead *the square* of each move which, because $(-1)^2 = 1$, no matter whether the move was up or down. After *n* moves, therefore, the *square* of the percentage moves to one side or the other is always positive and equal to $n \times 1$, i.e. n, irrespective how which random path the stock price took.

Therefore, we know exactly and without doubt after *n* moves *the square* of the total percentage move, no matter how many moves were up or down. The average percentage move up or down is therefore the square root of this known number, the i.e. $\sqrt{n}\%$.

If each move in the price takes a fixed amount of time, then the total number of moves *n* since you bought the stock is proportional to the time *t* elapsed, and so the average percentage move up or down is proportional to $\sqrt{t}$.

to the square root of *t*. More volatile stocks suffer greater changes in price and less volatile stocks move a smaller amount, but whatever their volatility, the distance they move away from their initial price increases with the square root of the time elapsed.

This square-root growth of risk is important; though it seems counterintuitive, it soon becomes a part of every quant's visceral intuition. If the risk to your capital is 10% of its value, then, in this model, the risk over two years is not 20% but only about 14%.

The Random Walk As A Model

This type of price evolution goes by several names: a random walk, diffusion, or Brownian motion. It describes how smoke particles drift through a room as they collide by chance with air molecules.

For smoke particles this description of the motion is pretty close to a true theory; it's the way things actually happen. For stock prices, however, this is *only a model*. It's how we *choose* to imagine the changes in stock prices. It may not be – it isn't – what actually happens, though there are similarities. Nevertheless, let's continue to elaborate the consequences of thee assumptions.

We have assumed that the news that influences stock prices arrives in regular measured bursts. Now imagine that minute bits of news arrive almost continuously. Then stock prices undergo smaller changes, but much more frequently. Under those circumstances the changes in the model's stock price mimics the more-or-less continuous motion of actual stock prices in a reasonable though imperfect way. Figure 6.5 compares the paths of stock prices from the model with the price of JCG, the ticker the stock of J Crew Group, a United States clothing retailer, from late 2006 to late 2010. The apparently naive either-up-or-down model does *more or less* mimic the riskiness of a stock's price.

I say more or less but perhaps I should say *less*. The mimicry fails in detail because the stock paths in the model are too smooth when compared with the observed movements of actual stock prices. Figure 6.5c is less smooth than Figure 6.5b, especially if you look closely. But you don't need large amounts of data to reach this conclusion; there are obvious conceptual flaws in the model of price movements. The news that drives stock prices doesn't always arrive in small increments. Sometimes, something important happens suddenly and discontinuously, in a flash with little prior hint. This kind of event doesn't fit into the framework of the smooth random walk economists like to focus on. The demand for a stock can occasionally become so great that its price leaps up; more commonly, the market's panic to sell can be so intense and contagious that the prices of all stocks crash downwards, quite unlike the calmly random motion of diffusing smoke. The Efficient Market Model's price movements are too limited and elegant to reflect the actual market accurately.

Financial modelers are not stupid. One can easily make more complicated models of risky stock prices that incorporate violent stock price moves and violent changes in risk. But, in using such models one gives up an easily understandable simplicity in exchange for a much more epicyclic but still imperfect depiction. There is as yet no widely accepted model of stock price evolution, and certainly no hope of one that predicts the future. The movements of stock prices are more like the movements of humans than molecules.

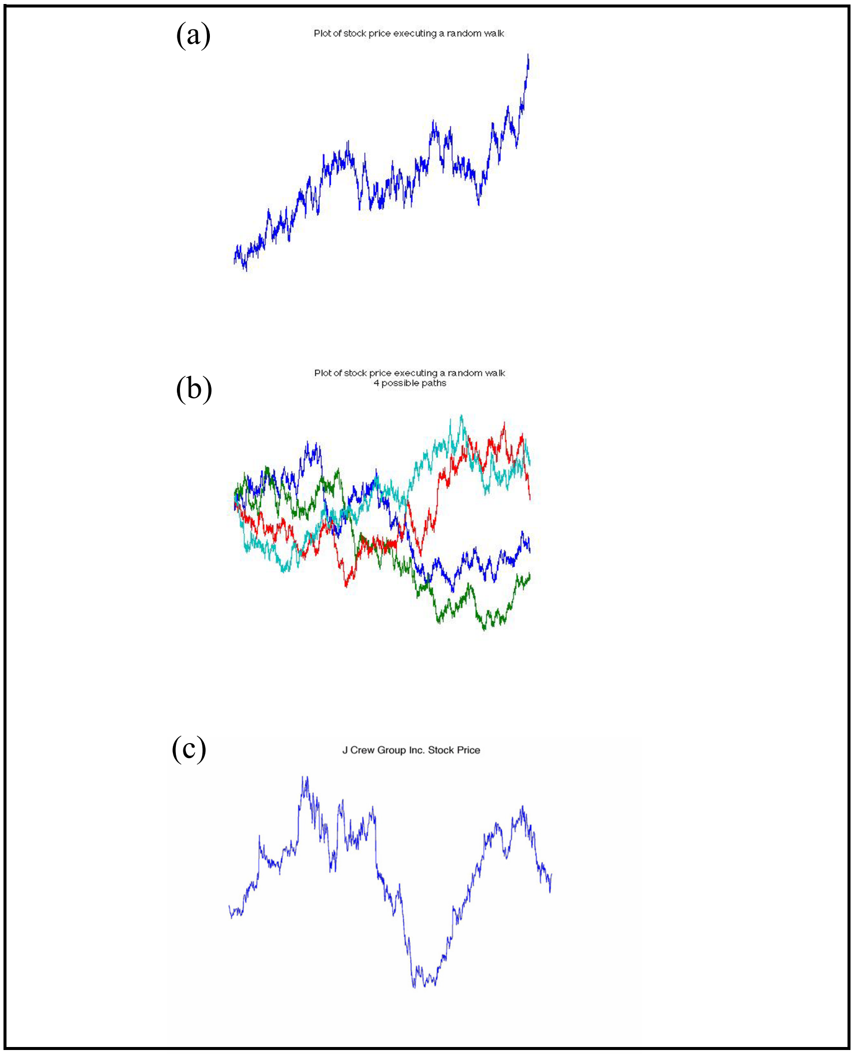
THE FUNDAMENTAL PRINCIPLE OF QUANTITATIVE FINANCE

Uncertainty Demands Higher Return

About the simplest and safest way to generate money from your money is to invest in a short-term three-month Treasury bill. You loan your money to the government and they promise to pay you a guaranteed rate of interest over the next three months and then return your principal. You can regard the act of loaning them money in return for the promised interest and repayment as *buying* a security from the U. S. Treasury. They pay you interest on the loan to recompense you for the loss of the use of your money, as well as to compensate you for the chance that something terrible will prevent them from repaying you. Since the chance of that is small – it is very unlikely the U. S. Treasury won't be around to repay what they owe when the loan matures – the investment is nearly riskless and consequently pays a low guaranteed rate of interest.

Putting money in a bank deposit account is a little more risky than lending it to the government, but still close to riskless. Lending money to a corporation (that is, buying a corporate bond) is riskier still. Corporations must promise to pay a rate of interest on a loan which is greater than the rate the Treasury offers, because, their promise is worth less. The more likely their perceived failure to pay, the more they must promise to pay.And in consequence, somewhat paradoxically, the more they must promise to pay the more likely that are to fail to do so.

FIGURE 6.5.(a) A single stock path simulated via a random walk. (b) Four typical simulated stock paths. (c) An actual four-year path for the price of J Crew Group Inc.



Suppose that, instead of lending your money to the government, you buy a share of stock. Stocks promise no guaranteed return at all. The company's earnings and the dividends it chooses to pass on to you as a stockholder will depend on fortune. If, after three months, you decide to offload the risk by selling the stock, you will have to take what the market is willing to pay for your share at that time. That will depend on the mood of investors, the state of the market, and the perceptions investors have about the company based on the three additional months of history. Buying a stock has a highly uncertain outcome.

Investors dislike uncertainty, and expected to be compensated for tolerating it – that's a fact of human psychology. Since the return of a stock is uncertain, whereas the return of a riskless bond is known in advance, investors will expect a greater return from a risky stock, *on average,* than from a riskless bond.

THE TIME AVERAGE VS. THE PORTFOLIO AVERAGE

Since future returns are uncertain, the fundamental question for finance is what *average* return should be expected for a given amount of risk.

Average can mean (at least) two different operations. The first kind of average is a *time* average: if you invest in one security and measure the returns year by year over many years, the time average is the average of the successive yearly returns. The second kind of average is a *portfolio* average: if you buy many different securities with the same risk, the portfolio average means the average of the returns in the same given year over all securities with the same risk.

Mostly, in finance, one assumes these two averages are equivalent. But the financial world is not particularly stable and so taking long-term averages for one security over time is a tricky ill-defined business: governments, administrators and regulators change, booms and busts come and go, people modify their behavior based on recent failures and successes, and so the patterns of returns and their statistical distributions tend to change with time.

In the long run, things may settle down, but no one knows how long the long run is. In contrast, in the physical world, as far as we know, laws remain invariant through time.

The Most Important Question: The Relation Between Risk and Return

The key question for any theory of finance is therefore:

> Exactly what average return should you expect to earn in the future for accepting a given amount of risk today?

Notice the question posed: what return should you *expect* to earn? Since you must transact today in order to achieve a return in the future, and since the future is uncertain, the question is about one's *expectations* of return rather than *realized* returns. Expectations about the future are a kind of model. When the realized returns differ wildly from the expected returns, the world awakes, startled at the inadequacy of the models behind the prices. One of the lessons of the credit crisis of 2007-2009 is that the models used to determine the value of mortgage securities involved future scenarios that didn't include those that actually came to fruition.

Answering the Question: Similar Securities Have Similar Prices

It takes a principle (not a theorem) to provide an answer., and there is, unfortunately, only one reliable principle in the theory of finance, the The Law of One Price:

> If you want to estimate the value of a security, use the price of another security that's as similar to it as possible.

