Luck, Skill, and Investment Performance

BRADFORD CORNELL

BRADFORD CORNELL
is a visiting professor of finance at California Institute of Technology in Pasadena, CA.
bcornell@hss.caltech.edu

Successful investing, like most activities in life, is based on a combination of skill and serendipity. Distinguishing between the two is critical for forward-looking decision making because skill is relatively permanent, while serendipity, or luck, by definition, is not. An investment manager who is skillful this year presumably will be skillful next year. An investment manager who was lucky this year is no more likely to be lucky next year than any other manager.

The problem is that skill and luck are not independently observable. Instead, all that can be observed is their combined impact, which in this article is called performance. The central question, therefore, is to determine how much can be learned about skill by observing performance. It turns out that a straightforward way to investigate that question exists and is based on application of the bivariate normal distribution. Though the results presented in this article are well known in statistics, they are not commonly applied in the context of assessing portfolio managers. As I will show, they can serve as the basis of a simple and useful model for assessing the skill of competing fund managers. To illustrate how the model works, I use the procedure to analyze the performance of large-cap equity managers tracked by Morningstar. It turns out, as one might expect given the volatility of asset prices, that the relative performance of managers in any given year provides little information about management skill.

A SIMPLE MODEL FOR ASSESSING LUCK AND SKILL

To develop the model, assume that a measure of performance, p, exists that reflects the sum of skill, s, and luck, L. This formulation has a straightforward interpretation in terms of portfolio management. In this context, p represents the observed return on a specific portfolio, s represents the added expected return due to the skill of the investment manager, and L represents the impact of idiosyncratic risk on the portfolio's return over the observed holding period.

More specifically, assume that both luck and skill are normally distributed in the cross section and that

$$p = s + L \tag{1}$$

where $s \sim n(E(s), sd(s))$ and $L \sim n(0, sd(L))$. By definition, the mean of the luck distribution is zero. Because $p = L + s$, p and s are distributed as bivariate normal with mean vector $[E(s), 0]$ and covariance matrix

$$\begin{matrix} sd(p) & \mathrm{corr}(p,s) * sd(p) * sd(s) \\ \mathrm{corr}(p,s) * sd(p) * sd(s) & sd(s) \end{matrix}$$

Because luck cannot be correlated with skill—otherwise, luck would have a predictable component—it follows that

$$sd(p) = sd(s) + sd(L) \text{ and } \mathrm{corr}(p, s) = sd(s)/sd(p) \quad (2)$$

For the bivariate normal distribution, it is well known from the statistical literature that[1]

$$E(s|p) = E(s) + \mathrm{corr}(p,s) * sd(s)/sd(p) * [p - E(p)] \quad (3)$$

Substituting the relations from Equation (2) into Equation (3) gives

$$E(s|p) = E(s) + [\mathrm{var}(s)/\mathrm{var}(p)] * [p - E(p)] \quad (4)$$

Because $E(L) = 0$, it follows that $E(p) = E(s)$, so Equation (4) can be written

$$E(s|p) = E(s) + [\mathrm{var}(s)/\mathrm{var}(p)] * [p - E(s)] \quad (5)$$

Equation (5) is the basic model. It states that when performance is observed in excess of the mean, the assessment of skill is adjusted upward, but not all the way to the observed level of performance, p. Instead, the assessment of s is adjusted upward by the observed superior performance, $p - E(s)$, times the ratio of the variance of s to the variance of p. Therefore, the assessment of skill, based on the observation of performance, depends critically on $\mathrm{var}(s)/\mathrm{var}(p)$.

Notice that in both limiting cases, Equation (5) makes intuitive sense. If $\mathrm{var}(L)$ is much larger than $\mathrm{var}(s)$, then $\mathrm{var}(p) >> \mathrm{var}(s)$, in which case $E(s|p)$ goes to $E(s)$. That is reasonable, on the one hand, because if performance is dominated by luck, then observation of performance should play little role in the assessment of skill. On the other hand, if $\mathrm{var}(s) >> \mathrm{var}(L)$, then $\mathrm{var}(s)$ is approximately equal to $\mathrm{var}(p)$, which implies $E(s|p)$ goes to p. That makes sense because if luck has a relatively minor impact on performance, then observed performance is a precise measure of skill.

Equation (5) has numerous applications in finance and is the basis for the phenomenon referred to as regression toward the mean. Regression toward the mean occurs because whenever the measure of performance, p, differs from the mean it indicates two things. First, it indicates that the above average performance represents above average skill. Second, it indicates that the above average performance reflects good luck. In other words, above average performance is evidence of *both* good luck and superior skill. However, whereas the skill element is permanent, the luck element is transitory. Therefore, the expected performance next period reverts back toward the mean from p because the luck variable has an expected value of zero. The greater $\mathrm{var}(L)$ relative to $\mathrm{var}(s)$, the larger the regression back toward the mean.

