The problem in attacking a framework, as opposed to a theory, is that there is no definitive test. The CAPM was a simple model with an unambiguous implication, and has been rejected so thoroughly it is now indefensible, yet still, strangely, prominent in the MBA curriculum. But the APT and general equilibrium (SDF) framework are healthy in academia because they can never be disproven, as any test of that theory is merely a test of one person’s guess as to how to reify risk. Any single bit of evidence, such as the low return to highly volatile stocks, in isolation, is just an anomaly, and may merely suggest a new risk proxy. Yet fundamentally, the theory of risk premiums is based on the idea that we do not like things that covary with our wealth, broadly defined, because they increase our net wealth volatility, broadly defined. It seems reasonable to presume, therefore, that priced risk is somehow positively correlated with volatility or betas.

Thus, the most damning evidence is the scope of the volatility-return failure across many asset classes. This evidence has never been presented as an argument for the failure of the conventional theory because as this theory cuts across several asset classes, each with their own measurement and stochastic characteristics, a neat statistical amalgamation test is difficult if not impossible. But the welter of data is broad, and examples of a positive volatility-return relation are the anomalies, rather than the observation that volatile securities have lower returns in a particular asset class.

Even an incorrect framework should be moderately consistent with reality. Whatever the true risk factor is in the current models, it should be somewhat positively related to volatility, or be positively correlated with the business cycle (risk should be greater in bear markets). These are basic, intuitive properties about risk, present at the creation of our conception of risk aversion, and the definition of risk, related to both our dislike of wealth variance, and our relative dislike for things that go bad when everything goes bad. As a first approximation, we should see vague, general evidence consistent with this view of risk if the standard theory is correct, because if risk is
totally uncorrelated with volatility or the business cycle, the as-yet unidentified risk factor makes little sense. After all, markets are presumably aggregating risk preferences, and if risk is related to something uncorrelated with the business cycle, and the stock market, and total volatility, it defies credulity that this is what people really think of as risk. In most prominent venues of risk and return we can imagine, the things that are volatile, that seem to do worst in recessions, do not have higher returns. Indeed, more often than not, they have lower-than-average returns. Let us consider the various venues where researchers have looked at how risk and return are correlated.

TOTAL VOLATILITY AND CROSS-SECTIONAL RETURNS

In 1992, when I was looking for a dissertation topic, I found that volatility was negatively correlated with returns. The key bias obscuring this fact was the size effect. If you controlled for size, you saw volatility negatively related to returns. My dissertation in 1994 was on documenting this finding, and arguing that it was relevant to the belief that mutual funds showed a preference toward highly volatile stocks—investors were buying hope, based on overconfidence, the amenability of volatile stocks to a sales pitch, and trying to take advantage of the highly nonlinear fund inflows from doing very well.¹

Why this was so incredible is that idiosyncratic variance should pick up mismeasured factor loadings, and mismeasured factor loadings should help explain the poor performance of factor models such as the APT.² For example, as mentioned, Stephen Ross and Richard Roll point out that it is possible—though not probable—inefficient estimates of the market portfolio are uncorrelated with returns, yet then residuals should still show a positive correlation with returns.³ That is, if some factor is misestimated, the measure of beta is not perfect because of imperfections in the risk proxy. But then the residual variance in such an equation should be positively correlated with returns. Furthermore, early in asset pricing, residual risk was thought to capture risk because of naïve investors failing to distinguish between systematic volatility and diversifiable volatility, seeing only a very volatile stock.⁴ Because many investors are not highly diversified, being undiversified in volatile stocks is more unpleasant than being undiversified in a low volatility stock. For both of these reasons, residual risk should be positively correlated with return, yet in fact the correlation for residual risk goes the wrong way.

I figured that knowing that low volatility stocks had lower returns than high volatility stocks, adjusted for size, had first-order implications for any strategy maximizing a Sharpe ratio, because the denominator in a Sharpe
ratio, the variance, would be lower for the low volatility strategy, and returns would be higher. I was in the economics department, and most of the finance professors at Northwestern considered this finding uninteresting, my story ad hoc. Though worthy of a dissertation, I did not get much positive feedback on my asset return finding, though my mutual fund finding related to this became the basis of a Journal of Finance article.\(^5\) It did not bother me because I figured that the investment world would love this result, and I liked the idea of becoming a portfolio manager as opposed to an academic.

As with any fact that really exists, others discovered this finding, too. Most notably, a paper by Ang, Hodrick, Xing, and Zhang in 2006 documented the relationship between idiosyncratic variance and returns cross-sectionally, and followed this up with another paper documenting it internationally.\(^6\) Please see Figure 4.1. While I was at Northwestern, Bob Hodrick was head of the Northwestern finance faculty, and when I was looking for thesis advisers, he was uninterested in my findings. Hodrick was a Chicago school, efficient markets type, who saw a risk factor or spurious result whenever someone saw an anomaly and my anomaly seemed no different.

But facts catch up with everyone, and it's a testament to the true scientific nature of finance that researchers eventually acknowledge the facts in a way that politicians or soft sciences do not. So now Hodrick was a co-author on a finding I found 10 years earlier. They found that stocks with higher volatility had significantly lower returns, and this held constant:

- Book to Market
- Leverage
- Liquidity
- Volume

![Figure 4.1: Annual Excess Returns for G-7 Equities](image)

**FIGURE 4.1** Annual Excess Returns for G-7 Equities
From Ang, Hodrick, Xing, and Zhang 2007, Table 7. Returns are annualized from 1963 to 2003.
* Turnover
* Bid-Ask Spread
* Coskewness (when returns go up, more positive skew)
* Analysts’ dispersion

That is a lot of things to control for, a pretty consistent finding. The results are in Figure 4.1, and I think the essence of any real finding is that it shows up in a simple-to-read graph, and the negative excess returns are very large. These are the average annual returns for quintiles sorted by idiosyncratic volatility. As is customary in academic studies, they only present excess returns, which are like the alpha in a market model that includes factors, in this case, the four-factor Fama-French factors. You take the total returns, and subtract the factor returns, because the excess return is that which is unexplained by the Fama-French factors. Because these factors are generally positively correlated to variance, which merely increases the alpha, the total volatility and beta are positively correlated, meaning highly volatile stocks have a higher beta and higher expected return. But that’s all a distraction in this case, because the excess returns are ridiculously large—15 percent annually for the highest volatility stocks. The following year, they documented this effect in 23 developed country markets and leave this finding as a global puzzle.

Like any result, you could find earlier premonitions in the literature. Most notably, Robert Haugen documented the low return to volatility and financial leverage back in the 1990s in a series of papers and a book. But this effect was somewhat hidden because it was given equal weight with several other independent explanatory variables, and so was not given prominence. The negative cross-sectional return to high volatility stocks is indisputable; the question is why.