The principle states that, in contrast to physics, there is no absolute knowledge in finance; instead, you must be a relativist and find an analogy. Here is how it works.

Given a "target" security whose value you want to estimate, you must find some other security (or portfolio of securities) that is "similar" to the target and, additionally, is liquid and fungible, so that its market price is easily obtained. Afiçionados call this portfolio *the replicating portfolio.* Given the "similarity", your best estimate for the value of the target security is the known price of the replicating portfolio.

What does it mean to say that two securities are "similar"?

In finance, two securities are similar when they have *the same future payoffs* under *all foreseeable future circumstances.*

A very naive example is the similarity between a \$100 bank account and a portfolio of two \$50 bank accounts at the same bank. Assume that the bank will pay you the same interest on all accounts[10]. Furthermore, in case of bankruptcy, both accounts will be covered by insurance provided by the Federal Deposit Insurance Corporation. If the FDIC itself has already run out of money, both accounts will suffer the same unfortunate fate. The similarity between the $100 account and the two $50 accounts is therefore close to perfect under all imaginable circumstances.

A more realistic example of a more limited similarity is (1) a loan of $100,000 to the U.S. Treasury paying 5% interest semi-annually for ten years and (2) a loan of the same amount to a friend starting a new business who also promises to pay you 5% semi-annually. These two securities will be equivalent as long as both parties can satisfy their obligations. But, in bad economic times, your friend will surely be more likely to default than the U.S. Treasury, and so there are foreseeable circumstances under which the future payoffs mismatch. The value of a loan to your friend is therefore worth less than the nominally equivalent Treasury bond, since there are some future scenarios under which the Treasury repays you but your friend doesn't.

Financial economists refer to their essential principle as the *Law of One Price*, or the principle of *No Riskless Arbitrage*, which more formally states:

> *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 definitely not a theorem; it's simply an assumption about the behavior of human beings that isn't gospel. It holds true in a rough way some of the time. 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. There are always short- or even longer-term exceptions that persist, usually for good reasons.

Using the Law of One Price to Value Apartments

Here's a simple financial model I invented to illustrate the methodology behind all valuation models.

---

10. These days they will likely pay you zero interest on an account this small, as well as charging you a monthly administrative fee that will rapidly reduce your balance to zero, but I ignore that for now.

Consider the problem of estimating the value of a grand eight-room penthouse apartment on Park Avenue in New York City. The most direct way is to put the apartment on the market and look at the offers. Next best is finding a recently sold similar apartment in the same neighborhood, and using the sale price as the approximate value of this apartment.

Suppose that there have been no sales of similarly grand apartments lately, and that you can obtain only the recent sale price of one of many fairly vanilla two-room apartments in Battery Park City. Suppose these apartments are liquid and change hands frequently. You will have to use the price of the Battery Park City unit and value the Park Avenue apartment relative to that.

You can approximately *replicate* the penthouse, in your imagination, by combining several Battery Park City units into one larger one the size of the penthouse. Let's suppose this requires about seven Battery Park City apartments. Then your first estimate is that Park Avenue penthouse should be worth seven times the Battery Park City apartment.

This model assumes that the price per square foot is constant, irrespective of apartment details. Price per square foot is what economists call an *implied* price per square foot. It's "implied" in the sense that it's not the actual cost per square foot of construction, but rather the price per square foot implied by quoting market prices of entire apartments in terms of price per square foot. It's a *model* because it assumes that price per square foot is all that matters.

That's not true. First, in New York City, very large apartments are scarce and sell at a premium; second, a Park Avenue location is more exclusive[11] and hence desirable; third, there are more staff and doormen per capita in a Park Avenue cooperative building, fourth …

The price of the replicating portfolio consisting of Battery Park City apartments is therefore an underestimate. With some practical experience you can make rule-of-thumb nonlinear corrections for size, adding an implied-price-per-square-foot skew so that the larger the apartment the greater the fair price per square foot.

11. "Exclusive" means desirable and expensive because it excludes more people, a good thing from a financial point of view. In a financial crisis, though, exclusivity means illiquidity. What you want to own in a widespread financial crunch are *inclusive* securities.

You can make similar corrections for location, park views, natural light and facilities to correct upwards the initial estimate of a factor of seven in price.

This is the way most practical financial models work. You compare what you don't know with what you do. You use a simple powerful assumption to get off the ground. Then you make common-sense heuristic adjustments. In the apartment model, I assumed that price per square foot is constant across apartments, which does capture a major part of the variation in apartment prices. Then I added corrections for other important factors that lie outside the model. Models for pricing stocks, bonds and options work quite similarly.

DERIVING THE RELATION BETWEEN RISK AND RETURN

A Treasury bond is nearly riskless, though less so than it used to be. Though the interest and principal in dollars will almost certainly be paid, the value of those dollars – the goods they will be able to buy in the future – is uncertain. Those dollars will likely buy only an unknown fraction of what they buy today.

Yet any other investment involves even more risk. Faced with a range of risky returns rather than one certain one, investors demand that the average of the risky returns be *greater* than the return they could earn by taking no risk. Otherwise, why take risk? This *risk-aversion* seems to be a fact of human nature.

You must think of *risk* as a fundamental quality that investors have to acquire in order to have a chance of profiting. The one natural question then is: *how much excess return should one expect for taking on risk greater than that of a riskless bond?*

To answer, I intend to use the Law of One Price, that two securities with the same future payoffs under all foreseeable circumstances must have the same current value. In the Efficient Market Model, everything foreseeable about a stock, and in particular the range of its future payoffs, is determined by its expected return μ and the *volatility* σ that represents its risk. All that differentiates one stock from another in the model of Figure 6.3 are the values of μ and σ. Nothing else matters. Two securities with the same μ and σ are in essence identical. The Law of One Price then effectively states:

> Any two securities with the same foreseeable *risk* should have the same *expected return*.[12]

In principle, if all one can know about a security is its volatility σ, that should determine everything about its return.

There is a shallow resemblance between finance's Law of One Price and Einstein's Principle of Relativity in physics. Relativity claims (1) that all observers, irrespective of their relative speeds, should, from their observations, deduce the

---

12. Attempting to find a law on which to base financial theory, I called this statement an invariance principle in my paper *The perception of time, risk and return during periods of speculation*, Quantitative Finance 2(4), August 2002, pp. 282-296.

same equations for the laws of nature, and (2) that light is special: nothing can move faster than it. The Law of One Price claims (1) that all securities with identical risk, irrespective of their composition, should provide the same expected return, and (2) that riskless bonds are special; nothing can have less risk than them.

There Is Only One Kind of Risk

All securities with the same risk must have the same return. But how can this principle tell us *how much* return?

In the Efficient Market Model there is only *one* kind of risk, volatility. All volatilities of any magnitude can be created from a single security with volatility via financial engineering. So, if you know the expected return for any one security, you can figure out the expected return for all risky securities.

Here is how it works.

- By financial engineering: You can dilute a more risky security with a riskless security to create a combination with less risk.
- By the Law of One Price, the expected return of the less risky security must be the same as the expected return of the more risky security that has been appropriately diluted.
- Suppose you know the expected return of the more risky security.
- Then you can calculate the expected return of a diluted version.
- Therefore, you know the expected return of the less risky security.
- Therefore, once you know the expected return that goes with any one level of riskiness, you know the expected return for any level of riskiness.

You Can Synthesize Variable Amounts of Risk

How do you synthesize a less risky security from a more risky one?

Imagine you have as ingredients

(i) a stock with volatility σ, and

(ii) a riskless bond with zero volatility.

Their possible returns over a short instant of time are displayed schematically in Figure 6.6.

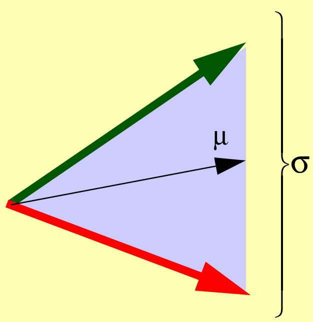
FIGURE 6.6. A risky amount of stock worth \$1.00 and a riskless bond worth \$1.00 can be combined to create a portfolio with an intermediate risk. By combining w cents worth of the risky stock with $\$1.00 - w$ cents worth of the riskless bond, where w is an amount between zero and \$1.00, you can create a \$1.00 portfolio with lower volatility $w\sigma$ and correspondingly lower return $r + w(\mu - r)$. Thus, adding volatility $w\sigma$ to a riskless bond increases the expected return over and above the return r of a riskless bond by $w(\mu - r)$. In other worlds, w cents risk gets you w units of excess return. Since this is true for one security, by the law of one price, it must be true for all securities.

(a) **A risky stock**[a]

risk: σ

expected return: μ.

The high return differs from the low return by an amount σ

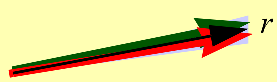


(b) **A riskless bond**

risk: 0

return: r.

The high and low returns are equal.

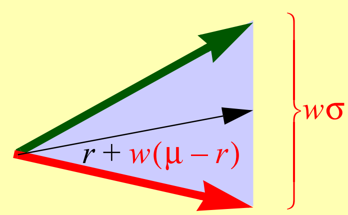


(c) **The mixture: a new security**

$w \times \text{Stock} + (\$1.00 - w) \times \text{Bond}$.

intermediate risk: $w\sigma$.

$$\text{expected return} = w\mu + (1-w)r \equiv r + w(\mu - r)$$



a. Green arrows denote returns higher than the mean; red lines denote returns lower than the mean; black lines denote the mean return. The distance between the high and low is related to the risk σ

Figure 6.6a displays the returns of a stock with risk σ; the high return is greater than the low return by an amount proportional to σ. That the high and low returns differ is precisely what one means by risk in the random walk model.

Figure 6.6b displays the returns of a riskless bond; its "high" and "low" returns are in fact identical. The bond is riskless.

