Behavioral Finance: Biases, Mean–Variance Returns, and Risk Premiums

Hersh Shefrin
Mario L. Belotti Professor of Finance
Santa Clara University
Santa Clara, California

Behavioral finance examines the biases that investors (individual and professional) incorporate in their investment decision-making process. These biases lead to inefficiencies in market pricing that are reflected in behavioral mean–variance portfolios. For an individual security, its risk premium reflects the extent to which its return co-varies with the return of a risky behavioral mean–variance portfolio.

Behavioral finance can shed light on many areas of investing, including valuation. In this presentation, I will discuss some of the key behavioral phenomena and how they relate to particular issues associated with analyst perceptions about returns (namely, representativeness and affect). I will then compare returns for what I call "behavioral mean–variance portfolios" with those of traditional mean–variance-efficient portfolios. In that respect, I will trace the importance of a concept known as "sentiment" and describe its impact on risk premiums in the market.

Because behavioral finance is complicated, it is worth keeping a few postulates in mind.[1] Behavioral finance postulates that in the short run, markets are inefficient and the potential exists for misvaluation. Behavioral finance also postulates that market values eventually revert to intrinsic values, although the long run might be very long. Therefore, investors need to be careful about looking at a stock that is overvalued and thinking that its price will quickly move to fundamental value. Instead, the stock price can move the other way: The mispricing can widen, and in some cases is likely to widen, before it narrows.

Representativeness and Affect Heuristic

Imagine that I flip a fair coin five times and the result is five heads. The question is, What are the odds of my tossing a tail on the sixth toss? The answer, of course, is 50 percent. The coin is fair.

But some people intuit a greater probability of a tail after so many heads in a row, even though the probability of tossing a tail is 50 percent because tails occur roughly half the time in any long stretch of coin tosses. Therefore, they think the odds favor a tail because they have not seen a tail for a while. One only needs to go to Las Vegas or Reno, Nevada, or Atlantic City, New Jersey, to see this phenomenon in place at any gambling table. This illusion is called "gambler's fallacy"—the tendency to overpredict reversals in situations of this sort.

Behavioral finance researchers believe that gambler's fallacy is a product of a psychological phenomenon called "representativeness." Representativeness refers to overrelying on stereotypes and is a principle that underlies particular rules of thumb that are used to arrive at judgments. (The term "heuristic" is a fancy name for a rule of thumb.) The stereotypical pattern for coin flips involving a fair coin is that heads appear half the time and tails appear half the time. Therefore, if a coin that is known to be fair is tossed several times in succession and a head has not appeared for a while, intuition

[1]For more information on behavioral finance, see Shefrin (2005a, 2005b).

This presentation comes from the Equity Research and Valuation Techniques conference held in Boston on 7–8 December 2006.

might suggest that a head is due. But in reality, a head is not due, and thinking that it is due is succumbing to gambler's fallacy.

In summary, representativeness involves overreliance on stereotypical thinking, and the stereotype for heads/tails is that 50 percent probabilities lead to half heads/half tails in terms of realization. This is why representativeness-based thinking leads some of us to succumb to gambler's fallacy when the odds are known.

Next, consider what happens when the odds are not known. In this case, representativeness-based thinking leads a person to try to uncover the process through observation. The person, in essence, begins to ask, for what process is the observed sequence representative?

Suppose now that instead of flipping coins, we are watching a basketball game. We are no longer observing a sequence of heads/tails; we are observing success rates for, say, three-point attempts. The concept of streaks (i.e., whether a particular player seems to be consecutively making or missing, in this case, three-point attempts) is well established among sports followers. So, the question becomes, How likely is Player X to make his next shot given that he has recently been hot (meaning that he has recently been successful in his three-point attempts)? The answer is that he is just as likely to make his next shot when he has been hot as he is when he has been cold.

That is the statistical fact of life in basketball, but it runs counter to the intuition of many—coaches, fans, and players, most of whom believe in hot hands, streaks, and slumps. But statistically, there is no such thing as a hot hand. This is not to say that we do not see streaks. We do see them, but the point is that they have no predictive power! The idea that streaks have predictive power is an illusion. The fact is, as a species, humans have very poor intuition about random processes, whether the process is coin tossing or whether it is three-point attempts in basketball. In particular, representativeness leads us to extrapolate recent performance.