**BETA-SORTED PORTFOLIOS**

When I found while researching my dissertation that volatility was negatively correlated with returns, I found that beta was also negatively correlated with returns at the high end. Yet I had to focus on idiosyncratic volatility because the idea that beta was actually slightly negatively correlated with returns was so preposterous, I would lose all credibility for even broaching the idea. My dissertation was consequently titled “Mutual Funds, Idiosyncratic Variance, and Asset Returns.” We also see that Ang et al. focused on idiosyncratic volatility. But the truth of the matter is that firms with higher betas have higher idiosyncratic volatility and vice versa. Indeed, the return by beta, when cross-tabbed with size, was reported as negative by Fama and French (1992), but that was presumably measurement error. The gist reported was that beta was uncorrelated with returns.
It is one thing to say beta does not work, but quite another to say that it is negatively correlated with returns. There had been too much work generating stories, anecdotes, and so forth to help explain beta to the great unwashed B-schoolers and their ilk, to now say—sign error! Best to say, beta is incomplete, and perhaps there's a risk gremlin lurking in low idiosyncratic risk. But actually, the returns to beta are worse than anyone expected.

I constructed beta portfolios using only those stocks with sufficient market capitalization to be investable. As an institutional investor has great difficulty investing in stocks with market caps below $500MM, this is in the 20th percentile of the New York Stock Exchange companies today. As these NYSE stocks are generally larger than those listed on the American Stock Exchange, or the Nasdaq, applying this cut-off backward in time is a nice way of focusing on those stocks that are truly investable, which is important in these types of tests because really small stocks tend to dominate the extremums of beta or volatility tests. No ETFs, REITs, closed-end funds, and so on. All common stocks. Next, I simply calculated the beta of each stock against the value-weighted NYSE-AMEX market return using the prior 36 to 60 months of data. The top 100 were high beta stocks, and the bottom 100 were low beta stocks. Please see Figure 4.2.

Figure 4.2 shows that historically, the cumulative return of low beta stocks significantly outperforms the high beta stocks. While the Internet bubble brought them back to an even performance, this was an aberration, and their subsequent performance in the Internet bust and the recent market cataclysm generates a net annualized 4.6 percent premium to low beta stocks. Like the tortoise and the hare, slow and steady wins the race. High betas are risky in the sense they have much higher volatility, and covary directly with the business, or market, cycle. Please see Table 4.1.

![Figure 4.2 Total Return to Beta Portfolios](image-url)
We can see that much of the poor performance from the high beta portfolio comes from the effect of geometric averaging, because the high beta stocks have such a higher volatility. Remember that the geometric average is a buy-and-hold return, whereas the arithmetic average assumes one can true up one's position daily, and the difference is the variance divided by two. Indeed, calculating the geometric return by way of a total return index, and applying the formula, generates about the same result. For the high beta portfolio, this adjustment lowers returns by 3.9 percent annually, whereas for the low beta portfolio, it is more like a 0.6 percent adjustment (it does not show up in the actual geometric returns because there is some positive autocorrelation in the low beta return). Considering the differences in beta, you could put on a negative beta portfolio at a premium (long low beta, short high beta), and then go long the aggregate market—which supposedly also generates a positive return—and generate a significantly positive zero beta portfolio.

It is striking that as prominent as beta is in financial pedagogy, there are no portfolios grouped by beta one can look up online or in Bloomberg, while value, size, and even momentum have a variety of different indexes and fund groupings. If beta is on par with these atheoretical factors, why is it not monitored the way these are? Because the high beta stocks have lower returns than low beta stocks once one merely controls for size, and so the beta factor proxy constructed in the way that the size and value factors are constructed would give it a negative price, which would turn the most fundamental parameter of finance (the equity premium) on its head. Nonetheless, today it is a dirty little secret, something all good equity quants know, but it is rarely addressed directly. It makes absolutely no sense in the conventional theoretic framework, and strikes at the very core of what is meant by risk, because if beta is inversely related to risk, then the standard theory is irredeemable. Thus, much prominence is given to Ang et al.'s piece on idiosyncratic variance, but the same result applied to beta would simply be too hard to square with the vast array of examples of how beta should work.
CALL OPTIONS

Theoretically, beta—or any covariance with the elusive risk factor—measures the "how much" of risk, and so if risk is priced, higher options with higher strike prices (that is, more out-of-the-money) have higher beta(s), which implies a higher average return. Therefore, far out-of-the-money call options should offer extremely high expected returns. As underlying stocks always, in practice, have positive betas against the market, all calls will have positive betas that exceed the beta of the underlying stock, and call betas will increase in the strike price as the calls get further out of the money. Hence, all calls will have positive expected returns and the expected returns will be larger for greater strike prices, because the betas, as a function of the call price, increase as you go out of the money.

For example, say you have a stock with a price of 100, and buy a call with a strike price of 120, expiring in three months. If the stock price rises to 110 over the next month, the call option will rise about 120 percent, while a long stock position rises only 10 percent. This is the implicit leverage in an option, that is, it is like being able to borrow 10 times one's capital and invest in the market. It is exactly the same bet as an equity position, just higher powered. This is why greedy retail investors with inside information prefer options: You get the most bang for your buck in options if you know where stock prices are going.

Coval and Shumway (2001) prove that expected European call returns must be positive and increasing in the strike price provided only that investor utility functions are increasing and concave and stock returns are positively correlated with aggregate wealth. They then document that this does not, indeed, happen.\(^7\) Please see Figure 4.3.

![Graph showing monthly returns and betas for call options ranked by delta](image)

**FIGURE 4.3** Monthly Returns, Betas for Call Options Ranked by Delta
Sophie Ni (2007) looked at data from 1996 through 2005, and found that the highest out-of-the-money calls, with one month to expiration, have average returns of —37 percent over a month. Figure 4.3 shows that if you bucket call options into groups based on their deltas, you find that call options, indeed, are indeed highly levered stock positions. Lower deltas mean the call option is less sensitive, in dollar terms, to a stock moving, but more sensitive, in percentage terms. Thus, an at-the-money call option with a delta of 0.5 moves 0.5 dollars for every 1 dollar move in the underlying, while an out-of-the-money option may have a delta of 0.08. On a percentage basis, since the at-the-money option has a price of around 5, while the out-of-the-money option a price of 0.25, the percentage change in price for the low delta option is much greater. The key to remember is that the average stock has a beta of 1.0, and betas on their call options range from 4 to 15—giving one 4 to 15 times the juice of the daily return. An option’s beta is the beta of the stock, times the omega, which is a measure of the percentage return in the option price given a 1 percent change in the stock price. If the omega on a Ford call option is calculated to be 3.6, then for every 1 percent change in the price of Ford, the price of the call option will rise by 3.6 percent.

Not only is the average return negative for call options, these returns get worse the more implicitly levered, and the riskier the options become, in contrast to what the weak assumption is as described by Coval and Shumway. Returns are negatively correlated with the betas. Call options have several times more risk, whatever that is, than their underlying stocks, yet negative average returns. Investors basically are overpaying for lottery tickets when they buy options, and just like the lottery, the average payout is worse the more risk one takes. If there’s a risk premium in equities, it certainly is not amplified in options in any way, because you lose money, on average, buying leverage market positions by way of call options.

