11-2014

Governance Matter: Morningstar Stewardship Grades and Mutual Fund Performance

Jerry Xiaping CAO
Singapore Management University, jerrycao@smu.edu.sg

Aurobindo GHOSH
Singapore Management University, AUROBINDO@SMU.EDU.SG

Jeremy GOH
Singapore Management University, jeremygoh@smu.edu.sg

Wee Seng NG
National University of Singapore

Governance Matter: Morningstar Stewardship Grades and Mutual Fund Performance[⊥]

Jerry Cao
Singapore Management University
Jerrycao@smu.edu.sg

Aurobindo Ghosh[1],
Singapore Management University
Aurobindo@smu.edu.sg

Jeremy Goh
Singapore Management University
Jeremygoh@smu.edu.sg

Wee Seng Ng
National University of Singapore

This draft: November 1, 2014

[⊥] This research was partially funded by the Center for Corporate and Investor Responsibility, Sim Kee Boon Institute for Financial Economics, SMU. We also gratefully acknowledge the Stewardship data provided by Morningstar Incorporated for the part of the data used for the project. The authors would like to thank the conference and seminar participants at the University of Paris at Dauphine, 2012 ESEM, Malaga, Spain, 2012 Finance Scholars Symposium, Singapore, 2013 IFC, IIM-Calcutta, 2013 Asian FMA, Shanghai, China, 2014 European FMA, Maastricht, The Netherlands. The authors also acknowledge the Best Paper award for fund management the paper received in 2014 at the 5th Financial Markets and Corporate Governance Conference, QUT, Brisbane, Australia sponsored by Victoria University, New Zealand. Usual Disclaimers hold. Previous version of the paper was circulated under the title "Grades matter in Performance: Morningstar Stewardship Grade and Mutual Fund Performance."
[1] Corresponding Author. Lee Kong Chian School of Business, Singapore Management University, email: aurobindo@smu.edu.sg, Tel: +65 6828-0863.

Governance Matter: Morningstar Stewardship Grades and Mutual Fund Performance

Abstract

Mutual fund investors have the arduous task of disentangling luck from ability of mutual fund managers' performance. In this paper we investigate the role of mutual fund corporate governance (measured by Morningstar Stewardship grade) in mutual fund performance. We propose an objective data-driven corporate governance score based on principal components of Morningstar Stewardship Grades. Furthermore, we establish corporate governance scores have Granger Causality on long-term risk-adjusted returns. The findings suggest that corporate governance grades of mutual funds carry information content beyond the usual star rating measures for predicting long-term mutual fund performance and provide an effective tool for selecting funds.

JEL Codes: G24, G34, G11

Keywords: Mutual fund performance, Corporate Governance, Stewardship Grades, Star Rating, mutual fund ratings, predictive, principal component, Granger Causality

.

1 Introduction

"*Like the stewards of a rich man, they are apt to consider attention to small matters as not for their master's honour, and very easily give themselves a dispensation from having it. Negligence and profusion, therefore, must always prevail, more or less, in the management of the affairs of such a company*" (Adam Smith, 1776, Wealth of Nation, pages 700)

Mutual fund investors have the daunting task of choosing which funds to invest in from thousands of available funds. The literature suggests that investors should fathom the future (short-term) performance of the funds based on current available evidence (Grinblatt and Titman, 1992, Hendricks, Patel and Zeckhauser, 1993, Goetzmann and Ibbotson, 1994, Brown and Goetzmann, 1995, Elton, Gruber and Blake, 1996, Carhart, 1997). Fama and Jensen (1983) argue corporate governance of that mutual fund such as board structure is less important than usual corporations since mutual fund investors use fund flow to discipline managers: they punish poor-performing managers by withdrawing flows and reward good managers with new flows into funds.

In this paper, we explore the role of Morningstar Stewardship Grades (a corporate governance score for mutual funds) in mutual funds. Different from Stewardship Grades, Morningstar Star Ratings have been widely used by retail and institutional investors alike as tools for selecting mutual funds. Del Guercio and Tkac (2008) show significantly large inflows in response to rating upgrades or initiation of top rating. Several papers propose the use of Star Ratings as a performance measure. Blake and Morey (2000) focus on gauging the predictive ability of these ratings. It was documented that poor Star Ratings

indeed indicate weak future performance but good ratings were rarely followed by superior returns for a sample of funds rated by Morningstar between 1992 and 1997 (Blake and Morey, 2000). Subsequent work examining funds rated after June 2002[2] found that best-rated funds outperform lower-rated funds over a three-year post-rating period (Morey and Gottesman, 2006).