What in this article is called luck can be interpreted as random measurement error in other contexts. Consider, for instance, the problem of estimating beta. In such a case, skill is equivalent to the unobservable true beta and performance is the estimate of beta. For betas, we know that the $E(s)$ is 1.0. Therefore, Equation (5) states that the best estimate of beta is not the observed regression coefficient (measured by p), but $E(s|p)$, which is a weighted average of the estimated regression coefficient and the overall mean of 1.0. Based on this property, Merrill Lynch developed a widely adopted weighted-average procedure for estimating beta. The task here, however, is not to estimate beta, but to assess the contributions of luck and skill in determining investment performance. In the next section, I consider the application of Equation (5) in that context.

APPLICATION OF THE MODEL TO MUTUAL FUND DATA

To illustrate the evaluation procedure, I apply it here to data on mutual fund performance. But before turning to the data, there is one important caveat. The model attributes investment performance exclusively to the combination of skill and luck. When the performance measure, p, is interpreted as the return on a portfolio, a third element comes into play, namely, the risk level of the portfolio. There are two ways to account for this. One approach is to perform the calculations in terms of risk-adjusted returns, which introduces the problem of deciding how to adjust for risk, a problem that the finance profession has not fully resolved after 40 years of research. The other approach is to perform the analysis on a comparable cohort of investment funds which is the approach I take in this article.

The data are drawn from a comprehensive 2004 Morningstar database of mutual fund performance.[2] To ensure that the data are relatively homogenous, the sample is limited to funds that invest primarily in U.S.

EXHIBIT
Cross-Sectional Statistics for Large-Cap Value Funds

Annual Data for Period Ended March 2004		
Number of Funds	Mean Return	Standard Deviation
1,034	25.35%	5.47%
Data for 15 Years Ended March 2004		
Number of Funds	Mean Return	Standard Deviation
341	10.03%	1.57%

Source: Morningstar.

large-capitalization value stocks. The Morningstar database includes performance data on 1,034 large-cap value funds during 2004. The first line of the exhibit presents cross-sectional summary statistics for the returns on these funds, which serves as the measure of performance, p. Also shown in the exhibit is the mean return of 25.02% and the standard deviation across the 1,034 funds of 5.47% for the year 2004.

To apply Equation (5), it is also necessary to estimate the standard deviation of s, which is more difficult because s is not directly observable. Two distinct approaches are available for overcoming this difficulty. The first approach is to rely on judgment rather than on specific data. For example, an investor may conclude that the stock market is sufficiently competitive that differences in skill among large-cap value managers should lead to no more than a 200 basis point differential from the mean for the vast majority of funds. That judgment translates into a standard deviation of s on the order of 1.0%. If the standard deviation of s is taken to be 1.0%, then the ratio of var(s) to var(p) is 0.033. That ratio implies that the observation of annual performance should have virtually no impact on the assessment of the relative skills of the 1,034 large-cap value managers. This result is largely consistent with a large body of literature on mutual fund performance beginning with the classic work of Jensen [1968] and continuing through the work of Nitzsche, Cuthbertson, and O'Sullivan [2007].

An alternative approach is to use long-run return data to estimate the standard deviation of s. If the fund return data were stationary and the sample period sufficiently long, then the cross-sectional standard deviation of s could be estimated with little error. Unfortunately, neither is the case. On the one hand, with a maximum sample period of 15 years, luck still plays a role in determining the cross-sectional distribution of returns. This random noise results in an upwardly biased estimate of the standard deviation of s. On the other hand, the 15-year sample is also impacted by survival bias. Whereas annual data are available for 1,034 funds, the 15-year sample contains only 341 funds. Because the funds that disappear from the sample are likely to be underperformers, both the mean 15-year return and the cross-sectional standard deviation are likely to be overstated. Given that the calculation presented here is only illustrative, no attempt is made to adjust for either of these offsetting effects on the estimated standard deviation of s.

The second line of the exhibit presents the cross-sectional summary statistics for the 341 funds in the 15-year sample. The mean annual return for the sample is 10.03% and the standard deviation is 1.57%. A standard deviation of 1.57% for s implies that the ratio of var(s) to var(p) is 0.082, which indicates that approximately 92% of the cross-sectional variation in annual performance is attributable to random chance.

CONCLUSION

The simple model I have presented provides a useful, practical tool for assessing the impact of skill and luck on portfolio performance. When the model is applied to a sample of large-cap value managers, the results indicate that most of the annual variation in performance is due to luck, not skill. This finding is consistent with findings reported in other papers on mutual fund performance. Nonetheless, the model described in this article provides another way of analyzing performance data.

The analysis also provides further support for the view that annual rankings of fund performance provide almost no information regarding management skill. Potential investors are better advised to consider the stated investment philosophies of competing firms than to rely on such rankings. In any event, at best, only minor revisions in estimates of skill should be based on annual performance data.

ENDNOTES

I would like to thank Wayne Landsman, Steve Stubben, and Elizabeth Tito for helpful comments on earlier drafts of this article. Any errors remain my own.

[1]See Mood [1974].

[2]The Morningstar historical data were graciously provided by Wilshire Associates.

REFERENCES

Jensen, M.C. "The Performance of Mutual Funds." *Journal of Finance*, Vol. 23, No. 2 (1968), pp. 389–416.

Mood, A. *Introduction to the Theory of Statistics*. New York: McGraw-Hill, 1974.

Nitzsche, D., K. Cuthberson, and N. O'Sullivan. "Mutual Fund Performance." Working paper, SSRN, 2007.