

# Stewardship As A Factor In The Financial Performance Of Mutual Funds

David P. Boyd, (E-mail: d.boyd@neu.edu), Northeastern University  
Mustafa Yilmaz, (E-mail: m.yilmaz@neu.edu), Northeastern University

## ABSTRACT

*Morningstar has assigned a Stewardship grade to 1165 mutual funds. This comprehensive construct can showcase funds that serve as stewards of investor capital and therefore warrant public trust. This study suggests that such funds may also preserve and augment shareholder return.*

## INTRODUCTION

When New York Attorney General Elliott Spitzer first leveled charges against some mutual fund firms in 2003, he exposed clandestine chicanery that would profoundly alter the public perception of asset managers and American business in general. Spitzer's clarion call precipitated a move toward heightened regulatory oversight and foreshadowed an emphasis on financial stewardship.

Through a strategy of "market timing," outside investors like Canary Capital Partners colluded with portfolio managers to make money trading in mutual funds. William Galvin, the Massachusetts Secretary of State, remarked that "letting hedge funds buy shares in mutual funds is like putting a shark in a goldfish tank" (Sloan, 2003, p. 2). Another fund manager who exhibited egregious behavior was Gary Pilgrim. He used an outside hedge fund as a vehicle to bet against the very mutual fund that he managed. As the fund fell 65 percent over a 20 month period, Pilgrim made \$3.9 million betting against it. Financial commentator Allan Sloan (2003) likened the situation to a captain who plowed his ship into an iceberg and then escaped in a private lifeboat to collect life insurance on his drowning passengers.

According to the Investment Company Institute (Jacobe, 2003), nearly one-half of U.S. households hold mutual funds. In the parlance of Secretary Galvin, mutual funds have become "the bank of necessity for middle Americans" (Sloan, p. 2). For such families mutual funds serve as fiscal and psychic bedrocks for encumbering tuition reserves and augmenting retirement streams. With the erosion of many traditional pension plans and the tenuous trajectory of social security, such funds assume pivotal prominence.

Feeling financially compromised and morally violated, fund holders reacted with alacrity. In a twelve-month period they snatched \$21 billion from Janus and \$40 billion from Putnam. Repugnance soon morphed into retribution. In a Business Week survey (2003) respondents averred that jail time should be meted out for executive malfeasance. The trading scandals pervasively impacted shareholder sentiment, with an Investment Company Institute (2004) survey highlighting reputation as the most pronounced component in public assessment of funds. According to a TIAA-CREF Survey about financial services firms, investors associated the word "trust" with "integrity" (73%) more than with "competence" (24%) (Roper Public Affairs, 2005). Not surprisingly, the press resounded with calls for a federal fiduciary standard that would assert the primacy of investors and curtail the legerdemain of managers.

In due course such changes came. Investors were not asking their assets managers and advisers to forego opportunity for enhanced remuneration. Rather they were urging performance-based fees so that advisers would make money only as a function of clients' expanded holdings. The Securities and Exchange Commission has agreed with the groundswell for fee clarification, ruling that fund companies must disclose the formulaic determinants of pay. Is pay based on performance that can benefit both the client and the firm? Or is it based on assets under management

that benefit only the firm? Most allegations of illegality have involved hedge funds. To Spitzer the pathway to propriety is thus clear. First block hedge funds from investing in mutual funds. Also prohibit firms from allowing hedge funds and mutual funds to fall under the same panoply.

Increased scrutiny has engendered more scrupulous behavior. According to a Morningstar (Young, 2006) study, mutual fund fees have declined for two consecutive years. Mutual funds are now seen as “a more ethical and transparent investment choice” (Pender, 2005). Total assets under management at mutual funds have ballooned to a record high of \$9.5 trillion (Young, 2006). A Stanford University study suggests that investors are no longer fleeing the scandal-tainted funds though neither are droves flocking to them (Greenberg, 2004).