In the research we examine how Stewardship Ratings serve as performance indicator for mutual fund performance especially during the financial crisis. The eruptions of the 2003 U.S. mutual fund scandals that involved late trading, market timing and other irregularities put corporate governance of mutual funds in the spotlight, and subsequently led to a series of regulatory reforms. One interesting development that ensued was the launch of the Morningstar Fiduciary Grades (renamed the *Stewardship Grades* in 2005) which evaluated funds based not on their past performance, but on their standard of corporate governance. Stewardship Grades, ranging from A (best) to F (worst), are calculated as the aggregate scores of five components – Corporate Culture, Board Quality, Manager Incentives, Fees and Regulatory History (*cf.* Morningstar, 2007).

Interest in the corporate governance of the mutual funds themselves is of a more recent vintage. Wellman and Zhou (2007) examine the role of the Stewardship Grades on mutual fund performance. They document that funds with top Stewardship Grade outperform those with poor grades by 19 to 23 basis points per month over the period Jan 2001 – July 2004, and by 10 to 16 basis points over the period September 2004 – December 2004. They find that among the five stewardship components, only Fees and Board Quality

[2] Morningstar changes its rating methodology in June 2002.

exhibit significant explanatory power. A recent work by Gil-Bazo and Ruiz-Verdu (2009) shows evidence that better governance is associated with fees that are more aligned with fund performance. Navone (2011) find that funds with better Board Quality grade is associated with less aggressive fees re-pricing by fund companies, although there is no evidence that better Board Quality grades translate into lower expense ratio. Zhou and Wang (2011) study the role governance plays in mutual fund voting. Chen and Huang (2011) employ both OLS regression and quantile regressions to examine the contemporaneous relation between fund performance and corporate governance using both overall Stewardship Grades and two stewardship component grades – Manager Incentive and Board Quality. While OLS regression reveal a strong contemporaneous association between overall Stewardship Grade and fund performance, they do not find evidence of any relation between performance and any of the stewardship components. Controversially, with OLS regressions Gottesman and Morey (2012) find no evidence that any of the stewardship components can consistently predict future performance.

We do not directly use Morning Star Stewardship Rating rather we construct an objective data-driven objective corporate governance score using principal component method. Our research thus provides the most powerful tests on predictive power of Stewardship Ratings on mutual fund performance, thus contributing to this debate[3] on the effectiveness of mutual fund corporate governance.

[3] Casavecchia and Tooman (2012) investigate how governance of Stewardship Ratings is associated with managerial herding behavior. Lai, Tiwari and Zhang (2010) find that funds with bad boards, a negative past performance is strongly predictive of future negative performance while following poor performance, funds with better boards are more likely to change their fund strategy compared to funds with bad boards. Kurniawan, How and Verhoeven (2012) explore whether governance matters to fund style drift. They show that style-drift is negatively related to individual stewardship components such as Board Quality, Fees Structure and Regulatory History.

The salient factors in corporate governance that affect firm value cannot be observed in isolation, in this paper we explore the main drivers controlling for other factors. The main objective of our study is to address the dearth of research in possibly predictive determinants of mutual fund ratings by investigating how well Stewardship Grades can predict future Star Ratings, and hence future fund performance. We address potential econometric issues like *endogeneity* associated with predictive regressions of panel data with a *Two-stage Least Squares* framework, and hence, dynamic panel data regressions to capture the feed-back dynamics of the relationship in a more comprehensive way.

In a panel data model, we find consistent predictability of US mutual fund performance using both monthly and yearly Stewardship Grades after controlling for fund specific characteristics. In the monthly data, we find that Stewardship Grades, while being quite persistent, does indeed Granger cause long term performance measures like the Star Rating.

