The Journal of Behavioral Finance
2003, Vol. 4, No. 4, 234–237



# Regression to the Mean: One of the Most Neglected but Important Concepts in the Stock Market

**Bernard I. Murstein**

*The meaning of "regression to the mean" is discussed, as well as the consequences of failing to recognize its effect on research. The lack of performance persistence among stocks and mutual funds is explained as evidence of a lack of valid variance in the performance of stocks, resulting in steep regression-to-the-mean effects. The ubiquity of regression to the mean is illustrated by showing that it is an important factor in marriage as well as in mutual funds.*

I have seen many references to the concept of regression to the mean, sometimes called reversion to the mean, but I have not seen any articles or books explaining what causes it. Indeed, some books make it appear that stocks have a moral obligation to descend from their lofty perch and head toward the mean to give some other stock a chance in the sun.

The true explanation may be more prosaic, but it is an important piece of information that investors need to succeed. To illustrate the concept, I use as an example a study I read in a professional journal a half century ago.

## A Prime Example

The study concerns a psychologist working with a group of children then called mentally retarded (I can't recall the citation). The author hypothesized that monosodium glutamate (MSG), a flavor enhancer added to soup and other foods, would raise the IQs of these students. Accordingly, he picked those with the lowest IQs, and administered doses of MSG to them. He then retested the students with a parallel form of the first test. To his delight, he found a significant elevation of IQs in this group, and was able to publish his results.

## Hypothetical Replication of Study

To understand what is wrong with this study, let's do a hypothetical replication. Begin with the assumption that every intelligence test score contains two components: one reflecting the student's true ability, and an error component. Error is assumed to be randomly distributed, and intelligence tests are fairly reliable. Reliability in this situation is the likelihood that a test will show consistency when given in parallel form on two occasions. The higher the reliability, the smaller the error component.

Suppose we have a class of 200 students, all with the same true ability, and we select the twenty with the lowest test scores in the group. But in selecting this group on the basis of their low scores, we have capitalized only on negative error, because in this class every student has the same true ability. Thus, whenever we *select a group because of its extremely high or low scores*, we capitalize on chance factors. The reason for extremely high or low scores is that random error operating helped some students, who became the high group, and hurt others, who became the low group.

Figure 1 shows the error scores (not the test scores) on the initial intelligence test. We can see that the average negative effect of sampling error for the students in the low group was a loss of 4 points (the left shaded side of the figure), with a range from 3 to an undetermined high number.

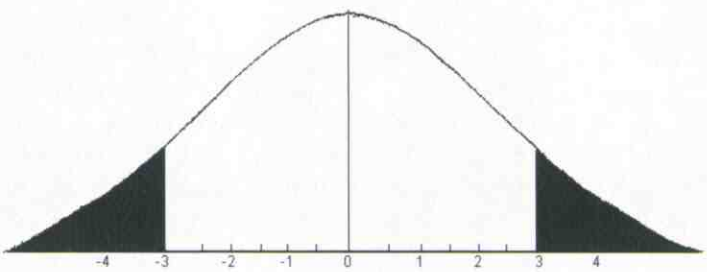
If we give the students a new perfectly parallel test and assume no practice effects, we would expect their mean error score on the new test to be zero, because error is randomly distributed. In the long run, the positive and negative errors should balance each other out. If the students were given MSG prior to the retest, it would appear that the low group gained an average of 4 points (from –4 to 0). From an error score of –4, they had regressed completely back to the mean for the group, 0. Remember that we are using only error scores, because we have stipulated that all 200 students have the same true ability.

The students with scores above the mean (on the right shaded side of the figure) had a mean score of +4. They also would have an expected error score of 0 because they regressed to the mean from the opposite direction. They would appear to be less smart on the second test. However, these students were absent from the actual study, so we could not observe this regression.