**SMALL BUSINESS**

Entrepreneurial investment, such as in small proprietorships (S corps and private LLCs) is generally a highly undiversified investment for most entrepreneurs. The reasons are straightforward, in that when one person has a significant effect on the business through his effort and competence, it is natural that he should have the most “skin in the game.” This is a classic issue of moral hazard because a business manager, who has significant upside and, without ownership, no downside, is motivated to take wild risks on the theory of heads I win, tails the banker loses. However, if the manager is the majority owner, his failure should affect his net worth too. About 75 percent of all private equity is owned by households for whom it constitutes at least
half of their total net worth. Furthermore, households with entrepreneurial equity invest on average more than 70 percent of their private holdings in a single private company in which they have an active management interest. Despite this dramatic lack of diversification, private equity returns are on average no higher than the market return on all publicly traded equity.9 Figure 4.4 shows the basic results, that over an 8-year period, if anything returns to private business, be it partnerships, proprietorships, S corps, C corps, and two entirely different sets of data, there is no demonstrable premium. Given an investor can invest in a diversified, and liquid equity portfolio, it is puzzling why households willingly invest substantial amounts in an asset with an equivalent return, but much higher volatility, including a positive correlation with the market.

The forced nondiversification of a private equity investment, from a pure portfolio perspective, implies a requisite higher return. How much higher than the average public equity return would we expect the average private equity return to be? Using standard utility models to calibrate the hurdle rate that would make a household indifferent between investing in a portfolio of a single private firm, a public equity index, and T-bills, or a portfolio of just the public equity index and T-bills, researchers estimate that private equity risk generates a hurdle rate of about 10 percent higher the public equity return.10 You should receive a huge premium for the large idiosyncratic risk you are taking, risk that unlike idiosyncratic risk in the market, is impossible to diversify away. Entrepreneurs appear to be taking extra risk, for no extra return.

**Leverage**

The Miller-Modigliani theorem states that regardless of the debt and equity proportions, the value of the firm is the same. As a firm increases its leverage
using more debt, its equity concentrates the variable returns of the business on a smaller and smaller base, making them both riskier: the equity’s beta and volatility will increase, the debt will have a higher chance of defaulting. The implication is that highly levered firms should have lower rated debt (junk), and more volatile equity, but because debt has a lower return than equity, the net, total return to all a company’s securities (debt and equity) is a constant. That is the theory. But as was just shown, levering up the equity does not appear to increase returns in the case of options, which are merely levered equity positions, and equity option betas were shown to be highly negatively related to returns. Please see Figure 4.5.

In Figure 4.5, we see that leverage is, anomalously, clearly negatively correlated with returns. These researchers held constant size and book/market, so that this market leverage should not pick up these well-known anomalies. Higher leverage implies lower returns for equity even though this should increase risk of that equity, and thus should increase returns. There have been only a handful of studies on this, which is understandable because no one likes to generate results that do not support well-established theories, and researchers tend to think the theory is correct, and so the analysis is probably wrong. Indeed, perhaps the biggest implication is from the dog that is not barking here. There have been no papers linking how leverage is positively related to expected returns, even though this result would have been consistent with a Nobel Prize-winning theory. Empirically supporting Nobel-winning theory for the first time is worthy of a publication in a top journal, and for academics, this is their number one priority. The absence of a positive finding in this context is perhaps more powerful than the handful of negative results.

![Figure 4.5](image)

**FIGURE 4.5** Annual Return to Portfolios Sorted by Market Leverage (Debt/Market Cap) Adjusted for Book/Market and Size Exposure
*Source: Penman, Richardson, and Tuna, 2007, Table 1.*
MUTUAL FUNDS

The original tests of the CAPM were on mutual fund returns, hoping to show that mutual fund performance would be explained by the new risk factor.¹³ That is, wouldn’t it be neat if economists could explain the returns, not as managerial expertise, but beta? Early work by Jensen looked at the value add of mutual fund managers, and denoted their abnormal performance \textit{alpha}, a term that has stayed with us.¹⁴ Subsequent work found no relationship between a stock’s return and its beta, and it was sufficiently uninteresting that in more recent work, the relation between beta and returns is addressed only as an aside in Malkiel (1995), who notes that a fund manager’s beta is uncorrelated with his fund’s average return.¹⁵

As with leverage studies, the absence of any volatility or beta correlation with mutual fund returns is most relevant here, because it highlights an absence of confirmation in an area examined since the very beginnings of asset pricing theory. Absence of evidence is evidence of absence, not proof, but suggestive, especially when you know there has been a systematic, thought­ful search for such evidence.¹⁶ Carhart’s 1997 study of mutual funds is most well known for introducing momentum as a factor, akin to Fama and French’s value and size factors, but it was also notable for the manifest irrelevance of beta in his analysis.¹⁷ Good or bad, a mutual fund’s beta was never an issue in explaining the results, and so that paper hardly mentioned the null results, and instead centered on the importance of momentum as a factor that explained the positive one-year persistence. Later studies of mutual funds by Russ Wermers (2000) don’t even address beta.¹⁸ The problem is not a lack of attention to this issue, but the absence of any empirical relevance of beta to average mutual fund returns, leading researchers to focus on factors that are correlated with return, which are then rationalized as risk.

If you look at Morningstar’s detailed analysis of mutual funds on their splash pages for particular funds, they note the degree to which the fund is in value versus growth, or in large versus small cap, the asset turnover, expense ratio, and the tracking error. No beta. Correlation with the market, or any risk factor, is left as an exercise for the reader. This highlights the situation that in practice no one cares about beta when investing in mutual funds, nor has there been any evidence that they should. People just want funds that generate high returns.

FUTURES

Futures are derivative securities, bilateral agreements, one side to buy, the other to sell, at a future date, a spot commodity at a prespecified price.
Futures returns are not driven by lower expected spot prices because such prices are reflected in a low current futures price. Unexpected deviations from the expected future spot price are by definition unpredictable, and should average out to zero over time for an investor in futures, unless the investor has the ability to correctly time the market.

What return can an investor in futures expect to earn if he does not benefit from expected spot price movements, and is unable to outsmart the market? The difference between the current futures price and the expected future spot price. Assume the current futures price is below the current spot price. This usually implies the expected spot price is above the futures price (we don’t truly observe the actual expected futures price, but this is generally true). On average, going long the futures makes money when it is below the current spot price because the futures price rises to the eventual spot price. At maturity, while the spot price may have fallen, the futures price has risen too. This is called normal backwardization because if you put the futures prices out like a yield curve, the more distant futures prices are below the current price. Please see Figure 4.6.