From the yearly panel data, we have several key findings. First, using Principal Component Analysis, we propose an effective yet procedure agnostic score based on the five components of the Stewardship as the first principal component (Baker and Wurgler, 2006)[4]. Second, employing a naïve fixed effects model, we establish a strong predictive relationship between corporate governance and risk adjusted four-factor alpha (besides the Star Rating) after adjusting for possible *endogeneity bias*. Third, with the use of a dynamic

[4] The first principal component that we calculate based on the total variation of the full 79 monthly data is FPC=0.35Board Quality+0.65Corporate Culture+0.37Fee Score+0.09Manager incentive+0.55Regulatory History. The loadings are almost the same up to second place for 7 yearly December data. We acknowledge that there is some level of selection in the data, however, the closeness of the yearly and monthly proportions give us enough credibility. The results are a little different after 2007 (not reported here) with Board Quality weighted down while Managerial Incentive is increased.

panel data model, we demonstrate that even in the presence of lagged performance measures, a strong relationship between Stewardship Grades and performance holds.

Our findings lend credence to the view that the Morningstar Stewardship Grades supplement the Star Rating as a mutual fund evaluation tool and are particularly effective during crisis periods.

The rest of this paper proceeds as follows. Section 2 provides a description and some statistics of the data we employ. Section 3 presents the methodology we use and Section 4 reports our findings. In Subsection 4.1, we explore the out-of-sample predictability of performance with the Star Rating and Stewardship. We discuss the relationship of short term performance and Stewardship in Subsection 4.2. In Subsection 4.3, we investigate the predictive panel regressions of the dynamic models. Section 5 concludes the paper.

2 Data

Morningstar provided monthly Star Ratings, including the 3-year, 5-year and 10-year ratings (whichever available[5]) and Stewardship Grades. We further obtained all Stewardship Grade components (Corporate Culture or CC, Board Quality or BQ, Fees Score or FS, Manager Incentive or MI and Regulatory History or RH), important fund information like average and longest manager tenure and various fund classifications, over the period November 2004 – May 2011. For simplicity and for subsequent reference, we

[5] Funds whose age is 3 – 5 years will receive a 3-year rating; funds with age 5 – 10 years will receive a 5-year rating; those with age 10 years or longer will receive a 10-year rating. The overall Morningstar rating is derived from a weighted sum of these ratings. More details can be found in Morningstar Factsheets on Ratings and Appendix.

shall enumerate the months as follows: November 2004 is month 1, December 2004 month 2 and so on, with the last month, May 2011 being Month 79.

We merge the Morningstar data with the Centre for Research in Securities Prices (CRSP) Survivorship Bias Free Mutual Fund database. The CRSP database includes the Fama-French-Carhart's four factors (Carhart, 1997), monthly total returns, monthly total net assets, quarterly expenses, quarterly portfolio turnover, date of inception and a 'Dead Fund Flag' that indicates whether the fund has ceased to exist. We include only funds whose fund identifiers from Morningstar (identifier = 'Ticker') and CRSP (Fund Identifier = 'Nasdaq') databases match. Only three funds are identified as 'Dead' fund.

Table 1 displays the frequency distributions of Star Ratings (Panel A) and Stewardship Grade (Panel B) for the January samples. We select only funds that receive both Star Rating and Stewardship Grade over the sample period[6]. We observe that only a small percentage of funds receive the best and worst mutual find ratings and Stewardship Grades. We can also observe an asymmetry in the proportions with the best grades proportions outnumbering the worst ones. This phenomenon can be assigned to both the selection issue and the non-imposition of a symmetric bell curve structure on the scores. For Stewardship Grades, the percentage of funds that receive the top grade of 'A' ranges from 6.1% to 10.2% compared to 9.66% to 16.46% for top star rate funds. For the worst grade of 'F', the proportions are 0.8% to 3.69% as compared to 1.93% to 6.7% 1-star rated funds. The two-way frequencies for both ratings in Panel C reveals that Star Rating and

[6] Choosing only funds that have both Stewardship Grade and Star Rating does entail some level of selection bias in the data as only the funds which are widely held, larger and more familiar to the Morningstar analysts get Stewardship Grades and only the ones with longer history gets favorable star rating, entailing some degree of *endogeneity* in standard models (Lutton et. al., 2011, p. 4-5)

Stewardship Grades are associated with each other for each year (Chi-squared test of Contingency Table results not included).