In basketball, the hot-hand fallacy leads to predicting that a player will continue to be hot because his recent performance has been hot. More generally, the hot-hand fallacy involves predictions of unwarranted continuation when observing processes that are unknown. In addition to representativeness, "recency bias" is also at work in the hot-hand fallacy. Recency bias is the tendency to overweight the importance of more recent events relative to less recent events.

The biases just discussed are not restricted to gambling and sports. They also show up in the investment world. **Figure 1** shows two charts. Panel A is based on a UBS/Gallup survey and represents individual investors' expectations about returns on the S&P 500 Index for the next year (graphed on the y-axis) against what the S&P 500 return was the previous year (graphed on the x-axis). Panel B is based on Livingston Survey data and represents the same type of relationship but for investment professionals.

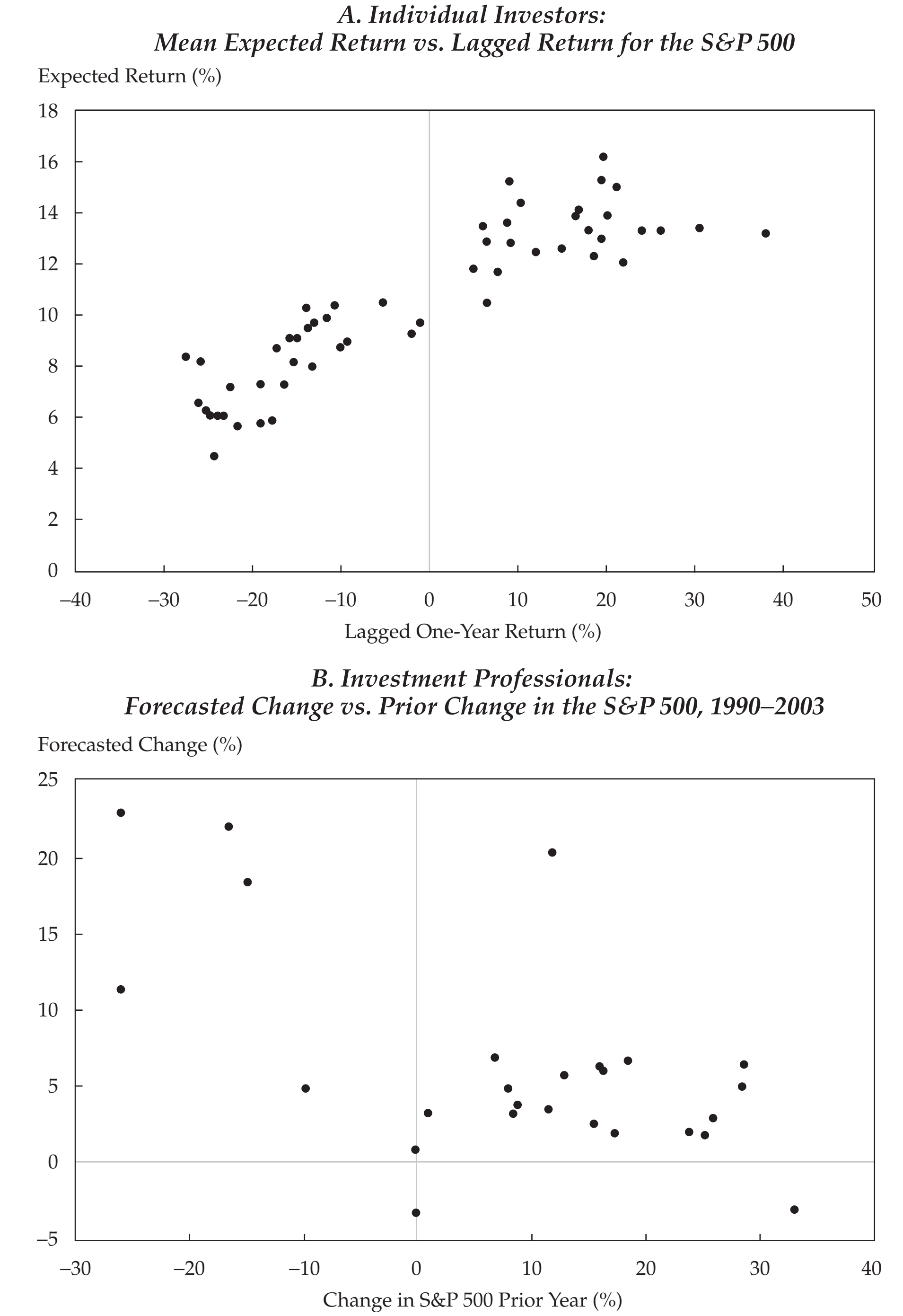
One does not need to display the regression lines to see the patterns. The regression line for Panel A would be positively sloped, and the regression line for Panel B would be negatively sloped. What does this mean? It means that when forecasting market returns, individual investors are prone to the hot-hand fallacy and that investment professionals are prone to gambler's fallacy. If I were to graph the relationship between actual year-over-year returns to the S&P 500, the slope would be fairly close to zero.