Figure 4.6 shows the term structure of futures for gold and copper in August 2008. Copper is in backwardization, while gold is in contango, a fun name that means the opposite. Historically, gold is always in contango, meaning, if you are long gold futures, you lose money on average as it rolls to maturity. Other commodities flop around, sometimes flat, sometimes in normal backwardization, and sometimes in contango. Harvey and Erb (2007) find that copper, heating oil, and live cattle were on average in backwardization, while corn, wheat, silver, gold, and coffee were in contango, on average.20
Now, on average, this term structure relating futures prices to maturity dates predicts future returns fairly well, so that a commodity in normal backwardation generates a predictable positive return for being long, whereas a commodity in contango generates a predictable positive return for being short. The question, obviously, is why are some futures in contango, and others in normal backwardization, from a risk perspective? A prominent early explanation put forth by none other than John Maynard Keynes, on why futures generate risk premium from being long, is that farmers grow wheat, say, and wish to hedge it by selling now, rather than waiting until the season is over. So a speculator buys the wheat now, and takes on the price risk, for which he must be compensated.

Futures allow operating companies to hedge their commodity price exposure, and since hedging is a form of insurance, hedgers must offer long-only commodity futures investors an insurance premium. Normal backwardation suggests that, in a world with risk-averse hedgers and investors, the excess return from a long commodity investment should be viewed as an insurance risk premium. It is easy to expand this to the other side, by focusing not on the producer of a commodity, but the purchaser. Say you are Boeing, and buy a lot of aluminum to build airplanes. If you hedge, you buy futures today, locking in a price. Thus, whether you hedge by buying if you are a consumer, or selling if you are a producer, futures have an insurance-like characteristic. The key is knowing, between consumers and producers, who dominates the futures contracts. One explanation of the futures returns is that for some commodities, producers dominate the demand for insurance, and thus futures, in the other, consumers dominate.

In a diversified worldwide market, however, this reasoning does not work in explaining equilibrium returns. Asset pricing theory tells us that returns are a function of risk. And as most investors are not aluminum consumers, or corn suppliers, the net covariance with the risk factor should be at work. For example, the needs of a company, its preferences, are unrelated to its returns, which are a function of the change in the expectation of a company’s cash flows in relation to these other things we care about (for example, the S&P500). This is due to arbitrage, and because asset prices are set by supply and demand, where investors should be allocating capital in a way so that the price of risk, from any source, is the same whether it comes from futures or equities. If one can get the benefits of the futures roll and not be involved in the futures commodity—as most investors are not—this should be like idiosyncratic risk is, the CAPM: diversifiable, and so unpriced. And the expected roll returns (so-called because the futures prices roll to the current spot price over time), based on the current relation of the futures to the spot, are uncorrelated with the prominent risk factors for equities (that is, the market, value, and size factors) or for corporate bonds.
(that is, the Baa-Aaa yield spread). Changes in inflation adversely affects the roll returns from normal backwardization, while adversely affecting the roll returns for contango.\textsuperscript{21}

In sum, there are predictable returns in futures returns, primarily from the movement in the futures price as the maturity date moves closer to the present, which is foreseen in the current relation of the futures price to the spot price. But what drives this, from a risk perspective, is a mystery.

**CURRENCIES**

A currency is not just a medium of exchange, but an asset with a return like a stock. The interest rate is like a dividend, the change in spot price, the capital appreciation. One would expect the return of currencies to be related to risk.

Uncovered Interest Rate Parity is a theory that connects current to future spot rates. This theory states that you have two ways of investing, which should be equal. First, you can invest in your home country at the riskless rate. So if the U.S. interest rate is 5 percent, you can make a 5 percent return in one year, in U.S. dollars (USD). Alternatively, you can buy, say, yen, invest at the yen interest rate (each currency has a different risk-free rate), and then convert back to USD when your riskless security matures. For this to be equal, you need something like

\[ r_{USD} = r_{yen} + \% \text{ change in yen} \]

Where \( r_{USD} \) is the U.S. interest rate, and so on. So, if you make 5 percent in USD, an U.S. investor should receive that same return in yen, by way of the interest rate in yen, plus the expected appreciation or depreciation in the yen against the dollar. If the interest rate in yen is 1 percent, this means one expects the yen to appreciate by 4 percent. When the foreign interest rate is higher than the U.S. interest rate, risk-neutral and rational U.S. investors should expect the foreign currency to depreciate against the dollar by the difference between the two interest rates. This way, investors are indifferent between borrowing at home and lending abroad, or the converse. This is known as the uncovered interest rate parity condition, and it is violated in the data except in the case of very high inflation currencies. In practice, higher foreign interest rates predict that foreign currencies \textit{appreciate} against the dollar, making investing in higher interest rate countries a win-win: You get \textit{appreciation} on your currency, and higher riskless interest rates while in that currency.
Now the rates of expected return through the two investment paths can differ according to risk, of course. So one can imagine, looking at the yen, or the dollar, or various European currencies in the 1970s, and so on, trying to tie each to some measure of a home currency’s risk factor: consumption, or the stock market and so forth.

Like high returns to low volatility stocks, it is difficult, but not theoretically impossible, to make sense of this. Robert Hodrick wrote a magisterial technical overview of the theory and evidence of currency markets in 1987. He summed up his findings:

We have found a rich set of empirical results. We do not yet have a model of expected returns that fit the data. International finance is no worse off in this respect than more traditional areas of finance.

That is seeing the glass half full, such as a book of models, trying to find something that would intuitively relate to a risk factor that predicts the perverse finding that futures curves predict currency movements. Indeed, Hodrick looked at CAPM models, latent variable models, conditional variance models, models that use expenditures on durables, or nondurables and services, and Kalman filters. None outperformed the spot rate as a predictor of future currency prices. Hodrick leaves off with the idea that “simple models may not work well.” Indeed this is true, and I think is the ultimate hope of these researchers that a little more math will uncover a solution that is merely complicated.

And so it continued for the next 20 years, and many hedge funds specialized in the carry trade, which was as simple as it was successful: lend capital to high interest rate currencies, enjoy the high riskless rates and currency appreciation on the spot rate; borrow capital at the low interest rate currency, and make money on the depreciation of this debt over time. Thus, in 2008, researchers took a different tack, and noted:

Overall, we argue that our findings call for new theoretical macroeconomic models in which risk premia are affected by funding and liquidity constraints, not just shocks to productivity, output, or the utility function.

By “our findings” they mean the carry trade continued to work 30 years after being identified by Farber and Fama (1979), and it has continued as a puzzle because no reasonable risk factor can explain it. Thus, they are looking at new conceptions, in this case, based on negative skew, because liquidity constraints is an academic euphemism for downside risk (and thus blowing through one’s capital).
Lotteries

The annual per capita lottery expenditure in the United States is about $170, and the rate of return is about —47 percent per dollar played. It clearly presents a challenge to the idea that people are looking at these games on a risk-return continuum, but they are in some sense risky decisions at their most basic level. These investments clearly cater to what is commonly called those seeking risk, or positive skew, in particular. There are two primary characteristics of lotteries. First, poor people play them more, in both relative and absolute terms, than wealthy people. A St. Louis Fed article finds that those with household income of less than $25,000 spent $575 on lotteries on a per capita basis. This spending was substantially more than spending by those with a household income over $100,000 ($196). The people who can least afford it, buy the most of it.