< Insert Table 1 >

Table 2 reports the descriptive statistics of the fund variables that we use in the empirical part of this paper. Based on Morningstar's 'US Broad Asset Class' , we divide the samples into 5 groups, namely balanced funds, bond funds ('municipal bond' or 'taxable bond'), 'international stock funds', 'specialty funds' and 'U.S. stock funds'. We notice that an overwhelming number of funds are of US equity type (624) nearly twice as much as the next biggest number of bond funds (315). As expected, expense ratio is 50% higher for US equity funds than bond funds and the absolute flow is about two and a half times more. Average manager tenure is between 6.5 and 7.5 years. The turnover ratio for bond funds (1.08) is nearly 50% more than the US equity funds (0.73). The monthly logarithm of age and size are comparable through all categories.

< Insert Table 2 >

The methodology for the Stewardship Grades for funds is independent from the Morningstar Star Rating for funds and thus should have no impact on a fund's Star Rating, through better governance might make the fund more attractive (Morningstar Fact Sheet, 2007). Using the six January samples (2005 to 2011), we compute the Pearson product-moment correlation coefficient between contemporaneous Star Ratings and Stewardship Grades. From Table 3 Panel A, in most cases there is at best some weak though statistically

significant positive correlation between Star Ratings and components of the governance measure. The overall Stewardship Grade and Corporate Culture score are significantly correlated with the Star Rating for all the six monthly data.

In Table 3 Panel B, we perform Granger causality tests with lag length of 2 on the raw scores of ratings from the January samples. For the Star Rating, the raw score can be estimated as follows

Raw score for Star Rating (SR) =

$$\begin{cases} SR_3 & \text{if fund has } 3\text{-}5 \text{ years of returns} \\ 0.6\,SR_5 + 0.4\,SR_3 & \text{if fund has } 5\text{-}10 \text{ years of returns} \\ 0.5\,SR_{10} + 0.3\,SR_5 + 0.2\,SR_3 & \text{if fund has} > 10 \text{ years of returns} \end{cases}$$

where SR_t is the t-year Morningstar Star Rating. For the Stewardship Grade, the raw score is simply the arithmetic average of the scores for the five stewardship component – Corporate Culture (CC), Board Quality (BQ), Manager Incentives (MI), Fees Score (FS) and Regulatory History (RH) (*cf.* Morningstar 2007).

We find in monthly data that raw Stewardship Grade and raw Star Rating strongly Granger cause each other which suggest there is a long term feedback relationship between the two variables. This result however has one caveat as Stewardship Grades are quite persistent (possibly non-stationary) and Star Ratings are not (i.e., stationary), hence such results could be biased. When we examine the difference series of both the Stewardship Grades and the Star Rating, we find no evidence of Granger causality. In an ongoing work (Ng and Ghosh, 2012) , we explore this interesting finding further by using a rigorous long panel data models and dealing with asymptotic results on large cross section (large N) and large time series (large T).

<Insert Table 3>

3 Methodology for Ranked Portfolio Tests and Regressions

Since, Jensen (1969) seminal work on the need for good corporate governance to reduce agency problems, scholarly work on corporate governance standards for firms has been focusing on a scorecard based approach like the Gompers Index or G Index (Gompers et al., 2003). Of the 24 provisions of the Investor Responsibility Research Center (IRRC) that the G-Index focused on, only 6 provisions forming the subsequent *Entrenchment* or E-Index turned out to be the main drivers for firm valuation (Bebchuk, Cohen and Ferrell, 2009). Some salient features of the main drivers, and the possibly *endogenous* control variables for corporate governance, deserve a re-evaluation.

First, although opinions are divided whether entrenchment reduces firm value, it has been documented that managers of firms with low value are often entrenched, hence it is challenging to decipher how much of this entrenchment is *causal* to the low value of the firm (*cf.* Bebchuk 2002, for a survey). This *correlation* could be an outcome of the simultaneous evolution of firm value and managerial incentives (Bebchuk et al., 2009). We account for this endogeneity with the *Two-stage Least Squares* framework applying variables like indicators of managerial ability (e.g., tenure) as instruments. These instruments are assumed to affect the variable reflecting firm performance (in the current context, the shareholders' risk adjusted return) only through the Stewardship Grade, i.e., satisfy the required *exclusion restriction* for a valid instrument (Wooldridge, 2010, p. 242, eq. 9.3).