There is very little predictive power in last year's return to the S&P 500, but representativeness leads both investment professionals and individual investors to attach too much predictability to the previous year's returns. For most individual investors, the process generating the market returns is unknown, predisposing them to the hot-hand fallacy. For most investment professionals, experience with stock market history leads them to be familiar with the process generating the market returns, thereby predisposing them to gambler's fallacy.

To understand how representativeness works at the level of individual securities, consider two stocks—Unisys and Intel, described in **Table 1**. The information provided about these two stocks most likely would lead an investor to conclude that Intel has been a better company than Unisys for this period, particularly based on the past five-year sales, market cap, and retained earnings. Suppose that investors form their judgments about risk and return by relying on representativeness. They thus believe that better stocks feature higher expected returns. In addition, they associate safe stocks with the stocks of financially sound companies. They would tend to judge Intel as a better stock than Unisys because they would think Intel is a better company than Unisys. So, they would end up viewing Intel as featuring a higher expected return than Unisys, but they also would view Intel as being a safer stock than Unisys.

The implication is that investors see risk and return as being negatively related—not positively related as is taught in traditional finance. The positive relationship between risk and return is a cornerstone of modern finance. But when most investors form judgments in practice, they get the relationship upside down. The cornerstone principle of modern finance is not accepted at a gut level.

Figure 1. Expected Return vs. Lagged One-Year Return



Source: Reprinted from Shefrin (2005a, pp. 66 and 78) with permission of Elsevier.

And interestingly, I have found this pattern to be the case whether I am looking at investment professionals or individual investors.

The above is also an example illustrating the "affect heuristic," which is another way of saying that decisions are made based on gut instincts. The affect heuristic serves to reinforce representativeness by allowing snap intuitive judgments based on emotion. The basis for these judgments is a set of affective or emotional tags that are mentally attached to images, objects, and concepts. In this regard, imagery is important. For example, adding dot-com after the name of a company in the second half of the 1990s made a profound difference in the market value of the company because it affected investors' immediate emotional response to those companies.

Table 1. Unisys and Intel: Comparative Data from April 2000

Item	Unisys	Intel
Beta	1.33	1.04
Market value of equity (billions)	$8.003	$441.860
Book value of equity (billions)	$2.088	$ 36.103
Book-to-market equity	0.26	0.08
Balance sheet retained earnings (billions)	–$1.62	$ 25.22
Prior six-month return	–67.6%	215.6%
Prior one-year return	92.8	222.3
Prior three-year return	60.2	56.2
Past five-year growth rate of sales	0.8	21.2

Source: Reprinted from Shefrin (2005b, p. 58) with permission of McGraw-Hill/Irwin.