## **METHOD**

An abiding legacy of the mutual fund scandal is the sanctity of stewardship. John Bogle, founder of Vanguard Mutual Funds, coined the phrase “stewardship quotient” in a Motley Fool interview (Gardner and Gardner, 2003). In calculating such a quotient, Bogle stipulated that the foremost criterion should be the fund’s cost. Morningstar subsequently seized the gavel and constructed an operational definition for “how seriously the mutual fund company takes its fiduciary responsibility to mutual fund investors” (Subramanian, 2004, p. 1). Morningstar views fund managers as stewards of investors’ capital. Such an approach accentuates whether the interests of the fund’s managers and board align with those of fund shareholders. Specifically, Morningstar’s Stewardship letter grades assign an “A” through “F” to mutual funds based on the aggregation of five components. Determination depends on both quantitative and qualitative information distilled by Morningstar analysts. The scale expressly incorporates some intangibles that are deemed important inputs to the investment decision.

Morningstar identifies Regulatory Issues as the first component of the Stewardship Grade. Have any fiduciary breaches occurred within the last three years and, if so, have steps been taken to remediate problems and forestall transgressions? Board Quality is the second component. Are independent directors in place? Does the board oversee a manageable number of funds rather than risk diligence due to a plethora of funds? Morningstar relishes boards whose independent directors invest in the funds they serve; these directors have “skin in the game.”

Manager Incentives serve as the third component. An emphasis on long-term performance rather than asset growth bolsters the fund’s grade. In addition, Morningstar welcomes portfolio managers who invest in their own funds. It is especially partial to managers who invest \$1 million and/or more than one third of their liquid net worth in the funds they run. Morningstar believes that fund managers who “eat their own cooking” will outperform abstemious counterparts (Spence, 2006). Recent studies lend support to Morningstar’s conviction that “skin in the game” enhances fund performance (Cremers et al., 2005; Laise, 2006). For the next component of Fees, Morningstar assesses relative expense by referencing a compendium of comparison groups. Additionally, it is expected that firms will reduce their expense ratios even as they grow their asset base.

The final component is Corporate Culture. This rubric subsumes a range of considerations. Above all the cultural ethos should champion long-term performance rather than chase “du jour” trends. Plaudits accrue to funds that close when they become too large - as 300 have done since 2000 (Young, 2006). According to Morningstar, should a fund’s asset base become too large, sheer size can dilute good ideas. Morningstar likes redemption fees that discourage churning but abhors 12b-1 fees that impeded client returns. Shareholder communications should always be timely and transparent, and the firm should eschew soft dollars. Behind the strong culture of successful families lies a “secret sauce” for investment performance (Dolan, 2006). For example, low cost, expansive research or technological resources can bring competitive advantages in the field.

To generate Stewardship grades, Morningstar analysts evaluate every fund along these five components, awarding component points in increments of 0.5. With the exception of Regulatory Issues, the minimum score a fund can receive in each component is 0 and the maximum is 2. For Regulatory Issues, the lowest possible score is –2. For each component, these points are described in qualitative terms as follows:

Excellent = 2 points  
 Good = 1.5 points  
 Fair = 1 point  
 Poor = 0.5 points  
 Very Poor = 0 or fewer points

The qualitative terms correspond to a single point score for each component. The sole exception is the Very Poor rating for regulatory issues where scores can range from 0 to -2. The overall Stewardship grade is obtained by summing the five component scores, and then generating letter grades according to the following designations of total point intervals:

A: 9 - 10 points  
 B: 7 - 8.5 points  
 C: 5 - 6.5 points  
 D: 3 - 4.5 points  
 F: 2.5 or fewer points

On its website Morningstar publishes the above qualitative terms for component evaluations and letter grades for Stewardship. Morningstar does not provide numeric point scores for individual components.