Second, corporate governance for firms is notoriously sticky or persistent. This feature has been effectively used to "fill in" interim yearly data between the irregular

publications of the IRRC volumes where the governance scores are assumed to essentially remain constant (Gompers et al., 2003, Bebchuk et al., 2009). For the current paper, we are in a unique position to assess the transitions of the corporate governance scores, both monthly as well as yearly. Hence, with the longer time dimension in the longitudinal or panel data, we find strong evidence of the *simultaneity* between our governance score and performance measures using Granger Causality tests. This evidence of co-evolution of corporate governance in mutual funds and their corresponding performance through different business cycles gives us a remarkable insight into their inter-dependence.

Third, in the current context mutual funds are part of financial sector where the major component of the firm performance is risk adjusted return rather than the *Tobin's Q*. However, the components of Stewardship like Managerial Incentive might be pivotal in the performance of the mutual fund for its shareholders. We also observe that limits to shareholder control according to the IRRC provisions like staggered board might be subsumed within Board Quality and Managerial Incentive, while Golden Parachute will most likely be linked with Corporate Culture, Managerial Incentive and Fee Score. Finally, Regulatory History probably is also related to Board Quality and managerial incentive. The other factors not in the E-Index (for example, Director indemnification and relevant contracts, Director's limited liability and severance packages) might also play significant roles (Bebchuk et al., 2009, Lutton et al., 2011). This does put us into a difficult position to extract the true components of stewardship protection, we delegate this responsibility and seek the advantage of aggregating to an objective or data-driven Stewardship score.

We propose to use a Principal Component Analysis (PCA) based methodology that will look at the variation of the entire evaluation dataset to determine the adaptive weights

on the components of Stewardship. This reduces the subjective bias that might be affected by recent or more noteworthy events. In a way we can call this proposed method a really *question agnostic* and *data dependent* framework.

To examine the predictive power of fund ratings, we employ a standard methodology in which we study the relation between in-sample ratings of funds with their out-of-sample performance, as measured by some standard performance metrics over some evaluation period. Our main benchmark is the four-factor model of Carhart (1997):

$$R_{it} - R_{ft} = \alpha + \beta_{i1} RMRF_t + \beta_{i2} SMB_t + \beta_{i2} HML_t + \beta_{i4} UMD_t + \varepsilon_{it}$$

which is an extension of the celebrated Fama and French (1993) three-factor model. In this model, $RMRF_t$ is the value of the market return in excess of monthly T-Bill rate (or the *market risk premium)*; SMB_t (small minus big factor) is the difference in returns across small and big portfolios (or the *size premium)*; HML_t (high minus low factor) is the difference in returns between high and low book-to-market equity portfolios (or the *value premium)*; UMD_t (monthly momentum factor) is the difference in average returns on two high ex-ante return portfolios and two low ex-ante return portfolios (defined as the *momentum factor* or momentum premium).

The SMB factor which is designed to capture the size effect is based on a portfolio comprising a long position in a portfolio of small-cap stocks financed by a short position in a portfolio of large-cap stocks. The HML factor which is meant to capture the book-to-market factor is calculated by building a portfolio that takes a long position in a portfolio of high book-to-market (value) stocks and a short position in a portfolio of low book-to-market (growth) stocks. The UMD factor, described in Jegadeesh and Titman (1993), is a

momentum factor estimated from a portfolio long in high-momentum stocks and short in low-momentum stocks.

Following the methodology in Elton, Gruber and Blake (1996), we perform a two-stage procedure to estimate the monthly out-of-sample performance measures of mutual funds. In stage 1, for each one-year evaluation period [t – 11, t], we regress each fund's monthly excess return on the monthly four risk factors over 36 months (that is, month t – 35 through month t) prior to the last month of the evaluation period. In the second stage, we add the average residuals over one-year prior to and including month t (that is month t – 11 to month t) to the estimated intercept term at month t from stage 1 to get the estimated one-year out-of-sample measure.

Our first approach to examining the strength of predictive power of ratings consists in forming portfolios by their mutual fund ratings and examining the portfolio mean out-of-sample performance over a 12-month evaluation (post-rating) period. Specifically, for each month over the period November 2004 (month 1) to May 2010 (month 67), sample funds are ranked by one or both of the Morningstar Stewardship Grades (abbrev. SG) and/or Morningstar Star Ratings (abbrev SR) and the difference in mean four-factor alpha for funds in any two groups is observed. A Newey-West robust t-test is then performed on the time series of differences.