Figure 6.6c shows the result of diluting the risky stock with the riskless bond by creating a mixture: its ingredients are the risky stock and the riskless bond. Specifically, the mixture contains w cents of investment in the stock and $\$1.00 - w$ invested in the riskless bond, so that the total mixture portfolio is worth $1.00.

Figure 6.6c shows the risk and return of the mixture, obtained by simply adding the expected returns of w cents in the stock and $(\$1.00 - w)$ cents in the bond. The new security, the portfolio consisting of the mixture, is riskier than the bond and less risky than the stock.

In Figure 6.6c you can see that mixture has a volatility equal to $w\sigma$, with a corresponding expected return $r + w(\mu - r)$ Focus on the red characters in these expressions, and you will see: adding an *extra* $w\sigma$ of risk *to a riskless bond* wins you an *extra* $w(\mu - r)$ *of expected return*.

By invoking the law of one price, namely, that securities with the same risk must provide the same expected return, one can conclude that *any security with risk* $w\sigma$ *must provide extra return* $w(\mu - r)$.

The Proportionality of Risk and Return

Figure 7 illustrates an example in which a risky stock provides an expected return of 7% with a volatility of 20%, and the expected return of a riskless bond is 2%. By choosing a weight $w = 0.75$ of the stock and then combining it with a fraction 0.25 of the riskless bond, the volatility of the mixture is reduced to 15%. The return of the mixture is $(0.75)7\% + (0.25)2\% = 5.75\%$. By the law of one price, one must expect an average return of 5.75% for *all* securities with a volatility of 15%..

More generally, $w\sigma$ of excess risk must win you $w(\mu - r)$. of excess return. Both contributions are proportional to w, the amount of risky stock in the mixture. Thus, the law of one price leads to following principle:

> **The Proportionality of Risk and Return**: For any stock, the expected return above the riskless rate must be proportional to its risk σ.

Thus, no matter how risky a stock, the ratio of its excess return to its risk is always the same.

FIGURE 6.7. See Figure 6.6. A risky stock, volatility 20% and expected return 7%, and a riskless bond returning 2% can be combined to create a portfolio with volatility 15% by combining 75 cents of risky stock with 25 cents of riskless bond. Since this portfolio has a weighted expected return equal to 5.75%, all securities with volatility 15% must have the same expected return 5.75%

(a) **A risky stock**

volatility: 20%

expected return: 7%

The high return differs from the low return by 20 percentage points.

(b) **A riskless bond**

risk: 0

return: 2%.

The high and low returns are equal.

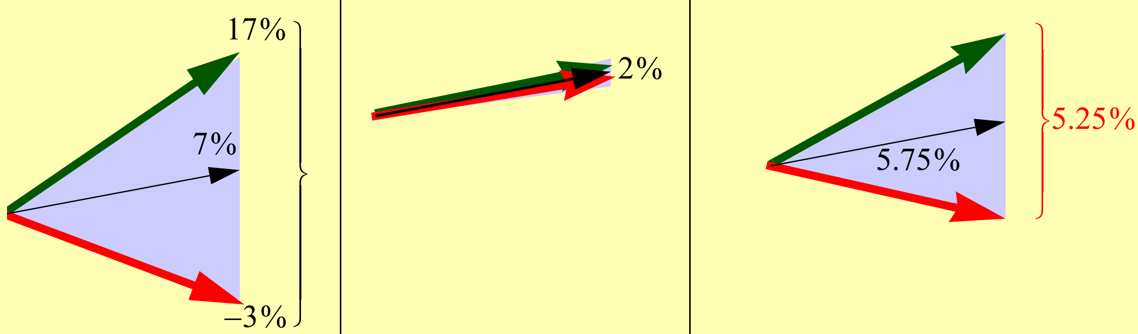
(c) **The mixture: a new security with** $w = 0.75$

$0.25 \times \text{Stock} + 0.75 \times \text{Bond}$.

intermediate risk:

$w\sigma = 0.75\sigma = 15\%$.

$$\text{expected return} = 0.75(7)\% + 0.25(2)\% = 5.75\%$$



The Sharpe Ratio

Finance theorists like to write this law of proportionality entirely in Greek symbols to make it seem oracular:

$$\frac{\text{excess return}}{\text{risk}} \equiv \frac{\mu - r}{\sigma} = \lambda \qquad \text{The Sharpe Ratio}$$

This states that the ratio of excess return to risk is a "universal" constant λ, the Greek letter '*l*', pronounced lambda. The number λ is called the Sharpe Ratio, after William Sharpe who first began to make use of this concept in the mid-1960s.

The Sharpe ratio measures the bang you hope to get for your buck of risk, and is therefore also sometimes called *the risk premium*.

We have derived the proportionality of risk and return from the assumption that equal risks must generate equal expected returns. It follows that taking greater risks should automatically win you greater rewards, on average. You don't have to be smart to take greater risks; in the model, any greater risk will do. The interesting question is: How do you tell if you are being smart?

The test is the value of your Sharpe ratio. If you can find portfolios of securities whose Sharpe ratios are greater than usual, then you have been smart. The theory of efficient markets has become so much a part of accepted market lore that professional money managers measure and report their Sharpe ratios with the hope of demonstrating their talent.

In the Efficient Market Model the *one* unknown is the Sharpe ratio. If you know that, you know everything. There is only one unknown because there is only one kind of risk. The type of business the stock represents, the people who run it, the variety of its endeavors, all genuine matters of import, are considered not to matter – or rather, they do matter, but only in the way they affect the value of the volatility of the company. The major naivete of this approach is that it simplifies and condenses the notion of risk into just one quantity, σ.

For coin flips, the chance of any particular combination of heads and tails *is* accurately expressible via a simple mathematical formula. For companies and businesses, risk has too many dimensions to be accurately defined by one number. The volatility of returns, σ, is too simplistic to measure the true risk of investments.

Nothing we have said so far can tell us what the value of λ is. Let me ignore that for now and continue to derive more consequences of the model.

The Importance of Sharpe Ratios

Here's an illustration using *realized returns*. Suppose that over the past ten years security A has produced an average annual return $\mu_A = 15\%$, with high returns corresponding to 20% and low returns corresponding to 10%, so that the annual volatility σ_A is of order 5% (in either direction) about the mean. Suppose also that the riskless rate r for one-year Treasury bills during this time was 5% per annum. Then the excess return for security A was $\mu_A - r = 15\% - 5\% = 10\%$ annually. The annual Sharpe ratio was $\frac{\mu_A - r}{\sigma_A} = \frac{10}{5} = 2$. This means you have earned two units of excess return for every one unit of risk each year. Suppose security B had earned an average 20% per year over the same period, with lows of 5% and highs of 35%. Then the excess return for B was $\mu_B - r = 20\% - 5\% = 15$, which *looks* better than the excess return of 10% for A. But … the volatility σ_B of B, approximately half the range between high and low returns, was $\frac{1}{2}(35\% - 5\%) = 15\%$. The Sharpe ratio for B was therefore $\frac{\mu_B - r}{\sigma_B} = \frac{15}{15} = 1$. B looks better in terms of returns, but A is better in terms of return per unit of risk taken, which, *if there is only kind of risk*, is what matters. If there is only one kind of risk, you are better off exposing yourself to it through security A than through security B.

AN ANALOGY: THE PLEASURE PREMIUM

Risk is a fundamental quality, something you can selectively choose to expose yourself to. It's a primitive, an atomic property that cannot be reduced to something simpler.

The Sharpe ratio λ which tells you how much reward you should expect from a given chunk of risk reflects human psychology and the physiology beneath it. Humans may demand more or less return from a given amount of risk at different times, depending on how they feel. If they demand more return from the same risk, the Sharpe ratio will rise and stock prices will fall; if they are willing to accept less return, λ will fall and prices will rise.

Risk is a biological primitive, something that people Desire. You could compare desire for risk to desire for *deliciousness* in food. Like risk, the *deliciousness* of food, the *pleasure* it causes, is an independent quality, distinct from caloric and nutritional value. Chefs in restaurants can charge more for greater expected *deliciousness*. If you wanted to be foolishly quantitative, you could define *the pleasure premium* as the amount of dollars diners are willing to pay, *over and above the bare cost of the ingredients themselves*, for deliciousness. This pleasure premium is analogous to the risk premium. You can make the following analogy:

Risk Of A Security	**Deliciousness Of A Meal**
Expected return on security: μ	Cost of meal: C
Return of riskless bond: r	Cost of ingredients: I
Excess return: $\mu - r$	Excess cost: $C - I$
Risk σ	Deliciousness[a]: Ω
Risk premium: $\frac{\mu - r}{\sigma}$ is the excess return per unit of risk.	Pleasure premium: $\frac{C - I}{\Omega}$ is the excess cost per unit of deliciousness

a. I use the Greek letter Omega to quantify deliciousness. Pronounce it 'Ohhhhh!mega'.

The pleasure premium will vary from day to day; in flush times diners may be willing to pay more for expected deliciousness; in bad times they may avoid deliciousness and stick to nutrition. If you built a quantitative model, the pleasure premium Ω could play a large part in determining luxury food prices.

Just as delicacies sell on the basis of expected deliciousness, so securities sell on the basis of expected return. But there is a difference between pre-prandial expected deliciousness and post-prandial realized deliciousness. Similarly, expected return differs from realized return (though expected pleasure is often itself a pleasure).

What is the right value for the Ω, the excess cost per unit of expected deliciousness? No theory can dictate what a sybarite should pay for pleasure. The best one can do is to say that, all things being equal, there should be a constant price per unit of pleasure. Similarly, no theory can dictate what return an investor should expect in exchange for taking on risk; that too depends on appetite and varies with time. The best one can do is to say that, all things being equal, the amount of expected return per unit of risk should be the same for all securities.