To test the face validity of the Stewardship construct, letter grades were converted to the customary numeric scale with A=4.0, B=3.0, C=2.0, D=1.0 and F=0.0. Component ratings were similarly coded on a 5-point numeric scale (we used a scale from 1 to 5, but a scale from 0 to 4 or any other fixed translation could have been used without affecting the results). Then a multiple regression model was constructed with Stewardship grades as the response variable and the five component ratings as explanatory variables. For reference, the resulting equation is shown below (the first number under each coefficient is the corresponding *t*-statistic, and the second number is its *p*-value or significance).

$$\text{Stewardship} = -2.050 + 0.280 \text{ Regulatory} + 0.223 \text{ BoardQual} + 0.235 \text{ Incentives} + 0.234 \text{ Fees} + 0.252 \text{ Culture} \quad (1)$$

|                         |        |       |       |       |       |       |
|-------------------------|--------|-------|-------|-------|-------|-------|
| ( <i>t</i> -statistics: | -37.66 | 27.07 | 16.82 | 31.95 | 33.75 | 17.63 |
| ( <i>p</i> -value:      | 0.000  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

This model explained 88.8 percent of the variability in Stewardship grades ( $n = 1165$ ,  $r^2 = 0.888$ ,  $p < .001$ ). The contribution of each component is highly significant and approximately equal to the other components. By a slight margin Regulatory Issues emerges as the highest contributor to Stewardship grade. These results indicate that the interpretation of the letter grades as the summative representation of five qualitative components appears to be reasonable and consistent, even without taking into account the specific manner by which Morningstar generates the letter grades.

Before conducting further analyses, the authors noted that assignment of Stewardship letter grades entails some loss of information conveyed by total stewardship points. For example, all funds with total Stewardship scores of 7.0, 7.5, 8.0, or 8.5 are assigned the same letter grade B. However, total Stewardship scores can be reconstructed with relative ease and complete accuracy from published qualitative ratings of the five components, except for funds that may have been originally assigned negative points for regulatory issues. Consequently, most of our analyses used numeric points for the five components and overall Stewardship rather than qualitative terms and letter grades. To do so, the regulatory issues score was assumed to be 0 for all funds rated Very Poor on this component. Among 1054 funds for which some annual return data were available, only 23 were rated Very Poor and among all 1165 funds listed with letter grades in May, 2006, only 27 were rated as such (return data were missing for 4 of these 27 funds). Thus, reconstructed Stewardship scores were accurate for almost all of the analyzed funds. As a consequence our results were not influenced in any significant way by the inclusion or exclusion of these funds in the data set. Since the Stewardship score was *defined* as the sum of five component values, these components explained 100 percent of the variability in Stewardship scores when used in a multiple regression model for Stewardship, as expected. These numeric distillates, of course, cannot be perfectly precise given the possible predisposing views of different analysts

for the various funds. While acknowledging the imprecision that might arise from analyst bias, the authors concede that Morningstar made a reasonable effort to describe its methodology and the authors are comfortable taking the ratings for what they are.

## **RESULTS AND DISCUSSION**

Having ascertained the utility of Morningstar’s protocol, we sought to explore the relationship, if any, between Stewardship grade and financial performance. Morningstar asserts that “good stewardship is good business practice” (Dolan, 2006) since a credible reputation will draw assets. A fund’s Stewardship grade should therefore serve as an input for investment decisions. When combined with other Morningstar commentary, these letter gradients “can help determine the difference between a great investment and one to avoid” (Morningstar Fact Sheet, 2004, p. 1). Based on these assumptions, the authors decided to test the hypothesis that the higher the Stewardship grade, the better the fund’s financial performance. As a measure of financial performance to use in the analysis, we considered using either the 3-year average return for 2003-2005 or the 5-year average return for 2001-2005. We ultimately settled on the former as a more reasonable compromise. First, numerous funds have been in existence for only a few years, and 5-year annual data are thus not available for them (111 funds). Second, the period 2001-2005 includes highly tumultuous times after the Internet bust and the 9/11/2001 terrorist attacks. These exogenous factors are presumed to be extraordinary in occurrence. Third, three-year horizon is the same as that used by Morningstar in assessing regulatory issues.

Clearly, the financial performance of a fund may depend on its type. For example, a Large Growth stock fund can be expected to yield higher returns than a Municipal bond fund. In its rating protocol, Morningstar uses more than 30 categories for classifying fund types, such as convertibles, large growth, long bond, moderate allocation, and specialty. In our analyses, we used two main categories of funds, namely Stock and Bond funds, represented as the variable S1\_B0. We assigned the value 1 for a stock fund and 0 for a bond fund. A total of 860 stock funds and 265 bond funds were classified in this manner. We excluded 40 funds whose type was not specified by Morningstar. Stock funds were further grouped into Growth, Value, Blend, and Other, with the last category consisting of specialty funds, emerging markets, and foreign funds. Bond funds were sub-categorized as Corporate, Government, and Municipal. Convertibles and bank loans were included in the Corporate bond category. No further sub-categorization was used to distinguish short, intermediate, and long-term bond funds.