The affect heuristic is a mental shortcut that people often use to judge benefits and risks. If I like something, I buy it; if I do not like it, I do not buy it or I short it. What drives this behavior is a subconscious coding system in our brains called "affective labeling." We know as a practical matter that benefits are associated with positive affect and that risks are associated with negative affect—and not just for stocks: When psychologists are looking at affect, risk and benefits tend to be negatively related. What is interesting, however, is that this general propensity carries over to stocks when there is a cornerstone academic principle that says the relationship should go the other way.

Risk and Return Perceptions

The traditional view is that expected return on a stock is determined by a four-factor model (return on the market, return differential between small-cap and large-cap stocks, return differential between value and growth stocks, and momentum). Individual investors associate low book-to-market value and high market capitalization with both good stocks and good companies. In addition, investors view stocks with low betas, large market caps, and low book-to-market values as being less risky than stocks with high betas, small market caps, and high book-to-market values. Therefore, investors have a good sense of what makes up risk, but they have a poor sense of how to connect that risk to expected returns.

Unlike individual investors, analysts treat the relationship between risk and expected return as being positive. Using beta as a measure of risk, analysts view the relationship as positive but nowhere nearly as strongly positive as academic textbooks suggest. Holding beta constant, analysts expect smaller-cap stocks to earn higher returns than larger-cap stocks, which, in view of the historical record, is reasonable. But analysts also expect growth stocks to earn higher returns than value stocks. Historically, this is not the case. Finally, analysts' target prices tend to be excessively optimistic.

A clear, strong, persistent, well-documented momentum effect exists in returns. But analysts establish target prices as if they believed in short-term reversals, not momentum. Given the strong evidence for momentum, this belief in short-term reversals means that analysts are prone to exhibit gambler's fallacy. If you want to know what distinguishes outperformers from underperformers among money managers, it is how they handle momentum. Parenthetically, the target-price time horizon is an issue here. For horizons up to a year, short-term winners tend to outperform short-term losers. But for horizons as long as three years, extreme past losers tend to outperform extreme past winners.

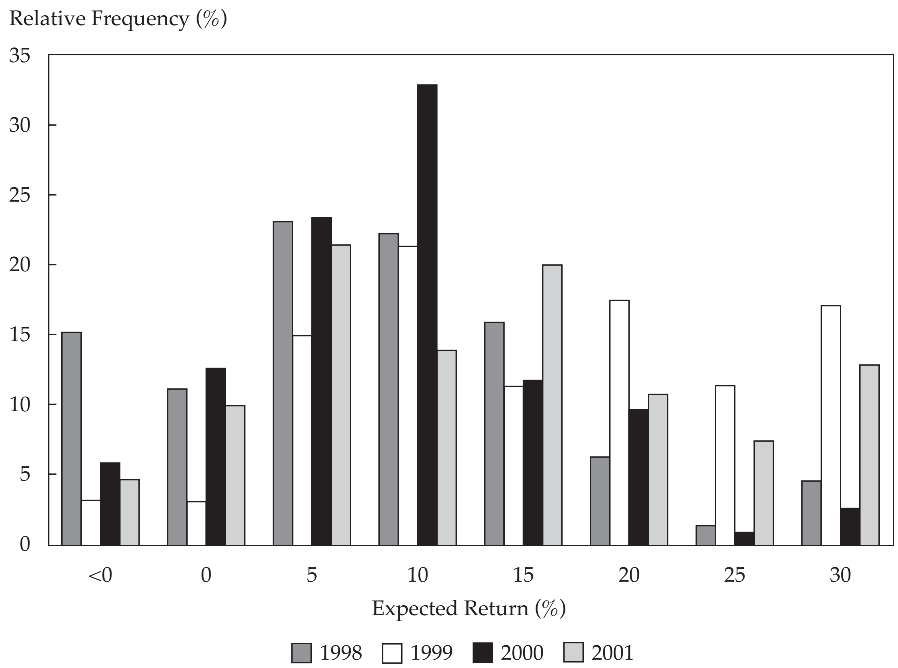
Investor Sentiment

In any view of the markets, opinions will differ with respect to such key variables as EPS and target prices. The question then becomes, Do these differences of opinion affect security prices, or are they instead self-canceling?