A study of the popularity (sales) of lotteries found that average payout (expected return) or variance did not matter, but the size of the top prize was highly significant. In other words, the $100 million Super Lotto has the most sales even though the probability of winning is so small it basically is outside the realm of intuition (1 in 150 million). Indeed, such lotteries are the only ones I play. People who bet seem to prefer those bets that offer the worst odds, but the greatest payout. Gambling seems to be totally outside the assumptions of risk aversion that underlie the risk-return assumption, and is one of the motivators of Prospect Theory, though all this theory does is say that when little sums are involved, people are risk loving, which is not so much a theory as a description of small stakes gambling.

Movies

A paradox in the movies is that their rate of return is around 4 percent and the risk is higher than most industries. Art DeVany (2003) found that between 1986 and 1999, G-rated movies generated lower volatility and higher returns than R-rated movies, though there was a clear preference toward R-rated movies (over 1,000 R-rated movies and only 60 G-rated ones). But movies have a strong Pareto distribution, where the mean is much higher than the median or mode. It seems everyone is betting on the next Titanic, because the very highest grossing movies are R-rated. Furthermore, the R-rated genre is more artsy, and so has reputational aspects to those involved that go beyond mere revenues.

WORLD COUNTRY RETURNS

The risk premium is an expectation, and returns are a realization. To the extent a return is high, it is never clear whether this return was expected or
rather just random luck. Nevertheless, one would expect that over the long run, or for lots of cases, the average return equals the expected return. If not, what is the point of an expectation?

Thus, it is interesting that among country equity returns, there is no clear risk premium. The United States had about the same average top line return relative to short-term debt from 1900 to 2005 compared to 17 other developed countries worldwide, about 5 percent.27 There is no clear return either within developed countries, or between developed and undeveloped countries. Erb, Harvey, and Viskanta (1995) show similar returns between developed and underdeveloped countries, using data from 1979 to 1992.28 As the arithmetic returns are much higher than the geometric returns, he highlights those monthly returns for displaying a risk premium based on the volatility, but if you look at the geometric returns, they are about the same as the developed country data (13.5 percent for developing, 12.3 percent for developed). Using a more updated set of data from 1989 through 2000, Bansal and Dahlquist (2002) report approximately similar arithmetic returns (15.8 percent versus 16.1 percent), but then using annualized geometric returns, the return for the developed countries was 13.8 percent versus 6.7 percent for the developing countries.29

Intuitively, investing in Nepal would seemingly be taking on extra risk, but in practice there is no risk premium for such forays historically. It is strange that there is not a pattern among their returns in regard to its volatility, because intuitively, those countries where the stock market index is especially volatile, would have a higher risk premium, as foreigners would not be able to invest sufficiently because of tax and institutional reasons, and eliminate the risk premium for this idiosyncratic risk. One can look for global risk factors that explain this, and the usual ones (for example, a world stock market index) do not work.

**CORPORATE BONDS**

The conventional corporate bond puzzle is that spreads are too high. The most conspicuous bond index captures U.S. Baa and Aaa bond yields going back to 1919, which generates enough data to make it the corporate spread measure, especially when looking at correlations with business cycles. Yet Baa bonds are still investment grade, and have only a 4.7 percent 10-year cumulative default rate after their initial rating. As the recovery rate on defaulted bonds is around 50 percent, this annualizes to a mere 0.23 percent annualized loss rate. Since the spread between Baa and Aaa bonds has averaged around 1.2 percent since 1919, this generates an approximate 0.97 percent annualized excess return compared to the riskless Aaa yield, creating the puzzle that spreads are too high for the risk incurred.
In the 1980s, Michael Milken was the point man for leading a revolution in finance, where firms with a 5 percent annualized default rate would have an active market, and firms could even issue bonds at this grade. As the old saw goes, banks only lend to firms that can pay them back, and so the probability of default, historically, for investment grade companies is well below 1 percent annualized. B and BB rated bonds have default rates of 6 percent and 1.5 percent, respectively, and before 1987, the market for these bonds was very illiquid, and all these bonds were “fallen angels,” bonds initially issued at investment grade whose ratings have fallen because of adverse financial performance.

There was a spirited debate as to what the actual default rate, and return, on junk bonds was during the 1980s, as Michael Milken and his firm Drexel Lambert were getting rich promoting these bonds, while others, like Warren Buffett and Ben Stein were saying these bonds had horrible returns. There simply were not a lot of data, so the debate was rather limited. My friend and former Moody’s colleague, Jerry Fons, wrote his dissertation on this market in 1985, and actually received payments from his university from people asking for copies of his dissertation because there was so little empirical data in this area.30 Not a lot of money, but a fun distinction, because most dissertations are not only never published, they are rarely read by anyone but one’s advisers.

Altman and Bana (2004) and Kozhemiakin (2007) note there is no premium to high yield portfolios relative to investment grade portfolios, a set of bonds with a 3.84 percent average annual default rate from 1970 to 2005.31 Furthermore, Altman (2006) notes that a bankrupt bond portfolio underperforms investment grade bonds.32 Both high yield and bankrupt bonds have more volatility and cyclicality than investment grade bonds, and do their worst when returns are most valued, in bad times. Junk bonds are intuitively, and academically, risky. Data from the Merrill Lynch High Yield Index show a 6.77 percent annualized return relative to the 7.18 percent return of their investment grade index from 1987 through December 2008. The risk premium is a negative 0.41 percent in these indexes, which seems odd given that high yield debt has greater volatility, and more cyclicality. This might be seen as mild support for the idea that risk begets return, but it assumes one can buy the index using closing prices during this period.

Yet the indexes are really an overstatement because such indexes have a systematic bias when portraying illiquid or unaudited asset classes. Even today, many times a junk bond’s bid-ask spreads is five points wide, and this transaction cost is implicit in the fact that for a set of mutual high yield mutual funds that currently exist (that is, with obvious survivorship bias), their total annualized return from 1987 through 2008 was 3.44 percent, whereas the Merrill High Yield Index rose 6.77 percent (see Figure 4.7) annually. Investment grade funds underperformed their index by much less,
in that the index return was 7.18 percent compared to fund returns of 6.48 percent. The simple idea is that illiquid assets have higher transaction costs, including higher management fees, that cause actual performance to be lower than that of an index. If it cost 5 percent to buy an asset (commissions, bid-ask spread, trade impact), given average turnover, this will diminish your returns by 5 percent, amortized over the average holding period in one’s portfolio. If your average holding is 3 years, over 30 years that’s not 10 trades, but 20, because you buy and sell each trade. Thus the closing prices of illiquid assets, such as in an index, will be a biased proxy for real returns if based merely on closing prices. If you take a couple percent off the high yield index due to price impact, commissions, and the bid-ask spread, there is no premium to these risky assets.