I am of course simplifying: there is patently more than one kind of deliciousness. There is sweetness and tartness, spiciness and blandness, smoothness and lumpiness, all of them different types of gustatory pleasure. And I have ignored other kinds of mentionable and unmentionable bodily pleasures. All of these have their pleasure premiums too. Similarly, there is more than one kind of risk and one kind of risk premium; there is stock risk and bond risk and currency risk and commodity risk and slope-of-the-yield-curve risk; and within the universe of stocks there is sector risk – health risk, technology risk, consumer durables risk …

In physics, the fundamental constants (the gravitational constant G, the electric charge e, Planck's constant h, the speed of light c) are apparently timeless and universal. It seems to me inconceivable that there will ever be a universal value for the risk premium, despite the advances of neuroscience.

MORE REFINED RELATIONS BETWEEN RISK AND RETURN

The Strategy of Replication

Before we proceed, here is a summary of the logical foundations of the Efficient Market Model:

- Financial securities must satisfy a common-sense invariance principle: equal risks must provide equal expected returns.
- You can use financial engineering to replicate a synthetic low-risk security by diluting a high-risk security with a zero-risk security.
- By the invariance principle, the synthetic low risk security must have the same expected return as a genuine low-risk security.
- This leads to the "law" of proportionality of risk and expected return.

The invariance principle resembles a law, more science or psychology than engineering. The use of dilution to replicate any less risky security from a riskier one is engineering rather than science. Replication lies beneath everything generally useful in finance, which is why I call finance a theory of relative rather than absolute value.

Finance theorists have discovered more sophisticated ways to replicate risk that we will now exploit in order to develop extensions of the more-risk-more-return Sharpe formula. Two additional methods of replication are *leverage* and *diversification*. What follows from using them are practical refinements that result in *The Capital Asset Pricing Model*, a much more comprehensive (but still imperfect) approximation to markets. The remainder of this section is devoted to these refinements.

You Can Increase Risk and Return With Leverage

Increasing risk via leverage is easy; it's the most common strategy for generating hoped-for profits, and perhaps the last refuge of desperate scoundrels who have lost their investors' money and increase risk in order to hope to make it back. You simply borrow money to fund your security purchases, amplifying your incentives by means of Archimedes claim to be able to move the entire earth with a long enough rigid bar.

This is how it works in finance.

Suppose you purchase a share of stock worth \$100 with your own funds. This doubles your money if the stock price doubles, and causes you to lose everything if it becomes valueless.

Suppose instead you provide \$50 of the purchase price from your own funds and borrow the other \$50 to buy the same stock. In that case you can double your money sooner. If the stock price increases merely by 50% to \$150, you can sell it, use \$50 to pay off the loan and pocket \$100, double your initial investment. Conversely though, you will lose everything you invested if the stock drops by 50%; the \$50 it's worth then is precisely what you borrowed and must repay. If it drops even further and you unwind the trade, you will have to pay back more than you invested.

A levered position – a portfolio consisting of a loan and the stock it facilitates buying – is more volatile than the stock alone. Leverage magnifies the intrinsic risk of the purchased security, and thereby, the intrinsic rewards. If you hope or expect to earn an excess return of 1% a year on a trade, but you would like to earn 10%, you can borrow nine times as much money as you invest yourself. That will convert a 1% return on the stock into 10%. Hedge funds and investment banks do exactly that to try to earn sizeable profits out of small perceived edges, and it's wonderful when it works. But, as Long Term Capital Management demonstrated dramatically in 1998, and as happens more mundanely every day in the investment world, leverage magnifies the possible punishment too. A relatively small 10% decline in the value of the stock can wipe you out completely.

If You Can Diversify, The Risk Premium Is Zero

Diversification is an old idea. Some wishful readers of the Bible have interpreted Ecclesiastes 11:1 as an investment guide that exhorts us to *Give a portion to seven, and also to eight; for thou knowest not what evil shall be upon the earth*. Though you could imagine this as advice to spread your investments over many diverse endeavors, it's misguided to imagine that Ecclesiastes gave investment advice. His commandment seems to me to be an exhortation to be charitable in spirit and deed, to be kind to one's fellow-sufferers in the face of an uncertain future.

More to the point, in *The Merchant of Venice*, Salarino says to Antonio:

I know, Antonio
Is sad to think upon his merchandise.

And Antonio replies:

> Believe me, no: I thank my fortune for it,
> My ventures are not in one bottom trusted,
> Nor to one place; nor is my whole estate
> Upon the fortune of this present year:
> Therefore my merchandise makes me not sad.

Antonio's strategy is diversification: spreading his fortune over many dissimilar securities rather than just one or two. Diversification is more intricate than simple dilution, and therefore not as reliable. By accumulating a portfolio of many uncorrelated securities – gold stocks, pharmaceuticals, utilities … – you can hope that some go up in price while others go down, and some to do nothing at all, because much of the time each sector and even each stock is sensitive to different economic conditions. In theory, if moves in the prices of individual stocks tend to cancel each other, the price of a diversified portfolio of many such stocks will hardly fluctuate at all. If that's the case, the portfolio will become riskless as the number of securities comprising it becomes large, and its total volatility σ_P will approach zero.

Since a completely diversified portfolio is riskless, it is "similar" to a riskless bond, and therefore, by the law of one price, is expected to earn the riskless rate: its net excess return is by definition zero. But if its net excess return is zero, and yet at the same time its excess return is the sum of the excess returns of its individual constituents, then *each* of the excess returns of the individual constituents must itself be zero.

This is puzzling. In the model, any excess return, for a security or a portfolio, is equal to its volatility multiplied by the Sharpe ration λ. Subject to this rule, how can the excess return of a portfolio of constituent securities be zero while each of the constituent's excess returns is λ times its nonzero volatility? The only solution is that the proportionality constant, the Sharpe ratio λ itself, be zero! *If λ is zero then each individual security's expected excess return must be zero too.*

Setting the value of λ to be zero in the definition of the Sharpe ratio on page 50, we see that for any stock,

$$\mu = r.$$

The expected return is just the riskless rate, a very strong result.

We've shown that if (i) a market is efficient, and (ii) the market allows for *full diversification*, then investors should expect to earn only the riskless return on *any* individual stock. That is, they should expect to do no better by buying a risky stock than by buying a riskless bond.

This is true if perfect diversification is possible. Markets don't quite allow for that, which brings me to *hedging*.

If You Can Hedge *and* Diversify …

Markets are not amenable to true diversification.

For diversification to be possible, movements in individual stock prices must be unrelated to each other. But securities are not mechanical systems that can be engineered to be disconnected. On the contrary, it is especially easy to see that correlation will rise during a financial crisis, when the demand for all stocks will decline as investors flee risk of all kinds. In a crisis all stock prices sink together as investors rush to the perceived safety of cash, government bonds or gold bullion.

Even in normal times, stocks tend to move in tandem, most of them going up when "the market" goes up and going down when "the market" goes down. By "the market" I mean the portfolio consisting of every single stock in the market. In practice, stock indexes that measure the average price of some large collection of stocks – the S&P 500 or the Wilshire 5000, for example – serve as reasonable approximations to the entire market. The reason for the comovement of all stock prices is Keynes's now clichéd but nevertheless accurate observation that it takes animal spirits to take risk. Investors pile into stocks with unison when the economy appears good, and divest in fear when conditions worsen. Full diversification is impossible, and to be more realistic we must reflect this herd tendency in our model of stock risk.

Because stocks move together, a portfolio of many different stocks cannot carry zero risk. But, if you think about it carefully, you can picture the change in an individual stock price as consisting of two components. One part of the price change will be linked to a change in the *entire* market, common to all stocks; the remaining part will be peculiar or *idiosyncratic* to the individual stock itself.

Figure 6.8a illustrates this model for price changes using the right-angled triangle of high-school geometry to represent the risk of a stock with volatility σ. The red hypotenuse of the *risk triangle* – the longest side, representing the total risk of the stock – corresponds to a change of σ% in the stock price. Sticking with the Greek symbols, the (blue) lower side of the triangle represents the fraction ρ (Greek letter *rho*) of the percentage change in stock price that is linked to the market's move. The (green) right-hand-side of the triangle, perpendicular to the blue side, represents the remaining fraction $\sqrt{1-\rho^2}$ of the idiosyncratic change in this stock price, unrelated to the price movements of all the other stocks in the market. The blue and green sides of the triangle are perpendicular to each other because, on average, the two sources of price changes have nothing to do with each other. Pythagoras's famous theorem relates the sides of the triangle to the hypotenuse. Note that risk triangles represent *percentage* changes in price, in other words the *returns* of the stock.

FIGURE 6.8. Risk Triangles (a) Percentage changes in price of an individual stock with volatility σ. A typical stock move is σ%. The correlation ρ represents the fraction of the total change, ranging between 0 and 1, that is linked to the market's change. (b) The triangle for a stock with total volatility $\sigma = 20\%$ and correlation $\rho = 0.6$.