Figure 2 shows the results of a survey of expected returns for the period from 1998 to 2001. Note that it is an aggregate of individual and professional investor expectations. So, for example, for the year 2000, approximately 33 percent of investors expected a 10 percent return in that year. At the same time, about 23 percent expected a 5 percent return. Thus, the expected returns were quite variable. Furthermore, the distribution shifted over time. In 1998, the distribution shows a fat tail on the left. This left-hand tail gets smaller over the sample period as the right-hand tail grows larger. And a small right tail in 1998 gets fat in 1999, shrinks, and then gets fat again in 2001.

I see these data as being the basis for market sentiment. Market sentiment can be thought of as aggregate market error. Some people may talk about the sentiment of individual investors, but when sentiment is talked about in any meaningful sense, it pertains to what the aggregate market error is. The distribution in Figure 2 is multimodal. It has thick tails with clusters of beliefs on the extremes. Expected returns in the data are not all in the middle. Instead, one can see clusters of optimism and pessimism at very high and very low returns.

Figure 2. Distribution of Average of Survey Expected Returns, 1998–2001



Source: Reprinted from Shefrin (2005a, p. 218) with permission of Elsevier.

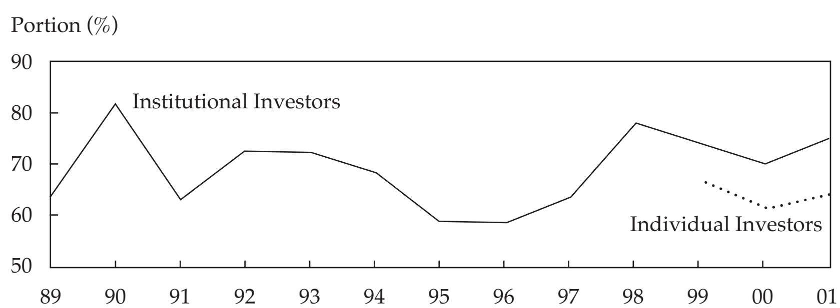
Evidence shows clustering by investors—both a cluster of overconfident optimists and a cluster of underconfident pessimists. In this regard, consider **Figure 3**, which represents the results of a survey conducted by Robert Shiller. Shiller's survey asked investors whether they believed the chance of a stock market crash in the next year exceeded 10 percent. The number of survey participants who answered "yes" was astonishingly high, suggesting that the market contains a great many underconfident pessimists.

Mean–Variance Returns

The nature of investor sentiment has profound implications for how security prices and risk premiums are determined. In this regard, traded options play an important role because options permit investors to bet on particular ranges of returns. Thus, investors have more flexibility than they would have if they were simply holding or shorting individual securities or market indices. These bets manifest themselves through options prices and have a significant impact on mean–variance-efficient returns, which are the basis for risk premiums.