Summary statistics for the main study variables are shown in Table 1 below. Mean values of the variables are similar for stock and bond funds with the conspicuous (and expected) exception of average 3-year return (20.15 percent for stock funds versus 6.11 percent for bond funds). The difference between the mean values of stock and bond funds also seems somewhat more significant for Incentives. On this variable, stock funds were rated somewhat higher than bond funds though the average rating in both groups is no better than “Fair.” To explore the existence of a relationship between Stewardship and fund performance, we first considered a regression model of 2003-2005 average return (AvRet3yr) against the Stewardship score (Stewardship) using all funds in the entire sample ( $n = 1098$  and 67 funds had missing return data). The result is summarized below:

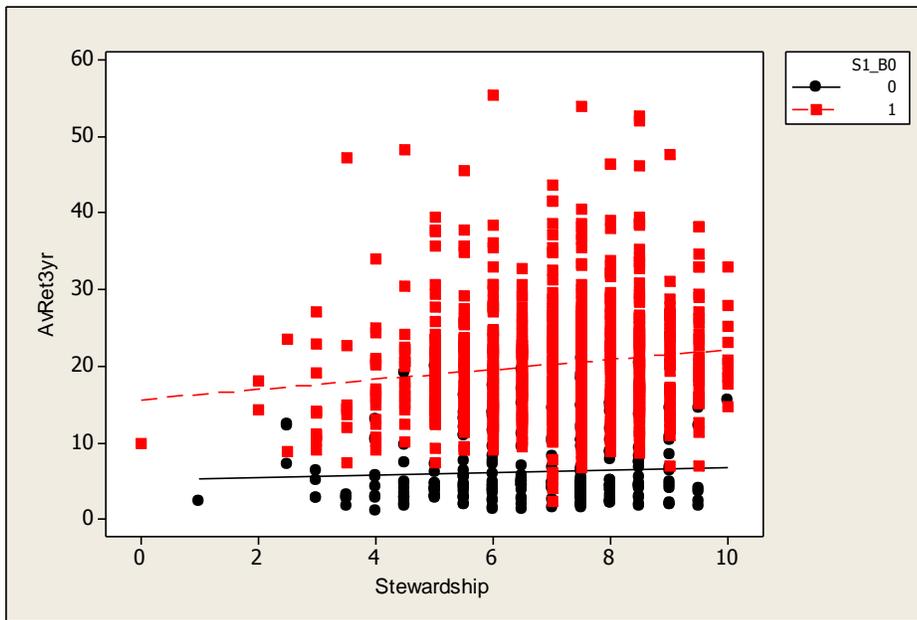
$$\begin{aligned} \text{AvRet3yr} &= 11.3 + 0.796 \text{ Stewardship} && (2) \\ (\text{t-statistics: } &9.08 \quad 4.52) \\ (\text{p-values: } &0.000 \quad 0.000) \end{aligned}$$

This model explained only 1.8 percent of the variability in average 3-year return ( $r^2 = 0.018$ ,  $p < .001$ ). Although small from a practical perspective, this statistically significant result is not a mere artifact of a large sample. According to the simple model above, a 1-point increase in the Stewardship score would be expected to lead to about a 0.8 percent increase in average return. To see which of the five component variables of Stewardship are responsible for its correlation with AvRet3yr, we used a multiple regression model where Stewardship was replaced with the five components it entails. In this model, Incentives and Culture were the only significant explanatory variables ( $t = 2.57$  and  $t = 2.35$ , respectively).