Illiquid and highly volatile assets often have this problem, and there’s a clear bias by data providers to ignore this risk in computing indexes. For example, Malkiel and Saha (2005) found that no hedge fund database providers had the last year of Long Term Capital Management in their data set, the −92 percent return, and so the indexes exclude the very risks that make hedge fund investors nervous. Malkiel and Saha estimate this bias adds approximately 6 percent to the annualized returns. Who creates these indexes? Usually groups that are allied with the product one is examining. For example, the Credit Suisse/Tremont Index that monitors hedge fund returns is maintained by the following:

The joint venture, Credit Suisse Tremont Index LLC, combines the considerable expertise of Credit Suisse, one of the world’s leading global investment banking firms, and the data research group of Tremont Capital Management, Inc., a full-service hedge fund of funds investment management firm.
It is an inevitable conflict of interest where those most knowledgeable, and have access to the best data, who provide data generally used by researchers, will be advocates of this field—one does not become extensively knowledgeable in something one thinks is irrelevant, inefficient, or fraudulent. There is oftentimes an arm’s length separation between the index provider and the portfolio managers, but there is no way they can be indifferent: without the market thriving, seeming to offer a good opportunity, their service will not have a long life. After all, it was conventional wisdom that active portfolio managers outperformed passive indexes until the 1980s—for several generations!—because the evidence was generally held and presented by the active managers and their industry groups. There are always good reasons why a certain investment should not be included in a database, and for someone with a rooting interest, these will tend to be poorly performing investments. So asset indexes are often biased, but little is done to note such biases because the index makers are advocates.

Therefore, the excess corporate risk premium puzzle pertains to one portion of the risk spectrum the difference between a 0.03 percent and a 0.3 percent annualized default rate, a distinction without a difference to most people. When one goes from a 0.3 percent to a 15 percent default rate, as one does when you go from BBB- to C rated bonds, there is no return premium at all. Given reasonable expectations of transaction costs, and the actual difference between the high yield indexes and actual high yield returns, it seem probable that people extend into higher credit risk with a lower average return. It is difficult to see how the little risk is priced, the big one not, if risk is to have any consistent meaning. If the corporate spread is a function of risk at one end, why is it not at the other, more intuitive end?

THE LONG END OF THE YIELD CURVE

The general shape of the yield curve is as follows. It rises about 1% until about three years, then flattens out. But this is deceiving, because bonds have positive convexity, and so their returns are very non-normal at annual frequencies. Furthermore, the higher maturity bonds have higher volatility, and this subtracts from their cumulative returns through the geometric averaging adjustment where we subtract the variance divided by two from arithmetic returns. We should expect long bond holders to have long frequencies, and so this adjustment on monthly data is important.

Thus, I took data on the U.S. government bond yield since 1958 through 2008 and constructed a set of annualized returns based on a buy-and-hold strategy. Each monthly return subtracted the Fed Funds rate, and included
both the monthly price appreciation plus the coupon yield. These data are presented in Figure 4.8.

As we can see, the returns to a funded position in bonds, an excess return, is very close to zero over this period. Indeed, buying six-month T-bills funded at the Fed Funds rate was a money loser, and for the one-year bond, about a zero return. But returns are increasing over maturity. Yet, the increase in yield from 5 years to 10 years is miniscule, from 0.55 percent to 0.65 percent, and it actually declines to 0.60 percent for the 20-year bonds. It is fair to say the returns to yield curve extensions past five years are basically zero. The price volatility, meanwhile, increases consistently as maturity increases, and thus the Sharpe ratio, the ratio of the return on the bond minus a risk-free rate, falls as the maturity increases beyond five years. Why would anyone hold long-term bonds, because you get the same yield (and return) but must take on more price volatility?

One explanation is that many investors have a specific time horizon they are interested in, say 10 years. It is merely a difference in perspective, the nominal payout in 10 years is fixed as least risky if you are fixated on consumption in 10 years, as compared to a security that rolls over the short rate many times. But now the theory is at a stand-off. For someone who needs a fixed payment in 10 years, the 10-year fixed rate bond is risk free. Changes in yield cause it to have price variability, which is no concern to the person with a fixed payment in 10 years. On the other hand, some people buy assets, and don't know when they will need it, or want its valuation to remain stable, as in the early naïve approaches to modeling bond risk. A
risk-free asset in this context would be an asset such as a floating coupon that reset every month, and a fixed price. So the issue is, what is risky: a fixed payment in N years and a floating current price, or a fixed current price and a floating terminal payment as the yield curve changes? In this framework, risk is as justifiable as one’s favorite color, because everyone has different investing horizons.

But even with the degrees of freedom implied by this reasoning, we face a puzzle. Why does the return, and thus risk, rise, and then flatten? Thus, using the aforementioned reasoning on fixed versus floating bonds, the two-year bond is riskier than the three-month bond because of its price volatility. But after three years, the price volatility does not matter. This cannot be because people are indifferent to a fixed payment in 20 years or 5 years, because if you assume five years, then the 20-year bond is much riskier, and vice versa if you assume people prefer fixed payments in 20 years.

The yield curve is usually used to demonstrate a risk-return premium by comparing two points on the yield curve, either the 1-year versus the 30-year, or the 1 month versus the 10-year. The key is, you span the 1- to 3-year fulcrum where the yield curve flattens out. Indeed, the six-month T-bill dominates the three-month T-bill in the sense that it generally gives one 20 basis points more return at immeasurably small levels of interest rate risk; the shorter end of the curve is a free lunch for investors who use the Sharpe ratio. But the returns to bonds and bills are often noted as proof of a risk premium, usually by merely noting that one has higher return and higher volatility. Price volatility, however, is clearly not the driver, because volatility continues to rise long after a return premium disappears, and no one thinks volatility as applied to stocks is related to risk any more.

One clever answer is that we observe nominal bonds, and to the extent there is an inflation risk premium that increases as maturities increase, because if there is really high inflation, like in the 1970s, where the real returns on long-term bonds were negative, the more so the longer the maturity. Short-term T-bills were able to keep pace with inflation because one continually rolled them over. Therefore, if one has a demand for a fixed, real, future payment, then not a nominal fixed payment, but rather, the short-term rollover strategy is preferred. Thus, after three years, there is an exact canceling out between the inflation hedging properties of the shorter-term bond (which can be rolled over at higher rates in high inflation times), which preserves future value, and the low risk that comes from a fixed nominal payment of a longer maturity bond. A clever set of assumptions about the nature of inflation versus real interest rate risk could get this to work.

Yet looking at the inflation indexed Treasuries, TIPS, which started around and in data collected since 2004, these show that the 5-to-20 year spread is virtually identical to the nominal Treasury spread over that period.
That is, the difference between the 20-year and the 5-year, when 2 percent nominally, was 2 percent for the TIPS. The yield curve for TIPS out two years looks just like the regular U.S. bond yield curve, just shifted downward. When the yield spread between short- and long-term bonds was near zero nominally for U.S. bonds, it was near zero for the TIPS. It would appear that inflation risk premia are not relevant to shape of the yield curve. The real Sharpe ratio declines for TIPS as one moves out in maturity just as for nominal Treasuries.