**Bernard I. Murstein** is May Buckley Sadowski Professor Emeritus of Psychology at Connecticut College.

Requests for reprints should be sent to Bernard I. Murstein, 46 Beacon Hill Drive, Waterford, CT 06385–4110. Email: bimur@conncoll.edu

**FIGURE 1**
**Hypothetical Distribution of Error for the Class on Intelligence Test Scores Showing Upper and Lower Deciles of Error**




The amount of error is a function of the reliability of the test. The more reliable the test, the less the actual test score will depart from a person's true score. In this sample, the amount of error is small: Only the upper decile (above the ninetieth percentile) had 3 or more points of "favorable" error. Those in the lowest decile (below the tenth percentile) had 3 or more points of "unfavorable" error (the average error for the high group was +4, and –4 for the low group).

Of course, real-life students don't all have the same true ability, so the lowest students would not regress all the way back to the mean. Their very low scores would be a combination of truly low ability and "unlucky" error variance. The amount they regress would depend on the amount of error in the test.

In short, there is no proof that the MSG caused the gain in IQ scores because the author capitalized on random error in choosing the students for the study. If he had *randomly* chosen a sample of 100 students and given them MSG, and given another *randomly* selected group no MSG, we would have found no regression to the mean and probably no significant difference between the means. If we did find a change in IQ or history score, however, we might have to consider MSG more seriously as a performance booster.

## Research on Stocks Illustrating Regression to the Mean

This example may be interesting, but what about the stock market? Consider Table 1, which shows how various mutual funds reported by Morningstar Mutual Funds on April 1, 1994, fared from March 1984–March 1994. Note that all the funds above the mean in 1989 were below the mean in 1994 and vice versa, which implies some rather interesting shifts. In the earlier example, when we did not assume all students had the same true ability, we did not predict that the upper half of the distribution would change places with the lower half. We only predicted that each group would regress somewhat to the mean. However, we were dealing with intelligence tests, which are fairly reliable.

In the Morningstar example, the fact that such an extreme switch occurred implies a lot of random error affecting fund performance, and probably only a small amount of true ability involved. With this assumption, dramatic shifts in performance are easily imaginable. Today's top dog may be tomorrow's underdog.

To show that Table 1 is not atypical, consider Table 2. Professor Malkiel (In the Vanguard, 2001) used a sample of 283 equity funds studied from 1990–1994 and identified the twenty with the highest returns. These twenty funds outpaced the S&P 500 by an annual average of 9.2%. He recalculated the returns of these funds for 1995–1999, and found that these formerly stellar funds trailed the S&P 500 by an annual average of more than 2% from 1995 to 1999.

**Table 1.** *Comparison of Returns for Mutual Funds by Classification*

| Objective | Five Years to March 1989 | Five Years to March 1994 |
|---|---|---|
| International Stocks | 20.6% | 9.4% |
| Income | 14.3% | 11.2% |
| Growth and Income | 14.2% | 11.9% |
| Growth | 13.3% | 13.9% |
| Small Company | 10.3% | 15.9% |
| Aggressive Growth | 8.9% | 16.1% |
| Average | 13.6% | 13.1% |

*Source:* Morningstar Mutual Funds, April 1, 1994.

**Table 2.** *A Shift in Performance: Early 1990s versus Late 1990s*

| Fund | Rank in 1990–1994 | Rank in 1995–1999 |
|---|---|---|
| A | 1 | 129 |
| B | 2 | 134 |
| C | 3 | 261 |
| D | 4 | 21 |
| E | 5 | 210 |
| F | 6 | 53 |
| G | 7 | 183 |
| H | 8 | 105 |
| I | 9 | 275 |
| J | 10 | 54 |

Table 2 shows how the top ten funds of the 1990–1994 period ranked over the 1995–1999 period.

Morningstar assigns funds from one to five stars to based on a historical balance of risk and return. The best funds receive five stars and the worst receive one star. The funds ranked "1" at the start of 2000 returned 6.9% over 2000, compared to –15.7% for the five-star funds.