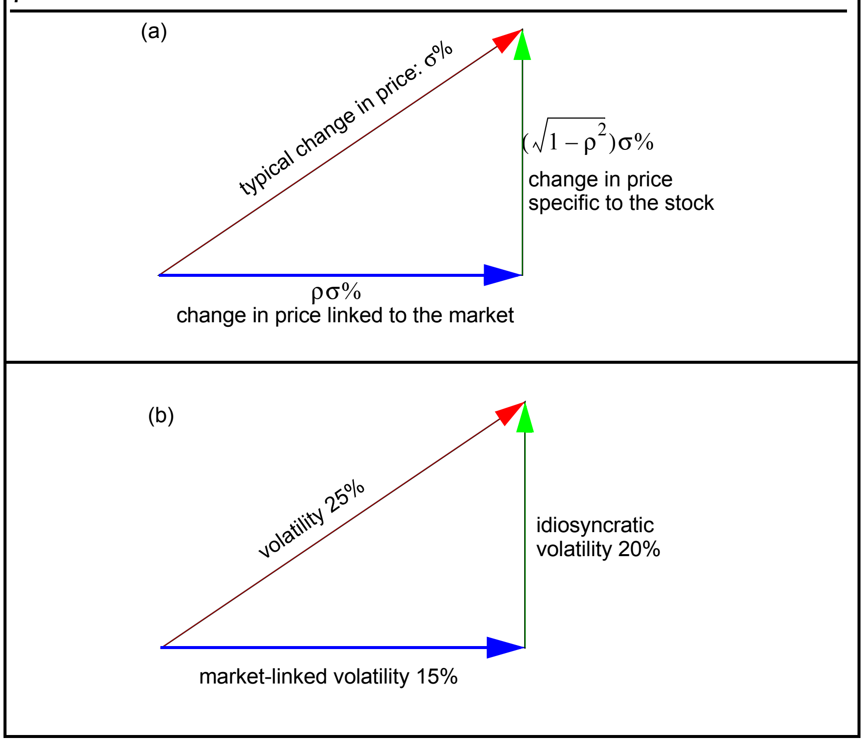


Figure 6.8b illustrates the case for a stock whose volatility is 25% and whose correlation with the market is $3/5 = 0.6$. The corresponding return linked to the market is $\rho\sigma = 15\%$, and the idiosyncratic return, is $\sqrt{1-\rho^2}\sigma = 0.8 \times 25\% = 20\%$.

This so-called correlation ρ is a convenient statistical way of describing components of return. Its value must lie between -1 and 1. The price of a stock with $\rho = 1$ always moves in the same direction as the market at every instant; the price of a stock with $\rho = 0$ moves on average independently of the market; the price of a stock with $\rho = -1$ always moves opposite to the rest of the market.

The correlation ρ is a crude empirical statistic rather than a fundamental quality; it's more like an ERA (Earned Run Average) in baseball, extracted from data, something that can change with time, something you notice and *measure* rather than a deep fundamental quality of a pitcher. The correlation depends at any given time on how market participants react to economic news that affects the market and the stock; it's a function of behavior.

Given that part of stock's return is market-linked and part idiosyncratic, let's see what happens when to portfolio of many stocks.

When all stock prices move, one part of each stock's return will correspond to its idiosyncratic risk, the green line in Figure 6.8. Every stock has its own independent green line.

But another part of each stock's return is linked to the risk of the market, the blue line that is the same blue line for all stocks. When you put together any portfolio of stocks, the idiosyncratic returns are uncorrelated and therefore diversifiable. However, the market-linked blue price changes of every single stock move together with the market. Thus, no matter how large and diverse you make your portfolio, the blue market risk will still be there.

To get rid of that, you have to *hedge the market*, that is, to invest in some other security that behaves antithetically to the market, goes up when the market goes down.

To do that, you have to *go short*.

Going Short: The Creation of the Antistock

If you own a share of stock, you profit when its price increases. What should you do if you think a stock is going to decline?

You could simply not own it, and thereby have no exposure to its price at all. But that's just staying unengaged. What if you want to stop being a neutral? Then you need the *antistock*, which behaves exactly like the thing itself *with all of its qualities reversed.* The antistock goes down in price when the stock goes up, and vice versa.

Much like a positron is created by temporarily borrowing an electron from the Dirac sea, an antistock is created by borrowing a stock from a broker who holds one and then selling it. You have sold something you borrowed, and will have to return it when asked. This transaction – borrowing something and selling it temporarily – is called *going short*, or *shorting.* You own a hole in the sea of stocks.

Going short is the analog of the invention of negative numbers in arithmetic. When you short a stock, it's as though you own an antistock $\bar{S}$, a mirror-reflected version of the stock whose value is $-S$., and goes down in price when the stock goes up.

Shorting a stock is contemptuous *a la* Spinoza, an intent to profit from what the stock lacks. No wonder then that so many corporate CEOs malign and fear 'shorts', those speculators who seek to profit from downturns in the fortunes of companies.

The financial crisis of 2007-2008 left many hypocrisies exposed. Investment banks run profitable businesses by coolly loaning their customers' stocks to hedge funds who want to short them in order to speculate on price declines. But during the crisis years of 2007 and 2008 these same banks and funds were only too happy to solicit the SEC to prohibit the shorting of their company's stock, for fear it would drive their own company's stock price so low it would threaten their existence.

The Market-Neutral Stock

Imagine you own a stock S which has both its own idiosyncratic risk and a risk linked to the market portfolio M. Before you can diversify over all stocks, you need to neutralize each stock's component of market risk. The trick is as follows.

How To Go Short

Going short allows you to profit when the stock price decreases:

- You borrow the stock from a broker who has it in inventory and promise to return it at some future date. If the stock is initially worth S_1 dollars, you now owe the broker that amount.
- To compensate the broker for making the loan you pay a small service fee.
- Then you immediately turn around and sell the borrowed stock for S_1 dollars and pocket the cash (less the fee). Ignoring the small fee to keep things simple, you owe S_1 dollars to the broker and you have S_1 dollars in cash, which you hold on to.
- At some later date you buy back the stock in the marketplace at a new price S_2, and give it back to the broker from whom you borrowed it. The transaction is over.
- If the stock has declined in price since the day you borrowed it, that is, if $S_2 < S_1$, after paying for the purchase you are left with $S_1 - S_2$ dollars, a *profit*.

  Had you simply bought the stock at an initial price S_1 and sold it at a lower price S_2, you would have *lost* $S_1 - S_2$ dollars.

Thus, shorting the stock produces exactly the opposite result from owning it. A short position profits by as much as a long position in the stock loses.

- For each stock *S*, combine it with just enough antimarket $\overline{M}$ (which moves opposite to the market) so that the combination has no market risk at all. In other words, for each stock *S*, create a portfolio S' that contains one share of the stock *S* and[13] β shares of the antimarket:

$$S' = 1 \times S + \beta \times \overline{M} \qquad \text{Market-Neutral Stock}$$

In terms of M, this is equivalent to $S' = 1 \times S - \beta \times M$.

This portfolio S' is a little basket that contains the stock and the antimarket, the amount of antimarket $\overline{M}$ carefully chosen to neutralize the market-linked risk of S. You can think of it as a synthetic version of the stock S that I call *the market-neutral stock*, containing only the residual part of the stock that is idiosyncratic, insensitive to the market. It's the part of the stock in Figure 6.8 that corresponds to the green line.

How many shares β of antimarket $\overline{M}$ are necessary to remove the market risk of the stock?

The Value of the Hedge Ratio Beta

Figure 6.9 shows the risk triangles for the returns on the stock S, the market M, the antimarket $\overline{M}$ and the market-neutral stock $S' = S + \beta\overline{M}$. The coefficient β is the number of shares of the antimarket $\overline{M}$ to be combined with one share of the stock S and has been chosen to be $\beta = \rho\frac{\sigma}{\sigma_M}$. This value is the ratio between the length of the heavy blue line in Figure 6.9a and the light blue line in Figure 6.9b, i.e. the ratio of the market risk of the stock to the market risk of the market itself. This value for β uniquely eliminates all the market risk of the market-neutral stock S' by cancelling the stock's own exposure to the market with that of the market itself.

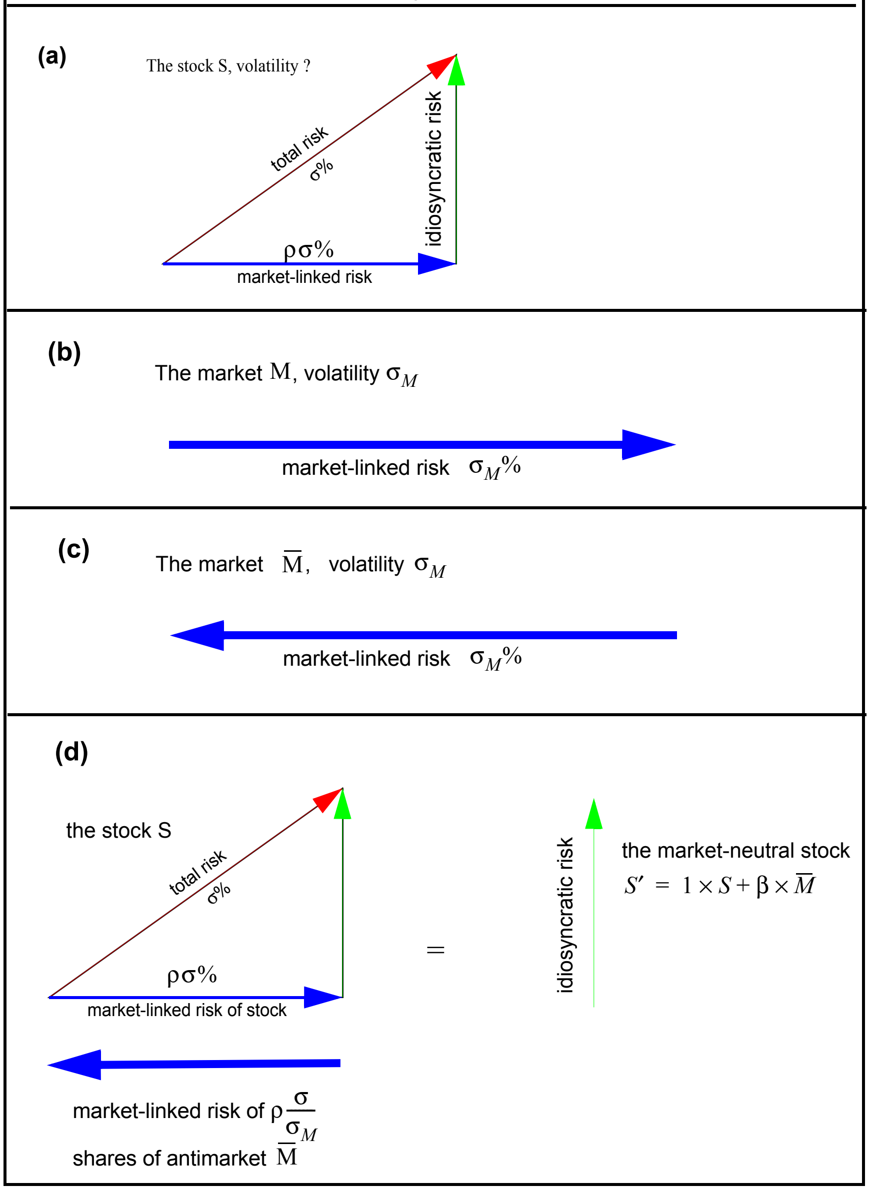
The Expected Return On the Market-Neutral Stock

Now let's figure out the expected return on the market-neutral stock S'. To be specific, assume we have a single dollar's worth of the stock S and a single dollar's worth of the market M, and we use these investments set up a hedged position $S' = 1 \times S + \beta \times \overline{M}$ that has no market risk. We can straightforwardly calculate the return on the basket S' from the returns of its constituents:

- The net cost of setting up the position $S' = 1 \times S + \beta \times \overline{M}$ is $1 - \beta$ dollars.
- What can you *expect* to earn on this initial outlay of $1 - \beta$ dollars, given that the expected return on the stock is μ and the expected return on the market is μ_M?