The bottom line is that an explanation has to explain the really low returns to the short end of the yield curve while explaining why price volatility and inflation risk do not matter, or cancel out after a certain point. This is impossible, so modern interest rate models are based on latent factors, yields as a function of their yields, and so they are atheoretical tautologies like saying value stocks outperform growth stocks because of the value risk factor. If the 10-year forward rate is above the current rate, you can say that that is because of inflation expectations, an expected increase in the real interest rate, and a risk premium. You cannot prove this is not true, but then you are taking a lot on faith. The specific reason why the expected return rises, or falls, or rises and then falls, is totally flexible, because the current yield is a function of itself, though suitably rationalized.

DISTRESS RISK

As mentioned in Chapter 3, distress risk, or the risk of financial default, was at one time thought responsible for both the size and value effect. Several studies subsequently documented that when measured directly, firms with higher distress risk have much lower returns, and so this explanation is no longer feasible. Indeed, while I was at Moody’s in 2000, I was able to use their database of ratings back to 1975 and find that the rate of return lined up almost perfectly with the rating, with AAA having the highest return and C the lowest. I updated those data using S&P ratings, and used the rating in June to form a portfolio over the following 12 months, a very straightforward strategy. I used 1987 as the starting point because before this, junk bonds, those with a rating below investment grade, were small in number, as there was a structural shift in the junk market in the late 1980s when these instruments started to have good pricing data. Please see Figure 4.9.

The returns are pretty constant until you get to the signature junk bonds, the Bs, and then it falls precipitously, and the Cs are even worse. Thus, the equity returns to firms with low financial strength are low, and their debt does not seem to compensate either. High risk, from a financial distress
perspective, appears generally flat, but then for the highest default risk companies, a strong negative relation to returns for agency-rated companies. When combined with the flat returns for the highest default risk bonds, it seems that the expected return on risky firms—debt and equities—is net lower than for less risky firms.

SPORTS BOOKS

Bets on high probability–low payoff gambles have high expected value and low probability–high payoff gambles have low expected value. For example, a 1-to-10 horse having more than a 90 percent chance of winning has an expected value of about $1.03 (for every $1 bet), whereas a 100-to-1 horse has an expected value of about 14 cents per dollar invested.36

This bias has appeared across many years and across all sizes of race track betting pools. The effect of these biases is that for a given fixed amount of money bet, the expected return varies with the odds level. For bets on extreme favorites, there is a positive expected return. For all other bets, the expected return is negative. The favorite long-shot bias is monotone across odds and the drop in expected value is especially large for the lower probability horses.37

Interestingly, sports like baseball and hockey, where the favorite odds are rarely greater than four to one, show no such bias. It seems the bias arises only in the extreme odds that are prevalent in horse racing, but absent in most sports. That is, even a bad team has a one in four chance of winning in baseball, whereas a bad horse’s chance will be more like 1 in 50. People
appear to pay for an opportunity for positive skew, big payoffs, but this is not risk aversion, but rather risk-loving, and only for the extreme odds.

**TOTAL VOLATILITY AND EXPECTED EQUITY INDEX RETURNS**

Thus far we have looked mainly at cross-sectional risk and return. This is because it is often simpler to look at one asset class, such as equities, that presumably have a similar relevant risk factor. If we compare bonds and equities, or bonds and houses, the differences make relative ranking of risk difficult, because their returns have different distributions; their data sources are different. But there is another prominent manifestation of risk and return: a time series. The stock market has what is known as serially correlated volatility: high volatility months follow high volatility months, low volatility follows low volatility. Unlike stock returns, volatility of assets is very predictable: I can predict when assets will have higher volatility tomorrow. Of course it isn’t perfect, but statistically, a big move in either direction today, implies a big move tomorrow.

The most basic risk models assume that the expected return on an asset is proportional to the expected nondiversifiable variance of the asset: the higher the variance, the higher the expected return. Modern models tend to focus on some abstract thing we don’t like, like declines in consumption, wealth, or output, but those bad states are generally coincident with higher volatility, as volatility increases when the economy is doing poorly. And so what do we see? Some research documents a null relationship between volatility and future returns, and some find a negative relationship.\(^{38}\)

Implied option volatilities are forward looking. They are market predictions of volatility, and thus, are true expectations. One may think these volatilities are biased, as they are driven by market supply and demand considerations, but if they were biased, it would be easy to make money arbitraging the difference by forming an arbitrage portfolio of the option and the underlying, and exploiting the difference. Indeed, in the late 1980s, many traders would buy yen options and hedge them with futures, because on average the implied volatilities were low. Please see Figure 4.10.

Using daily data from 1986 through 2007, I took all 5,300 observations from the VIX, which is a weighted measure of the implied volatility of liquid stock at the money put-and-call options. I then compared this to the stock market return over the next 12 months. I then grouped these observations into deciles based on implied volatility, and looked at the resulting average future stock returns. Figure 4.10 shows that there is definitely no positive correlation between expected volatility and future returns. If theory
suggested the opposite, one would probably see papers suggesting that the pattern is significant, but because there's no theory for this, researchers don't make such interpretations.

Steve Sharpe and Gene Amromin actually looked at survey data, and found that investor expectations were totally inconsistent with this assumption. They found that when investors have a more favorable assessment of macroeconomic conditions, they tend to expect higher returns. Second, they found that the expectation of more favorable economic conditions has a strong negative effect on expected stock market volatility. Another example is a Gallup poll put out by Paine Webber. In 1998, at the beginning of the stock market boom, they surveyed an expected return of 13 percent from investors. After back-to-back 20 percent-plus returns, when the Nasdaq doubled, investors raised their expectations to 18 percent in February of 2000, right before the peak. Two years later, after a 50 percent correction, and a 50 percent rise in the VIX (a measure of expected volatilities), they anticipated only a 7 percent expected return. So from a Sharpe ratio perspective, when investors expect a high numerator, they expected a low denominator! Win-win. They expect good times to be high returns and low risk, and bad times to be low returns and high volatility. Investors do not see risk and return as a trade-off, but rather, as Warren Buffett has articulated, low risk means a good return for many investors, in that if you think an asset will rise 20 percent, as opposed to 10 percent, it has less risk of returning less than zero.
This flies in the face of the axiom that stock market returns should compensate investors for exposure to anticipated macroeconomic risks: when times look good, people expect good returns, and when times look bad, people expect bad returns. Intuitive at some level, but counter to theory that at some level, risk implies a higher expected return.

**UNCERTAINTY AND RETURNS**

Keynes's and Knight's writings on this have generated a lot of intuition and excitement because there is something profound about their idea that the risks we take are distinct from the risks in lotteries or roulette tables, with their explicit \textit{a priori} odds, but they have not generated any concrete models in terms of risk conceptions that are priced. A recent attempt to apply this concept was Harvard economist Martin Weitzman, who uses some impressive mathematics to show that uncertainty in parameter estimates can explain the seemingly large equity risk premium. While in practice the distribution of future growth rates has its mean and variance calibrated to past sample averages, Weitzman shows that proper Bayesian estimation of uncertain growth parameters adds an irreducible fat-tailed background layer of uncertainty that can explain the seemingly large equity risk premium. In effect, the 17 percent annualized volatility of the stock market feels two or three times that because of the fat tails implicit in a nonergodic process like the stock market. But while this might get the equity risk premium into standard utility models, it also then implies that systematic volatility should be positively correlated with future aggregate returns using those same utility functions, which Figure 4.10 shows is not the case. That is, if uncertainty is positively correlated with volatility, which it almost certainly is, why is volatility, if anything, negatively correlated with returns across most assets? Why is aggregate volatility uncorrelated with future returns for an asset class like equities? Why is analyst disagreement \textit{negatively} correlated with returns?