Comparable results occurred in Jegadeesh and Titman [1993]. These authors compared the 10% of highest-returning stocks from the preceding six months with the 10% poorest-returning stocks, as shown in Figure 2. This figure represents the difference between the highest- and poorest-returning stocks during the succeeding thirty-six months. The portion of the curve above the abscissa represents the tendency of the highest-returning stocks to continue to outperform the poorest-returning ones; the portion below the abscissa represents the tendency of the poorest-returning stocks to outperform the highest-returning ones.

The results over thirty-six months are intriguing. Winners, as shown in Figure 2, followed a momentum path, continuing to rise and reaching their apogee about seven to eight months after earnings announcements. Thereafter there was a steady drop, crossing into negative territory after sixteen months, and proceeding down to the time the study ended at thirty-six months after the initial announcements. The highest-returning stocks lost ground to the poorest-returning stocks, but did not necessarily become losers (not shown in the graph). Likewise, the poorest-returning stocks often gained on the highest-returning stocks, but did not necessarily become winners. Both seemed to be regressing toward the mean.

These studies show that the future cannot be predicted very accurately from the past. With so much change in fund performance, we can conclude that *differences in the performance of various mutual funds and clusters of stocks are due largely to chance.* With so much error, investors chasing after the top mutual funds or clusters of stocks based on last year's performance are wasting their time. Their best hope is to focus on the funds with the lowest expenses, rather than those with the best records.

Pzena [1995] depicts the reversion to the mean for return on equity (ROE) for the S&P 500 over a five-year period divided into quintiles (see Figure 3). Note that each quintile continues to hold its rank with respect to ROE, but by Year 5 the spread has dramatically narrowed, making the highest ROE group a net loser, and the lowest group a net gainer.

Why these mispricings take place is not completely clear. Overconfident judgments in complex situations may be one possibility. Heuristic biases are another. Analysts may judge certain companies to be "good" or "bad," and it may take time for them to realize their mistakes. Thus, they continue to be surprised for several quarters in a row, before realigning their thinking.

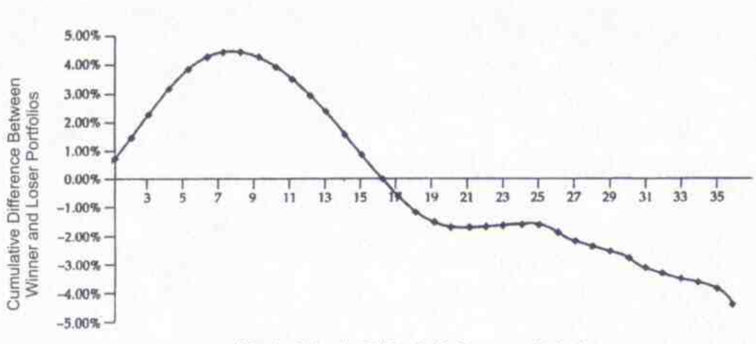
Dreman and Lufkin [1997] recently demonstrated that mispricing is not due to a change in fundamentals (e.g., growth in earnings, sales, cash flow, profit margin, or ROE). They also find that investors and analysts tend to overreact or underreact to news, and only later do they tend to correct these exaggerations. Investors typically fail to recognize that extremely good or bad news capitalizes on chance factors that cannot be sustained in the future.

In sum, mispricing is often the result of investor inability to differentiate between a company's true strengths and chance factors that inflate its future prospects. There is also a certain amount of pack-following, and, in the case of mutual funds, chance events may be attributed to fund managers' management skills.

## Mispricing in Non-Stock Environments

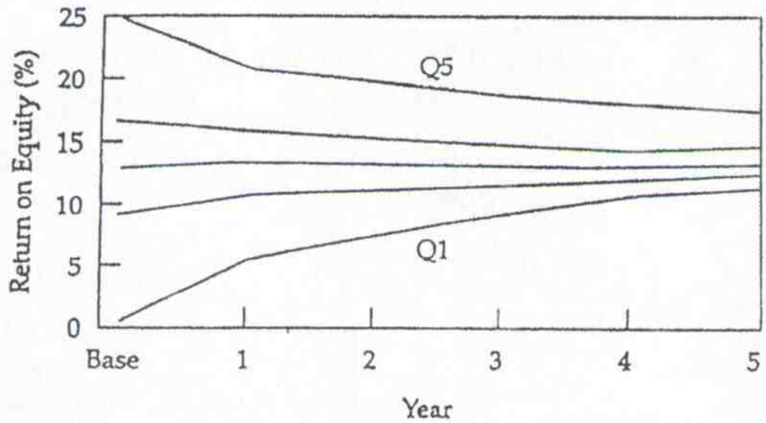
In a broad sense, mispricing is ubiquitous. When a private company goes public it has an initial public of-