13.A Greek letter as usual, beta.

FIGURE 6.9. **(a) Risk triangle for a stock *S*. (b) "Triangle" for the market M collapses to a line, because its total risk *is* its market risk. (c) "Triangle" for the antimarket $\overline{\mathrm{M}}$, whose risk is opposite to that of M. (d) Combined triangles of the stock S and $\rho\frac{\sigma}{\sigma_M}$ shares of the antimarket $\overline{\mathrm{M}}$. The market risks cancel. The resultant "triangle" for the market-neutral stock S' is shown at right, and collapses to a line because its total risk is its idiosyncratic risk.**



- The \$1 invested in the stock S has an expected profit of μ dollars.
- A position of β dollars in M has an expected profit of $\beta\mu_M$ dollars. The equivalent position in $\overline{M}$ therefore has an expected loss of $\beta\mu_M$ dollars.
- Therefore the net expected profit on S' is $\mu - \beta\mu_M$ dollars on an initial outlay of $1 - \beta$ dollars.

- The ratio of expected profit to purchase price produces the expected return

$$\frac{\mu - \beta\mu_M}{1 - \beta} \qquad \text{Expected return of a market-neutral stock}$$

… You Get The Capital Asset Pricing Model

We are almost there.

Earlier in this chapter, on page 57, I showed that if all stocks in the market are uncorrelated, you can cancel their risk by diversification. In that case, *any* stock's expected return μ must equal the riskless rate r.

Efficient Market proponents like to look at this from a moral point of view. When you buy an individual stock that is uncorrelated, you are exposing yourself to a risk that could in principle be mitigated. You cannot expect to be rewarded with a higher rate of return for refusing to protect yourself. You can earn excess return only on *unavoidable* risk.

Similarly, if all stocks are in fact correlated with the market, you cannot expect to be rewarded for not hedging market risk. This leads to the following argument.

- In actual markets, all stocks are correlated.
- As a result, even a diversified portfolio of stocks carries market risk.
- However, each market-neutral stock, as illustrated in Figure 6.9d, bears only idiosyncratic uncorrelated risk.
- Now n proceed as before. Assemble (in theory) a large diversified portfolio of market-neutral stocks. Their idiosyncratic risks are uncorrelated and stock-specific, and will cancel. The volatility of this diversified portfolio is close to zero.
- Therefore, the expected return from each stock's idiosyncratic risk that we calculated in the last bullet point on page 65 must equal the riskless rate. That is,

$$\frac{\mu - \beta\mu_M}{1 - \beta} \equiv r$$

This result can be rewritten by using simple algebra as

$$\underbrace{(\mu - r)}_{\text{stock's excess return}} = \underbrace{\beta}_{\text{stock's beta}}\underbrace{(\mu_M - r)}_{\text{market's excess return}}$$

Capital Asset Pricing Model

Though I haven't derived it in the same way as its discovers did many years ago, this is the famous Capital Asset Pricing Model, the triumph of so-called Modern Portfolio Theory developed by Treynor, Sharpe, Mossin and Lintner in the early 1960s. Finance afiçionados refer to the model via the affectionate moniker CAPM ("Cap Em").?

CAPM deals with *expected* returns. I have emphasized the word 'expect' to stress that it is a theorem about people's rational expectations, based on a model about how people *behave*. It is a relative-value model: it doesn't tell you what to expect absolutely, but only what to expect relative to something else. Contrast it with Maxwell's equations, which tell you exactly *how* light behaves, absolutely.

Utility companies that supply gas or electricity, something everyone needs to survive, animal spirits or no animal spirits. Utility stocks are stodgy and safe. Consolidated Edison (stock ticker ED), whose fondly remembered 1960s motto "Dig We Must!" described their disruptions of New Yorkers life as a utilitarian necessity, on average seems to suffer percentage price moves approximately one third as large as those of the market as a whole. It's relatively insensitive to the market, but not entirely so. The beta of ED is therefore about 0.3, and so (if you believe CAPM) you should expect only about one third of the market's excess return when you invest in it.

AAPL (Apple) a company that, in contrast, lives by inventing and creating demand for usefully vogue products you can survive without, depends on the high spirits of the public. Its percentage stock price moves are on average about the same as those of the market. Therefore, the beta of AAPL is about 1, and so, according to CAPM, you should hope for three times as much excess return from AAPL as you should from ED.

What CAPM Tells You

The left hand side of the equation, $\mu - r$, is the excess return you can *expect* to earn when you invest in some security S.

The right hand side contain $(\mu_M - r)$, the excess return you *expect* from by investing in the entire market M.

CAPM tells you that the expected excess return of a stock S should be the excess return of the entire market, *multiplied by the stock's beta*.

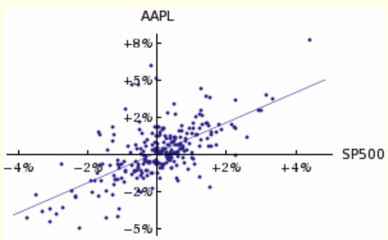
Beta is $(\rho\sigma)/\sigma_M$. Its numerator $\rho\sigma$ is the market-related risk of the stock. Its denominator σ_M is the market's intrinsic risk. Beta therefore measures the *relative market risk of the stock*.

CAPM says that the greater the relative market risk of a stock, the more you should proportionately expect to earn. All that matters for the return of a stock is its beta (β) to the market, the source of the *only undiversifiable risk*, and nothing else.

A TEST: APPLE AND THE S&P 500

Here is a plot of the prices of Apple (AAPL) and the S&P 500 (ticker ^GSPC) between Sep 13 2009 and Sep 13 2010. (All data and figures below are taken from www.wolframalpha.com.)

Here is a scatter plot of the daily return of Apple and the S&P 500 over the same period.



Let's see how well the Capital Asset Pricing Model worked for Apple stock during that year.

The mean volatility of the S&P 500 during that year was about 18%. During that period the S&P 500 returned 7.5%. If we include dividends earned by the stocks in the index, we can raise that number to about 11%. If we regard the S&P 500 as a reasonable proxy for the entire market M, then we can set $\sigma_M = 18\%$ and $\mu_M = 11\%$ in the CAPM formula.

Let's be generous and say that the annual riskless rate was about 1%, so that $r = 1\%$.

During the same period, Apple returned 55%, so we set $\mu = 55\%$. The mean volatility of Apple stock was about 27%, so we set $\sigma = 27\%$. The correlation of Apple's returns with the S&P was $\rho = 0.7$. Thus Apple's beta to the market was $\beta = \rho\frac{\sigma}{\sigma_M} = (0.7)\frac{27}{18} = 1.05$.

CAPM claims that $(\mu - r) = \beta(\mu_M - r)$. Strictly speaking, CAPM claims that this must hold for *expected* returns. But no one accurately knows what people expected a year ago, and so it is common to check the validity of CAPM for *realized* returns.

The value of the left hand side of the equation is $\mu - r = 55 - 1 = 54\%$.

The value of the right hand side of the equation is $\beta(\mu_M - r) = 1.05(11 - 1) = 10.5\%$

The value of the left hand side is about 44% greater per year than the right hand side. Apple returned that much more than CAPM demands. Afiçionados refer to the amount by which a stock's return exceeds the beta-inspired value on the right hand side as *alpha*, the Greek letter α.

A kind way to look at this is to say that CAPM holds over the long run, on average, with fluctuations from year to year and fluctuations from security to security. To test it, you should test it statistically, not for one stock at one time.

When in fact academics examine the historical returns of stocks, they find that stocks with higher beta don't always provide higher returns.

An unkinder way to look at it is to say that's it's not very good. Newton's law it ain't.

A still kinder way to look at it is to say that CAPM needs extending. Market risk is *not* the only non-diversifiable risk. Experienced investors are always busy trying to detect patterns in the universe of stock returns, patterns that they themselves, by their behavior, cause and, in their search for them, even accentuate. Investors perceive stocks as belonging to groups smaller than the entire market, groups whose constituent stocks tend to be bought or sold together by investors. The greatest group of co-movers is the entire market. There are smaller co-moving groups too. Large capitalization stocks form a group; so do "small caps". so do value and growth stocks. Computer stocks, for example, all tend to move together too, because certain kinds of news are particularly good (or bad) for computer companies as a whole, while not being as significant for, say, health care stocks. Each of these additional groupings and the factors that represent them can be hedged before diversification to lead to more complex version of CAPM.

CAPM's Infiltration

Most finance professors consider the Capital Asset Pricing Model to be one of the great achievements of financial theory. It creates a framework for thinking about markets which is now commonplace, not only for academics but for practitioners too. It also provided the base that led to the far more successful Black-Scholes options valuation model which is one the few genuine triumphs of financial engineering and quantitative finance.