Differences of opinion should proxy for parameter uncertainty, a perhaps better estimate of risk. Using analysts' earnings forecasts as a proxy for differences of opinion among investors, Karl Diether et al. (2002) find the quintile of stocks with the greatest opinion dispersions underperformed a portfolio of otherwise similar stocks.

Each month, they take stocks and sort them into five groups based on size (market cap), and then within these groups, sort again into quintiles on the basis of analyst forecast dispersion, as measured by the ratio of the standard deviation of analyst current fiscal year annual earnings per share forecasts to the absolute value of the mean forecast. They find that
the stocks with the higher dispersion in analysts' earnings forecasts earn significantly lower returns than otherwise similar stocks. Specifically, the highest dispersion group had a 9.5 percent annual return deficit over the 1980-to-2002 period.

Diether et al. note higher estimate dispersion is positively related to beta, volatility, and earnings variability, yet, because the returns go the wrong way (lower return for higher volatility), they note "our results clearly reject the notion that dispersion in forecasts can be viewed as a proxy of risk." Thus, in spite of being correlated with all things intuitively risky, like beta, volatility, and size, but uncorrelated with value or momentum, the correlation with returns proves it is not correlated with risk, because the one thing we know about risk is that it is positively correlated with returns.

**IPOs**

An initial public offering has a great deal of uncertainty, especially for an economist wishing to apply a factor sensitivity to it, because there is no time series. One usually applies a factor based on its characteristics, such as size, book or market, and perhaps industry. But without a track record, these assignments are highly uncertain in the Keynesian/Knightian/Ellsbergian sense. One would expect, given uncertainty aversion, for these stocks to have higher than average returns to compensate for this risk.

Jay Ritter at the University of Florida has a wonderful web site with data and articles on this issue going back 20 years. These data ignore the first day increase, because this one-day pop is only available to insiders with access to IPOs at the issuance price, a favor usually doled out by the big brokerages to their favorite customers. You simply cannot buy an IPO at the issuance price unless you are already paying a brokerage a lot of money in commissions or have in some way "paid" them a favor, so the relevant return for investors is from the close of the first day of issuance is used. Examining IPOs from 1970 to 2008, the geometric annual average return for these IPOs five years after issuance are 3.7 percent below size-matched firms. People who buy IPOs pay a premium, perhaps on the hope of buying the next Yahoo! or Google.

**TRADING VOLUME**

Another metric of disagreement is the amount of trading volume in a stock, normalized by its stock volume. A stock may turn over 50 percent of its stock a day, suggesting it is a useful factor proxy or there is a lot of disagreement about its prospects. On average, if disagreement leads to more buys and sells, we should see a higher return to these stocks. We do not, of course.
In the United States, since 1997, I created an index of stocks in the top 1,500 that had the highest trading volume, or shares, outstanding, every six months. The annualized geometric return for these high volume stocks was 1.6 percent, versus 9.7 percent for the low volume stocks.

**Volatility as Shorthand for Risk**

Introductory texts on risk and return, such as Brealey, Myers, and Allen’s popular MBA finance textbook, give the following examples as _prima fascia_ proof that risk generates return (Table 4.2).47

This is a typical overview, as Malkiel’s _Random Walk down Wall Street_ has a similar exposition. The implication, not mentioned explicitly but clearly implied, is that things that have higher volatility have higher risk, and _thus_ higher returns. In Brealey, Myers, and Allen, the explanation is merely that “portfolio performance coincides with our intuitive risk ranking,” though they do put the standard deviation numbers up there with the inevitable implication that risk and volatility are one and the same, or at the very least correlated. Malkiel mentions that higher returning stocks are “more variable.”

As we saw in Chapter 3, the excess return generated by small stocks is a fact, but it is not clear in what sense it is risky. Small stocks have greater volatility and a greater beta than regular stocks, but volatility and beta are not measures of risk, because if they were, there would be a massive arbitrage opportunity in many areas where systematic and idiosyncratic volatility actually generates negative return premiums. As to the very short end of the yield curve, the very high quality corporate bonds, to the extent these are proof of risk begetting return, fail to generalize within their own asset class. More highly volatile bonds, either because of increased default risk, or duration, do not generate higher average returns when extrapolated.

**TABLE 4.2** Typical Misleading Presentation of Risk and Return

<table>
<thead>
<tr>
<th></th>
<th>Return</th>
<th>Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Stocks</td>
<td>17.3%</td>
<td>33.4</td>
</tr>
<tr>
<td>Stocks</td>
<td>13.0</td>
<td>20.2</td>
</tr>
<tr>
<td>Corporate Bonds</td>
<td>6.0</td>
<td>8.7</td>
</tr>
<tr>
<td>Government Bonds</td>
<td>5.7%</td>
<td>9.4</td>
</tr>
<tr>
<td>T-bills</td>
<td>3.9%</td>
<td>3.2</td>
</tr>
</tbody>
</table>

Data from Brealey, Myers, and Allen’s MBA finance textbook (p. 154).
to B-rated, or 30-year bonds. The authors seem to be aware of this: They very conspicuously avoid saying directly that risk is volatility, because they know risk is not volatility or beta *per se*. That these three prominent researchers are involved in this sort of deception suggests a massive amount of cognitive dissonance. The intuitive description of risk is volatility, and this describes Table 4.2 pretty well, but when you dig deep, you see academics running from volatility as a metric of risk like the plague, although they don’t mind using it at the introductory level.

In the 1980s, you might see standard deviation and risk used interchangeably, as this followed directly from the graphs generated in the construction of the CAPM, and the efficient frontier. Return was on the vertical axis, standard deviation on the other. But now researchers are aware that the CAPM is an empirical failure of the first order, and so, while the standard deviation ($\text{vol}$) is often used as shorthand for risk, serious researchers know this is not the measure of risk relevant to expected returns.

If one’s education was unaware of utility functions, or Modern Portfolio Theory, one would have to look at these data and say that volatility is inversely correlated with returns. The theory that metrics of risk such as volatility or covariance is positively correlated with average returns fails spectacularly when applied to cross-sectional equities, movies, beta, developing country equities, U.S. time-varying equity returns, gambling, lotteries, options, financial leverage, financial distress, currencies, mutual funds, small businesses, analyst forecast dispersion, IPOs, and futures. These are not minor lacunae, but the heart of the risk-return theory. As Richard Feynman stated: “It doesn’t matter how beautiful your theory is, it doesn’t matter how smart you are. If it doesn’t agree with the experiment, it’s wrong!”