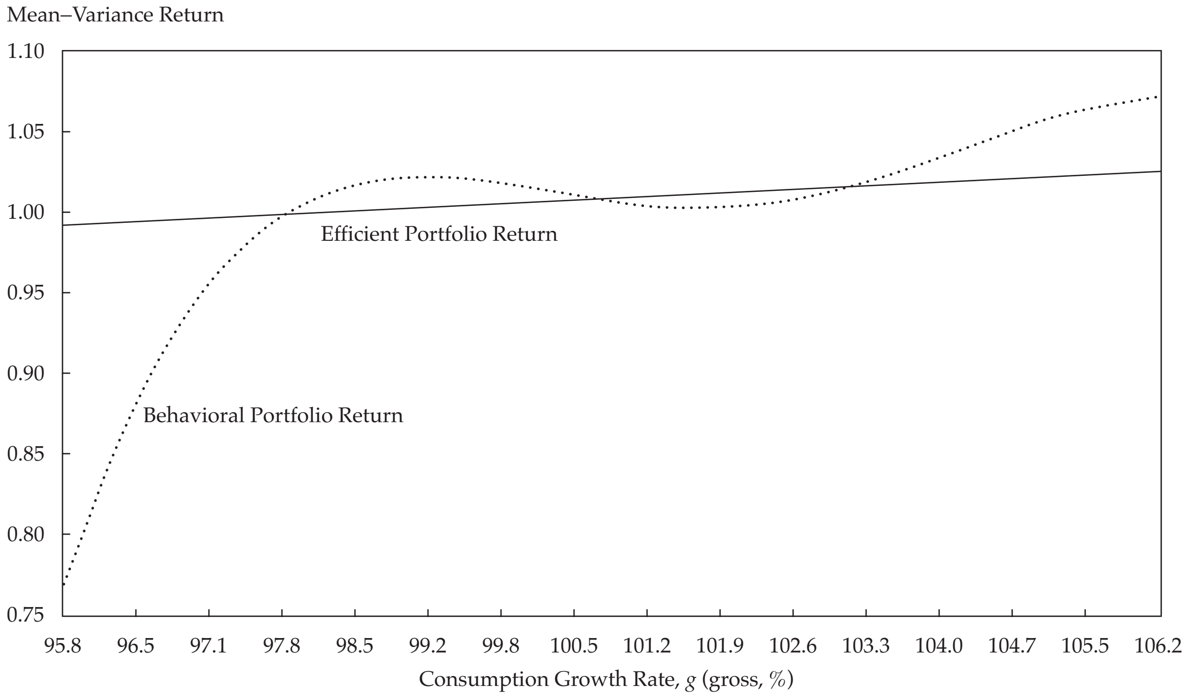
To demonstrate the impact that sentiment can have in contrast to a traditional mean–variance portfolio return pattern, consider **Figure 4**. This figure depicts how the gross return (per quarter) to a mean–variance portfolio varies with the growth rate of the underlying economic fundamentals. Here, gross

Figure 3. Time Series of Percentage of Investors Expecting a Crash Based on Shiller Crash Confidence Index, 1989–2001



Source: Reprinted from Shefrin (2005a, p. 360) with permission of Elsevier.

Figure 4. Gross Return to a Mean–Variance Portfolio: Behavioral Mean–Variance Return vs. Efficient Mean–Variance Return



Source: Reprinted from Shefrin (2005a, p. 245) with permission of Elsevier.

return is 1 + net return. For example, a gross return of 1.10 represents a 10 percent return. The solid line represents the traditional mean–variance return pattern. It can be thought of as representing the return pattern of a portfolio that is a mix of the risk-free security and the market portfolio in a market that is free of sentiment. This particular mean–variance portfolio is quite conservative, with the weight associated with the risk-free security being heavy. The strange, oscillating, lumpy and bumpy curve (the dotted line) displays the return pattern associated with a behavioral mean–variance portfolio. Its return pattern is relatively conservative in the middle portion and is not "too different" from its traditional counterpart. But the return pattern for a behavioral mean–variance portfolio is very different at the extremes, with behavioral mean–variance returns being more volatile than those associated with the traditional mean–variance portfolio. (The choice of a conservative traditional mean–variance portfolio is intended to highlight the nature of the incremental behavioral mean–variance volatility.)

The peaks and the valleys in a behavioral mean–variance-efficient portfolio actually represent opportunities to exploit mispricing. The very low returns at the left of Figure 4 occur because bearish investors who overestimate volatility bid the price of out-of-the-money index put options above their efficient levels. The high returns at the right of Figure 4 occur because all investors, including bullish investors who underestimate upper-tail volatility, bid the price of out-of-the money index call options below their efficient levels. Investors who hold true mean–variance-efficient portfolios will seek to exploit these inefficiencies by shorting out-of-the-money index puts and going long out-of-the-money index calls. Thus, a mean–variance-efficient portfolio that is attempting to exploit these inefficiencies will perform badly if a lower-tail event occurs and perform well if an upper-tail event occurs.