Table 1 Summary Statistics For Stewardship And Fund Performance Variables

| Variable             | Stock Funds<br>(n=860) |                 | Bond Funds<br>(n=265) |                 |
|----------------------|------------------------|-----------------|-----------------------|-----------------|
|                      | Mean                   | Std. Dev.       | Mean                  | Std. Dev.       |
| Stewardship          | 6.98                   | 1.54            | 6.70                  | 1.61            |
| Regulatory Issues    | 1.68                   | 0.52            | 1.62                  | 0.58            |
| Board Quality        | 1.43                   | 0.38            | 1.37                  | 0.34            |
| Incentives           | 1.04                   | 0.59            | 0.88                  | 0.55            |
| Fees                 | 1.41                   | 0.64            | 1.48                  | 0.65            |
| Culture              | 1.42                   | 0.46            | 1.35                  | 0.49            |
| Aver. Return 2003-05 | 20.15<br>(n=828)       | 7.59<br>(n=828) | 6.11<br>(n=263)       | 4.40<br>(n=263) |

Figure 1 Scatterplot Of Average 3-Year Return Versus Stewardship Scores



The simple regression model (2) above does not distinguish between stock and bond funds, an obvious shortcoming in view of the large difference in their average returns as seen in the last row of Table 1. The effect of Stewardship might become murky if it is considered by itself without regard to other potentially pertinent variables, such as the basic characteristic of the funds. To account for fund type, we included the corresponding variable S1\_B0 in the model with the following result ( $n = 1091$  and 74 funds had missing data):

$$AvRet3yr = 2.60 + 0.525 \text{ Stewardship} + 13.9 \text{ S1\_B0} \tag{3}$$

(*t*-statistic: 2.62 3.91 28.34)  
 (p-value: 0.009 0.000 0.000)

This model explained 43.5 percent of the variability in average 3-year return ( $r^2 = 0.435$ ,  $p < .001$ ), indicating that whether a mutual fund is a stock or bond fund is an important factor in explaining the variability in average return. The difference in the average returns of stock and bond funds Table 1 is mostly explained by the variable S1\_B0, but the contribution of Stewardship beyond fund type remains significant, at least from a statistical

perspective. These findings are also evident in the Figure 1 scatterplot of AvRet3yr versus Stewardship using S1\_B0 as a grouping variable:

The slight difference between the slopes of the regression lines in the graph suggests that there might be a small interaction effect between Stewardship and S1\_B0. When an interaction term was included, the model was not significantly improved. Also, the exclusion of some possible outliers in the data set (such as the two points near the lower left corner of the graph) did not discernibly alter results.

To assess the effects of various types of stock funds in explaining average return, we regressed AvRet3yr using data for stock funds only, including Stewardship as an explanatory variable along with various combinations of the dummy variables for Growth, Value, Blend, and Other (with no more than three of them in any model to avoid extreme multicollinearity). These results showed that average returns from Growth and Value funds did not differ significantly. Blend funds had somewhat lower returns (approximately 0.8% lower) while Other funds had somewhat higher returns (approximately 1.1% higher).

Similarly, we ran a regression model for AvRet3yr using only bond fund data, including the Corporate and Municipal dummy variables along with Stewardship. The result was as follows ( $n = 263$  and 2 cases had missing data):

$$\begin{aligned} \text{AvRet3yr} &= 3.27 - 0.082 \text{ Stewardship} + 5.89 \text{ Corporate} + 1.59 \text{ Municipal} \\ (\text{t-statistics: } & \quad 3.06 \quad -0.57 \quad 8.67 \quad 2.29) \\ (\text{p-values: } & \quad 0.002 \quad 0.567 \quad 0.000 \quad 0.023) \end{aligned} \quad (4)$$

This model explained 30.3 percent of the variability in AvRet3yr ( $r^2 = 0.303$ ,  $p < 0.001$ ) with Corporate bond funds yielding the highest average returns, followed respectively by Municipal and Government bonds. Interestingly, Stewardship does not emerge as a significant variable in explaining the performance of bond funds. Fund type appears to be much more important. This finding can also be gleaned from the almost horizontal slope of the regression line for bond funds in Figure 1. When Stewardship was excluded from the last regression model above, there was negligible degradation in the coefficient of determination ( $r^2 = 0.302$ ). A similar situation applied when the three bond fund types were considered separately: again Stewardship was not a significant explanatory variable for average 3-year return. Therefore, a key conclusion is that the Stewardship variable is relevant for Stock funds but not for Bond funds. The fixed decision parameters of bond funds no doubt limit discretionary choices for portfolio managers. Equities customarily exhibit greater variability in return than their bond counterparts. Within stocks' wider range Stewardship could exert a more causative effect on return.