**FIGURE 2**
**Cummulative Difference Between Winner and Loser Portfolio Rates of Return at Announcement Dates**




Source: Jegadeesh, N. and S. Titman, S. Returns to Buying and Selling Losers: Implications for Market Efficiency. *Journal of Finance*, 48, (1993), pp. 65–91.

**FIGURE 3**
**Reversion to the Mean by S&P 500 Quintiles**




Source: Sanford C. Bernstein & Co.

fering (IPO). Companies usually wait for the best possible moment to go public: when demand seems to be greatest, when fortune seems to smile on them, or when earnings are at their highest. In short, companies capitalize on favorable error variance to make their futures seem brighter. But the truth is that the average IPO considerably underperforms all the indexes. The average investor doesn't realize this because losing IPOs are all but ignored by fund managers, while winners are publicized endlessly for both profits and ego.

As noted in Murstein [2002], the IPO process is very similar to the search for a marital partner, which can include regression to the mean and mispricing! When a couple is dating, each person attempts to put his or her best foot forward (positive error), so they can appear to have greater desirable assets than they truly have (Murstein [1976]). Moreover, individuals tend to marry when their compatibility seems highest. If their compatibility is less than optimal, they either break up or continue to date without moving toward marriage or cohabitation. Because they often misperceive the true character of their partners, they are setting themselves up for a regression-to-the-mean effect, when a more accurate perception of their partners replaces the idealized, distorted portrait.

Countless studies have indicated that the first ten years of marriage represent a negatively decelerating curve of happiness, which reaches its nadir about a decade after the wedding. I believe marital happiness is dependent on the discrepancy between the expectations for marriage and their perceived fulfillment. Many couples cannot live up to the high expectations engendered during courtship. Regression to the mean in this situation can result in bitter disappointment, or a lowered, more realistic appraisal of the partner.

In similar fashion, many companies cannot live up to the expectations raised during their IPO period, when everything was going their way and investors expected the "honeymoon" to last forever. The fundamentals of these companies don't necessarily change dramatically, but investor expectations don't allow for bad luck. Failure to live up to unrealistic expectations leads to severe punishment for stocks, just as it does for many marriages. The dream of quick and eternal wealth cannot survive the more mundane realities of 10% per year earnings growth. The jilted public seeks retribution and the market price of disappointing companies can be savaged.

In sum, to overcome the regression-to-the-mean effect, astute investors must overcome the hyperbole surrounding new investment fields and try to separate out unbiased empirical data. Ignoring the selection process when investing in "high-flying" stocks can lead to a rude awakening.

## References

Dreman, D.N., and E.A. Lufkin "Do Contrarian Strategies Work Within Industries?" *Journal of Investing*, 6, (1997), pp. 7–29.

*In the Vanguard*. "Investing for the Future? Don't Fixate on the Past." 6, Summer 2001.

Jegadeesh, N., and S. Titman. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *Journal of Finance*, 48, (1993), pp. 65–91.

Morningstar Mutual Funds, April 1, 1994.

Murstein, B.I. *Who Will Marry Whom? Theory and Choice in Marital Choice*. New York: Springer, 1976.

Murstein, B.I. *Getting Psyched for Wall Street: A Rational Approach to an Irrational Market*. Leawood, KS: Cypress Publishing Group, 2002.

Pzena, R.S. "Behavioral Biases and Investment Research." In A.S. Wood, ed., *Proceedings of the AIMR Seminar Improving the Investment Decision-Making Process*. Marina del Ray, California, April 4, 1995.