I didn't fully understand how deeply CAPM had become embedded in the practical world of business until I recently tried to use Bloomberg, Yahoo and Google to find the volatilities and correlations of Apple and the S&P 500 in the box above. To my astonishment, there was no easy and direct way to find the correlation ρ between a stock's returns and those of the market, or to obtain a stock's volatility σ. What all those websites gave you easy access to was the stock's beta, the *product* of ρ and σ, *divided* by σ_M. It is a sign of the political power of models like CAPM that commercial web sites tools give you the value of beta, a parameter in the model, but not the more fundamental volatility or correlation statistics[14].

Despite this, CAPM has been important. Its most significant impact has been the introduction of alpha and beta. Metaphorically, beta represents the driver of the return that you earn for simply making the choice to go into the market, for simply deciding to commit, smartly or dumbly. Alpha, in contrast, represents skill, the return you earn by being smarter than the market as a whole. Inspired by CAPM, investors now ask themselves whether their manager is providing merely dumb beta, or smart alpha too.

Acquiring beta should be cheap, a commodity or a service; anyone can buy the entire market. Finding alpha takes skill and is worth paying for. Alpha is exercised by picking better than average stocks, or picking ordinary stocks at the right time. Money managers calculate their alpha to measure their performance. In that sense, they too have drunk the academics Kool-Aid.

Is CAPM True?

You can ask whether CAPM is true. I ask in return whether you believe that everyone is rational, whether you think everyone has the same information, and most importantly, whether you believe that the risk of a security is described purely by the volatility of its returns.

My answer is: "No, not really." Most significantly, the risk of a security is not just the stock's volatility. Volatility is exclusively important only if stock prices change according to geometric Brownian motion, and there is ample and undeniable evidence that they don't. Investors need to worry about more than the volatility of a stock; they need to worry about fear and greed and contagion and market memory and downward jumps in the stock price. Volatility alone is a poor measure of the really bad risks. Periodic market crashes testify to that regularly.

---

14. I have remarked on the marketing power of models before, in my book *My Life As A Quant*, (Wiley, 2004) pp. 198-9.
"I came to see that creating a successful financial model is not just a battle for finding the truth, but also a battle for the hearts and minds of the people who use it. The right model and the right concept, when they make thinking about value easier, can stick and take over the world. A firm whose clients start to rely on the results that its model generates can dominate the market. This is what happened with Salomon's concept of option-adjusted spread—a short while after they invented it, every other firm on the Street was writing their own version of the model to do the same analysis, because clients demanded it."

It's not surprising, then, that CAPM doesn't correctly account for the returns of investments. The tenets of efficient markets and CAPM hold better in undramatic liquid markets where informed investors who do careful analysis drive trading and the setting of prices. But, during times of panic, fear and limited liquidity, the assumptions fail. CAPM is a useful way of thinking about a model world that *is not* the world we live in, but may sometimes describe it more or less well. As Fischer Black wrote, markets are efficient when prices lie between one half and twice the "correct" value.

That's about as good as it gets. It is irresponsible to pretend things are better.

> Capable men live in a sort of despair over the fact that they are bound by the rules of their office to teach and communicate things which they look upon as useless and hurtful.

Maxims and Reflections: *Goethe*

OPTION PRICING AND BLACK SCHOLES: THE ORIGINAL DERIVATION

There are several ways to understand the derivation of the Black-Scholes model. Here is the original one.

Black was a great believer in equilibrium of markets. He derived the Black-Scholes equation in more or less the following way.

- A stock and its call *both* provide access to the underlying risk σ of the stock in the next instant.
- In equilibrium, the stock and the option should both provide equal bang for the buck, that is equal return per unit of risk. Because, the argument goes, if one provided more return per unit of risk than the other, savvy people would buy the cheap one and sell the rich one until they came into equilibrium. Or, put another way, you can dilute an option at any instant to create the risk of a stock, which you know, and hence you can find the expected return of the option.
- In other words, the stock and option should have same instantaneous Sharpe ratio:

$$\frac{\mu - r}{\sigma} = \frac{\mu_{Call} - r}{\sigma_{Call}}$$

- Simply from stochastic calculus for geometric Brownian motion, one can show that there is a relation between the stock's return and volatility and those of the option, because the option is a derivative of the stock. Hence

$$\mu_C = \frac{1}{C_t}\left\{\frac{\partial C_t}{\partial t} + \frac{\partial C_t}{\partial S}\mu_S S_t + \frac{1}{2}\frac{\partial^2 C_t}{\partial S^2}(\sigma_t S_t)^2\right\}$$

$$\sigma_C = \frac{1}{C_t}\left(\frac{\partial C_t}{\partial S}\sigma_t S_t\right)$$

- Combining the above two equations leads to

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2} = rC \qquad \text{Black-Scholes equation}$$

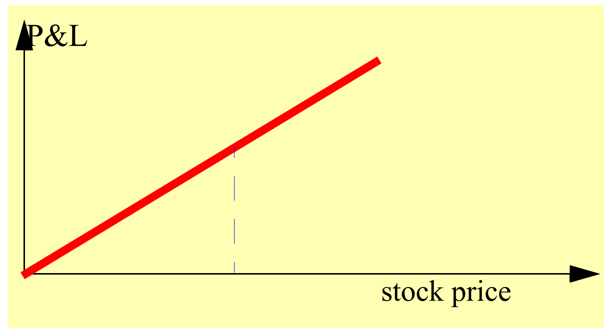
which is simply a statement that if two securities provide equal risk, they must provide equal return. You can solve for call price C given its payoff at expiration, and then find out how the call price C should vary with volatility, stock price, etc.

- Black derived the equation and then took several years to guess the solution and show that it satisfied the equation.

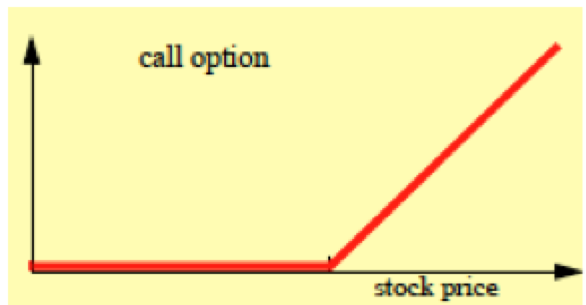
OPTION PRICING AND BLACK SCHOLES: THE SECOND DERIVATION IN THE STYLE OF MERTON

This derivation is based on creating an instantaneously riskless investment out of an option and a stock, and then, since it's riskless, it must replicate a riskless bond and its return, by the law of one price. Equivalently, then, the option can be instantaneously replaced by a portfolio of a riskless bond (short) and stock.

- Stock prices can go up or down.
- A stock's P&L is linear in the stock price

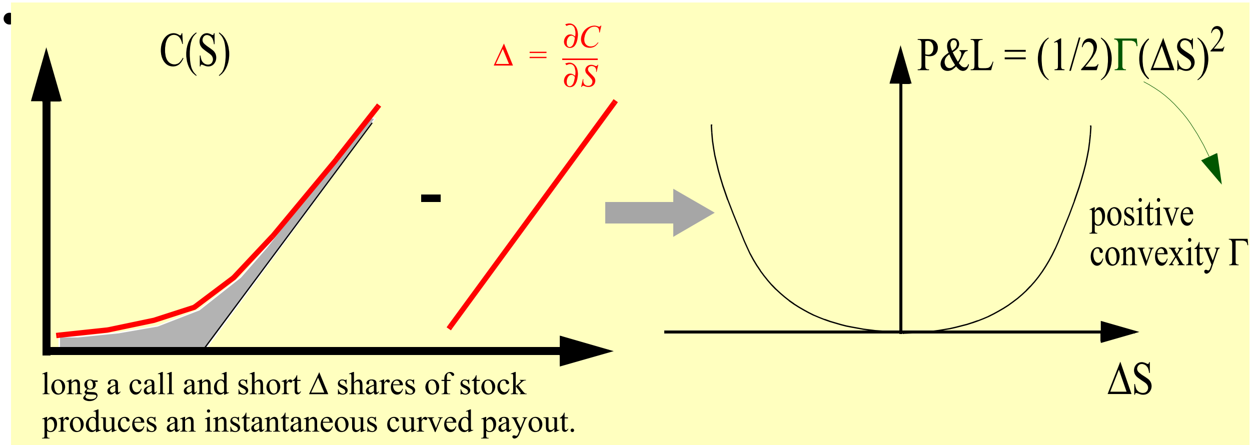


- A call option has only upside, no downside. It's payoff is *convex.*

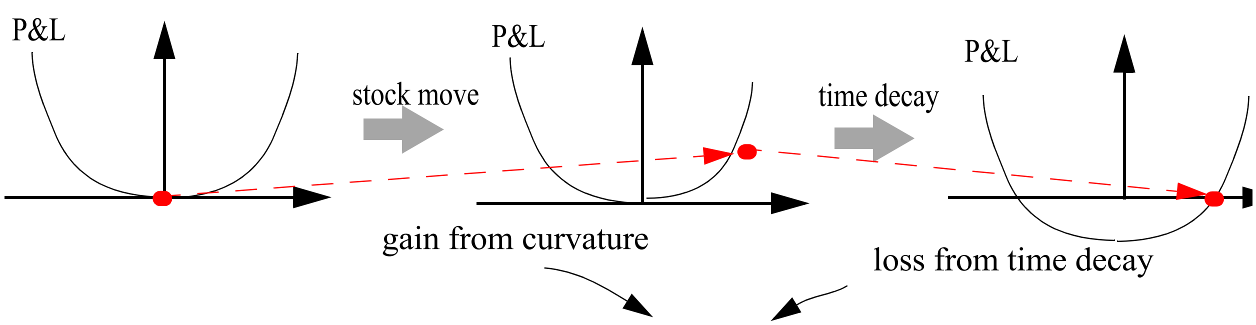


- What should you pay for downside protection, or for convexity?
- You can use **The Law of One Price** to tell you, by the following method:
- Buy the call (see figure below)
- Hedge away the stock risk by shorting $\Delta = \frac{\partial C}{\partial S}$ shares against it.
- That portfolio has a parabolic payoff over next instant that profits whether the stock moves up or down.
- The magnitude of the payoff depends on the curvature $\Gamma = \frac{\partial^2 C}{\partial S^2}$ on the magnitude of the move ΔS, i.e. the stock's volatility

•



- Suppose you think a stock will have volatility Σ : in time Δt you think it will move $\Delta S = \pm\Sigma S\Delta t$ where Σ is the volatility you expect for the stock. Then there are two ways the portfolio can have its value change in time.
- First, you expect to earn $\frac{1}{2}\Gamma(\Delta S)^2 = \frac{1}{2}\Gamma(\Sigma^2 S^2 \Delta t)$ from the stock's movement.
- But second, the value of the cal itself can decline as time passes, and so you expect to lose $\Theta\Delta t$ from *time decay*, where $\Theta = \frac{\partial C}{\partial t}$
- Since the portfolio is riskless, independent of stock moving up or down, it musst behave like as riskless bond (by the law of one price) and therefore earn the riskless rate *r*.
- To keep it simple, for now, set *r = 0*, Then this must happen:



$$\frac{1}{2}\Gamma\Sigma^2 S^2 \Delta t + \Theta = 0$$

- This is again the Black-Scholes equation, as before.