But despite the differences between the composition of traditional and behavioral mean–variance portfolios, it is possible to apply traditional mean–variance concepts in a behavioral world. For example, it is possible to compute a behavioral beta for a security. To do so, simply compute the covariance between the return to the security and the return to a behavioral mean–variance-efficient portfolio. Then, divide by the return variance of the behavioral mean–variance-efficient portfolio to obtain a behavioral beta. From a behavioral beta, it is an easy step to get to a behavioral capital asset pricing model (B-CAPM). Expected return for an individual security can be found by adding the equilibrium risk-free rate to the product of behavioral beta and the risk premium on the behavioral mean–variance-efficient portfolio.

If you think about a bell-shaped distribution, the right tail and the left tail are mirror images of each other. They are symmetrical—a situation called "zero skewness." But some distributions, like those associated with the payoff of lottery tickets, do not have this payoff pattern. Suppose you hold a lottery ticket that will pay you $100 million with a probability of 1 in 50 million if you win or $0 if you lose. The expected payoff for the ticket is $2. The length of the left tail of the payoff distribution is a mere $2 (the difference between the expected value and the left-side payoff). In contrast, the length of the right tail of the payoff distribution is long: $100 million – $2.

The distribution for the lottery payoff is said to be positively skewed because the right tail is longer than the left tail. By the same token, the distribution for shorting an out-of-the-money index put option is negatively skewed because, although the option will expire unexercised most of the time, a small probability of a crash exists, in which case the investor who shorted the put will lose a lot of money—more than the premium received from shorting the option. Thus, in Figure 4, the net returns associated with growth rates lower than those displayed on the *x*-axis of the graph can fall below 100 percent.

According to the shape of Figure 4, because of its negative skew, a positive premium must be associated with the risk of shorting out-of-the-money index put options. Think of it this way: The potential for large losses (negative skew) represents an unattractive feature to investors. To induce investors to invest in situations with negative skew, a large positive risk premium must be expected. Therefore, the risk premium for a stock will include a component whose magnitude depends on the extent to which the stock's return mimics the negatively skewed return pattern from shorting index puts.

The co-skewness of a stock measures the impact its return distribution has on the already negative skewness of a behavioral mean–variance portfolio. Co-skewness has a beta-like formula that involves the squared market return instead of the market return. The reason for measuring co-skewness is to adjust for the quadratic-like pattern in the left and right extremes of the mean–variance pattern depicted in Figure 4.

Individual investors form behavioral portfolios, not mean–variance portfolios. Behavioral portfolios reflect the needs for downside protection and upside potential (again indicating the importance of the co-skewness statistic). In a behavioral portfolio, individuals protect their downsides by avoiding negatively skewed returns. They meet their needs for upside potential by seeking securities that offer positively skewed returns. In other words, the behavioral portfolios that individual investors select are not behavioral mean–variance portfolios—in fact, just the opposite.

The problem is that individual investors do not find the idea of shorting index puts very attractive. It is not even attractive for many investment professionals who can afford to do it. That is one reason why a positive risk premium is associated with shorting index puts and a negative risk premium is associated with lottery-like stocks (whose return patterns feature positive skewness).

As far as analysts are concerned, the key issue in this discussion is estimating risk premiums for securities, not constructing portfolios. The point of the discussion is that the risk premium for a security actually has a component that reflects the degree to which the stock mimics the return pattern from shorting index puts. Therefore, if you are trying to understand what determines the risk premium for a security, you will want to measure its co-skewness because the co-skewness indicates how much of the mean–variance left-tail exposure displayed in Figure 4 you will add to a portfolio when you include the stock in that portfolio.

The reason that risk premiums reflect co-skewness is analogous to the reasoning used to explain risk premiums by using the CAPM. In the CAPM, the risk premium for a stock is determined by how much market risk the stock adds to a well-diversified portfolio. The measure of how much market risk a stock adds is beta. But we also want to measure how much mean–variance skewness the stock adds to a properly diversified portfolio, and for that, we measure co-skewness, which, I remind you, has a beta-like formula.

Much of what is known about the pricing of skewness comes from Harvey and Siddique (2000). They found that the correlation between co-skewness and mean returns of portfolios sorted by size, book-to-market equity, and momentum is a remarkable and surprising –0.71. Thus, much of the explanatory power of size, book to market, and momentum may derive from co-skewness—an explanation that is consistent with the behavioral approach.