## CONCLUSION AND FUTURE RESEARCH

It is worth emphasizing that while our findings appear to be statistically robust, the absolute numbers are modest. Several comments should be made. First, financial returns are influenced not just by fund management, but also by many external factors such as market conditions, interest rates and geopolitical dynamics. Second, in the highly crowded and competitive financial markets, explanatory increments of even a small magnitude can propel a firm past benchmark peers. Finally, with today's trend toward transparency, readily available information about Stewardship will increasingly affect investor proclivities.

Morningstar's new Stewardship Quotient identifies funds where fiduciary concern is foremost. Should a fund face regulatory or managerial issues, it can suffer a credibility loss that diminishes customer draw. The present study demonstrates that the relationship between Stewardship and financial return is real even though Stewardship is not the main determinant of financial performance. Investors increasingly incline toward funds that project a reputation for integrity. This study suggests that such funds may also augment shareholder return, however modestly. If subsequent research confirms this result, then trusted firms will further their appeal in the marketplace. Besides the prospect of enhanced return, funds with a high Stewardship quotient appear to preserve capital and can therefore imbue investors with peace of mind.

Since Morningstar assessments are retrospective, a salient issue remains untested: Will the factors that galvanize historical performance also serve as prognosticators of success? Longitudinal designs will be necessary to address this question. Financial markets are complex mosaics where multiple forces weave strands. If fiduciary responsibility is part of this tapestry, then it can help investors find safe passage amid turbulent seas.

## REFERENCES

1. Business Week, The Mutual Fund Scandal and You, *BusinessWeek Online*, September 30, 2003.
2. Cremers, Martijn, Driessen, Joost, Maenhout, Pascal J., and Weinbaum, David, Does Skin in the Game Matter? Director Incentives and Governance in the Mutual Fund Industry, EFA 2005 Moscow Meetings, September 2005. Paper Available at SRRN: <http://ssrn.com/abstract=676983>.
3. Dolan, Karen, How Strong Is Your Mutual Fund's Corporate Culture, *Morningstar.com*, June 29, 2006.
4. Gardner, David and Gardner, Tom, What's Your Fund's Stewardship Quotient, *The Motley Fool.com*, November 12, 2003.
5. Greenberg, Gregg, Time Helps to Heal Fund Scandal Wounds, *TheStreet.com*, August 31, 2004.
6. Investment Company Institute, Shareholder Sentiment of the Mutual Fund Industry, *Fundamentals*, 13, 4, pp. 1-4, October 2004.
7. Jacobe, Dennis, Will Scandal Hurt Mutual Funds, *Gallup Poll Briefing*, December 23, 2003.
8. Laise, Eleanor. Another Way to Assess a Mutual Fund, *The Wall Street Journal*, July 26, 2006.
9. Morningstar, *Fact Sheet: The Morningstar Stewardship Grade for Funds*, Morningstar, Inc., Chicago, Illinois, 2004.
10. Pender, Kathleen, Mutual Funds Cleaning House, *San Francisco Chronicle*, January 2, 2005.
11. Roper Public Affairs, *Trust in America 2005: The TIAA-CREF Report*, GfK NOP, New York, New York, September 2005.
12. Sloan, Allan, The Mutual Fund Scandal: Unfair Fight, *Newsweek*, MSNBC.com, December 8, 2003.
13. Spence, John, Fees Are Falling, as Are Old Habits, *The Wall Street Journal*, July 3, 2006.
14. Subramanian, Sam, Mutual Funds: What Investors Need to Know about Morningstar Fiduciary Grades, AlphaProfit Investments, September 29, 2004.
15. Young, Lauren, Mutual Funds: Back in Action, *BusinessWeek Online*, June 26, 2006.

NOTES