- It says that if you want to potentially gain from curvature, you must potentially lose from time decay. There is *no free lunch.*

$$\frac{\partial C}{\partial t} + \frac{1}{2}\Sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} = 0$$

time decay and curvature are linked

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\Sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} = rC$$

non-zero rates

Thus the Black-Scholes equation can be looked at from three different viewpoints:

1. Stocks and their options have equal Sharpe ratios.

2. Equivalently, the gain from curvature must be balanced by the loss from time decay.

3. Equivalently, you can replicate an option by creating a portfolio by borrowing money and buying the appropriate number of shares of stock, and then rebalancing the portfolio as time passes and the stock moves.

Note. The Black-Scholes model works much better than CAPM, even though it is based on CAPM. The reason is that two stocks are very different, riskwise, from each other, and so saying that their Sharpe ratios must be equal is kind of iffy. But an options and its underlying stock really are closely related, do bear similar risks (even if those risks are not exactly those of Geometric Brownian Motion) and so saying they should have equal Sharpe Ratios is more plausible.

A MODEL IS ONLY A MODEL

Theories describe the real world. Darwin's Theory of Evolution describes how life actually evolves and how species originate. Maxwell's Electromagnetic Theory succinctly represents actual electromagnetic fields. Dirac's hole theory is a theory of the electron as we know it. Even Freud's attempts to describe what he believes *is* the structure of the psyche. True or untrue, they all focus on actuality. CAPM, in contrast, is a model that makes strong assumptions about a possible world, and derives the consequences. That world could be ours, but it isn't.

You have to use models to estimate value in finance, but you have to understand that they are only models, not a theory and not the system itself. On January 7, 2009, in the midst of the financial crisis and the collapse of the market for collateralized mortgage obligations, Paul Wilmott and I published the manifesto below.

THE FINANCIAL MODELER'S MANIFESTO

by Emanuel Derman and Paul Wilmott

Preface

A spectre is haunting Markets – the spectre of illiquidity, frozen credit, and the failure of financial models.

Beginning with the 2007 collapse in subprime mortgages, financial markets have shifted to new regimes characterized by violent movements, epidemics of contagion from market to market, and almost unimaginable anomalies (who would have ever thought that swap spreads to Treasuries could go negative?). Familiar valuation models have become increasingly unreliable. Where is the risk manager that has not ascribed his losses to a once-in-a-century tsunami?

To this end, we have assembled in New York City and written the following manifesto.

Manifesto

In finance we study how to manage funds – from simple securities like dollars and yen, stocks and bonds to complex ones like futures and options, subprime CDOs and credit default swaps. We build financial models to estimate the fair value of securities, to estimate their risks and to show how those risks can be controlled. How can a model tell you the value of a security? And how did these models fail so badly in the case of the subprime CDO market?

Physics, because of its astonishing success at predicting the future behavior of material objects from their present state, has inspired most financial modeling. Physicists study the world by repeating the same experiments over and over again to discover forces and their almost magical mathematical laws. Galileo dropped balls off the leaning tower, giant teams in Geneva collide protons on protons, over and over again. If a law is proposed and its predictions contradict experiments, it's back to the drawing board. The method works. The laws of atomic physics are accurate to more than ten decimal places.

It's a different story with finance and economics, which are concerned with the mental world of monetary value. Financial theory has tried hard to emulate the style and elegance of physics in order to discover its own laws. But markets are made of people, who are influenced by events, by their ephemeral feelings about events and by their expectations of other people's feelings. The truth is that there are no fundamental laws in finance. And even if there were, there is no way to run repeatable experiments to verify them.

You can hardly find a better example of confusedly elegant modeling than models of CDOs. The CDO research papers apply abstract probability theory to the price co-movements of thousands of mortgages. The relationships between so many mortgages can be vastly complex. The modelers, having built up their fantastical theory, need to make it usable; they resort to sweeping under the model's rug all unknown dynamics; with the dirt ignored, all that's left is a single number, called the default correlation. From the sublime to the elegantly ridiculous: all uncertainty is reduced to a single parameter that, when entered into the model by a trader, produces a CDO value. This over-reliance on probability and statistics is a severe limitation. Statistics is shallow description, quite unlike the deeper cause and effect of physics, and can’t easily capture the complex dynamics of default.

Models are at bottom tools for approximate thinking; they serve to transform your intuition about the future into a price for a security today. It’s easier to think intuitively about future housing prices, default rates and default correlations than it is about CDO prices. CDO models turn your guess about future housing prices, mortgage default rates and a simplistic default correlation into the model’s output: a current CDO price.

Our experience in the financial arena has taught us to be very humble in applying mathematics to markets, and to be extremely wary of ambitious theories, which are in the end trying to model human behavior. We like simplicity, but we like to remember that it is our models that are simple, not the world.

Unfortunately, the teachers of finance haven’t learned these lessons. You have only to glance at business school textbooks on finance to discover stilts of mathematical axioms supporting a house of numbered theorems, lemmas and results. Who would think that the textbook is at bottom dealing with people and money? It should be obvious to anyone with common sense that every financial axiom is wrong, and that finance can never in its wildest dreams be Euclid. Different endeavors, as Aristotle wrote, require different degrees of precision. Finance is not one of the natural sciences, and its invisible worm is its dark secret love of mathematical elegance and too much exactitude.

We do need models and mathematics – you cannot think about finance and economics without them – but one must never forget that models are not the world. Whenever we make a model of something involving human beings, we are trying to force the ugly stepsister's foot into Cinderella's pretty glass slipper. It doesn't fit without cutting off some essential parts. And in cutting off parts for the sake of beauty and precision, models inevitably mask the true risk rather than exposing it. The most important question about any financial model is how wrong it is likely to be, and how useful it is despite its assumptions. You must start with models and then overlay them with common sense and experience.

Many academics imagine that one beautiful day we will find the 'right' model. But there is no right model, because the world changes in response to the ones we use. Progress in financial modeling is fleeting and temporary. Markets change and newer models become necessary. Simple clear models with explicit assumptions about small numbers of variables are therefore the best way to leverage your intuition without deluding yourself.

All models sweep dirt under the rug. A good model makes the absence of the dirt visible. In this regard, we believe that the Black-Scholes model of options valuation, now often unjustly maligned, is a model for models; it is clear and robust. Clear, because it is based on true engineering; it tells you how to manufacture an option out of stocks and bonds and what that will cost you, under ideal dirt-free circumstances that it defines. Its method of valuation is analogous to figuring out the price of a can of fruit salad from the cost of fruit, sugar, labor and transportation. The world of markets doesn't exactly match the ideal circumstances Black-Scholes requires, but the model is robust because it allows an intelligent trader to qualitatively adjust for those mismatches. You know what you are assuming when you use the model, and you know exactly what has been swept out of view.

Building financial models is challenging and worthwhile: you need to combine the qualitative and the quantitative, imagination and observation, art and science, all in the service of finding approximate patterns in the behavior of markets and securities. The greatest danger is the age-old sin of idolatry. Financial markets are alive but a model, however beautiful, is an artifice. No matter how hard you try, you will not be able to breathe life into it. To confuse the model with the world is to embrace a future disaster driven by the belief that humans obey mathematical rules.

MODELERS OF ALL MARKETS, UNITE! You have nothing to lose but your illusions.

The Modelers' Hippocratic Oath

~ I will remember that I didn't make the world, and it doesn't satisfy my equations.

~ Though I will use models boldly to estimate value, I will not be overly impressed by mathematics.

~ I will never sacrifice reality for elegance without explaining why I have done so.

~ Nor will I give the people who use my model false comfort about its accuracy. Instead, I will make explicit its assumptions and oversights.

~ I understand that my work may have enormous effects on society and the economy, many of them beyond my comprehension.

APPENDIX: COMPARING A THEORY WITH A MODEL

DIRAC VS. CAPM

Dirac doesn't say:

*if (*there is a sea)
then (the electron behaves as follows).

Dirac employs no 'ifs'. He says:

This is the equation the electron satisfies; in consequence there is a sea; in consequence there must be positrons.

CAPM says:

If {
investors are rational *&*
all investors have access to the same information at the same time *&*
they can go long or short as much of any security as they like for no fee *&*
they can lend and borrow unlimited amounts of money at the same risk-free rate of interest &
they are naturally risk-averse in the sense that they demand a higher promised return for investing in a risky security
}
and if
the risk of a security is described entirely and only by the volatility of its returns
and if
the only undiversifiable risk is market risk
then
the excess return one can expect to earn over the riskless rate when buying a security is equal to the security's beta times the excess return you can expect to earn from the whole market.