The CAPM explains only 3.5 percent of cross-sectional returns. Notably, the addition of co-skewness brings that up to 68.1 percent, just shy of what the regular linear three-factor Fama–French model explains—that being 71.8 percent. Moreover, co-skewness alone can explain a portion of excess returns that the Fama–French model cannot. Within the Fama–French framework, the correlation between unexplained return and co-skewness is 0.53, which is large. Consistent with the behavioral approach, the momentum effect also has an explanation in terms of co-skewness. Recent winners feature lower co-skewness than recent losers.

The above findings, taken together, provide strong support for a behavioral asset-pricing theory in which the basis for risk premiums is the behavioral mean–variance frontier.

Conclusion

Many investors form judgments that suggest that they mistakenly believe that risk and return are negatively related. In contrast, the judgments underlying analysts' target prices are consistent with risk and return being positively related. At the same time, analysts appear to make some errors in judgment when setting target prices. In particular, analysts establish target prices as if they expected growth stocks to outperform value stocks and recent losers to outperform recent winners. But the historical evidence shows that value stocks outperform growth stocks and recent winners outperform recent losers.

Behavioral mean–variance portfolios embody investor errors and feature negatively skewed return patterns. Therefore, the risk premium for a stock features a component that reflects how much "mean–variance skewness" the stock contributes to a well-diversified portfolio. The measure of a stock's contribution to mean–variance skewness is called its "co-skewness." The more negative the co-skewness is, the higher the stock's risk premium will be.

This article qualifies for 0.5 PD credits.

References

Gilovich, Thomas, Robert Valone, and Amos Tversky. 1985. "The Hot Hand in Basketball: On the Misperception of Random Sequences." *Cognitive Psychology*, vol. 17, no. 3 (July):295–314.

Harvey, Campbell R., and Akhtar Siddique. 2000. "Conditional Skewness in Asset Pricing Tests." *Journal of Finance*, vol. 55, no. 3 (June):1263–1295.

Shefrin, Hersh. 2005a. *A Behavioral Approach to Asset Pricing*. Burlington, MA: Elsevier Academic Press.

———. 2005b. *Behavioral Corporate Finance*. New York: McGraw-Hill/Irwin.

Question and Answer Session

Hersh Shefrin

Question: For a basketball player with a "hot hand," does your analysis take into account the skill set of this player relative to the constant 50/50 probability when flipping a fair coin?

Shefrin: The difference between the coin and the basketball player is that the success probability for the basketball player might be 68 percent, not 50 percent. The issue, however, isn't the probability. The issue is whether that basketball player's probability of making his or her next basket stays at 68 percent regardless of whether he or she has been "hot" or "cold" during the current game.

Should we continue to use the historical average of 68 percent when we're trying to predict what's going to happen next with this particular basketball player? Or should we attempt to adjust the historical average for the most recent performance? The original study by Gilovich, Valone, and Tversky (1985) using data from the Philadelphia 76ers (and since replicated for many other teams and sports) shows that you should not adjust the historical averages.

Question: Six flips of a coin may all come up heads, yet the odds on the next flip remain 50/50. So, somewhere along the way, doesn't there have to be more tails than heads coming up?

Shefrin: You need to understand the law of large numbers. If you mark down a 1 for each head and a 0 for each tail and then divide that sum by the number of coin tosses, the result will approach one-half for a large number of flips. It is not that some sort of a correction process is occurring; it is simply that the impact of the first six flips becomes miniscule.

Question: How do we use our awareness of the biases identified by behavioral finance to become better money managers?

Shefrin: The particular biases that I have been discussing are so well documented because they are hardwired into our brains. There was a time in our history as a species when these biases helped us survive. They were well suited to an earlier environment. It is just that they are not as well suited to the modern environment.

The problem is that whenever a conflict exists between how we think and how we feel, how we feel wins. We need to develop ways to counter our emotions. One way to do that is to work in groups where there is recognition within the group of these biases and a willingness to work together to mitigate the effects of these biases.