Funds with SR (respectively SG) = 1 or 2 is in the bottom SR (respectively SG) group. Funds with SR (respectively SG) = 5 is in the top SR (respectively SG) group. The remaining finds are placed in the middle rating group. Funds in the top SG_SR group are those in both top SR and top SG groups. Similarly, funds in the bottom SG_SR group are those in both bottom SR and bottom SG groups. The remaining funds are placed in the

middle SG_SR group. Similar criteria apply to the five Stewardship Grade components - Corporate Culture (CC), Board Quality (BQ), Manager Incentives (MI), Fees Score (FS) and Regulatory History (RH).

As we seek to find a linear sum of the five stewardship components that possibly possesses a stronger predictive power than the overall Stewardship Grade, we employ the Principal Component Analysis on the time series of stewardship components to construct a new corporate governance score which we name the First Principal Component (FPC). Funds are sorted by their FPC and divided into three portfolios - approximately 30% in each of the top and bottom groups and the remaining 40% in the middle group. The next step is to compute, for each rating, the difference in mean out-of-sample return of portfolio. This produces a time series (from month 1 to month 67) of difference in returns between groups 2 and 3 and 1 and 3. We use the symbols 3_2 and 3_1 to denote these differences.

Another standard way to assess predictability of ratings is to run a regression of out-of-sample return on rating dummies. Blake and Morey (2000) perform a cross-sectional dummy-variable regression of the form

$$S_{it} = \mathbf{b}^{T}\mathbf{d} + e_{it}$$

where S_{it} (in %) is the out-of-sample performance measure of fund i at time t and **d** is a vector of rating-based binary dummy variables. In their model, $\mathbf{d} = (d_1, d_2, d_3, d_4)$ is a vector of binary response variables with $d_k = 1$ if a fund has a Morningstar Star Rating of k-star. The best rating group (5-star) is used as the control group. Under the hypothesis that rating is predictive, the following condition on the estimated regression coefficients of d_k hold:

$$b_1 < b_2 < b_3 < b_4 < 0.$$

In this study, we consider the following specification

$$S_{it} = \mathbf{b}^{T}\mathbf{d} + \mathbf{c}^{T}\mathbf{x} + e_{it} .$$

Our regression model differs from the preceding in several ways. First, we use the raw scores of ratings instead of dummy variables. We estimate different regression specifications in which different ratings, including the Star Rating, the overall Stewardship Grade, the five stewardship component grades and the First Principal Component grade, are used. Second, we include **x**, a vector of control variables that are found in the literature to be potential determinants of fund performance.

We are mindful that any results on predictability could be driven by factors such as fund size and fund age. Control variables in **x** include prior one year expense ratio and turnover ratio reported in the CRSP mutual fund database, prior one month absolute fund flow defined as $TNA_t - (1 + R_{i,t-1})TNA_{t-1}$ (TNA_t and R_t being the total net assets and total monthly return provided by CRSP), prior one month natural logarithm of net asset, prior one month natural logarithm of fund age (in months), prior one year average manager tenure and time dummy variables.

Third, instead of treating our data as cross-sectional data at different observation time, we perform panel data regressions which is known to be more informative than its cross-sectional counterparts. We have a unique dataset that provides us with monthly values of the overall Stewardship Grade and the five component grades, hence we are able to use standard panel data models. Our panel methods help us identify the effect that Stewardship Grades as well as the individual components have on standard risk-adjusted returns such as the four factor alpha (Carhart, 1997) or a longer-term performance measure like Morningstar Star Rating. Finally, for the sake of ensuring the robustness of our results and addressing the issue of potential endogeneity, we employ static fixed effect regression,

two-stage least square regression and dynamic panel regressions with instrumental variables. For a detailed description of these regression models we refer to Wooldridge (2010).

We repeat the same analysis on our study of the relationship between the two Morningstar ratings by regressing Star Ratings on the Stewardship Grade or its component grades. As we are dealing with the time series of ratings that are not necessarily stationary, especially the Stewardship Grade or its component grade as we have observed from the data, we perform unit root tests on both series using their raw scores. As expected, while there is no evidence that the Star Rating is non-stationary, we cannot reject the hypothesis that the Stewardship Grade is non-stationary using panel unit root tests (see Baltagi, 2008). In fact, when we further test for stationarity of the first order difference of the Stewardship score time series, we find that the Stewardship score is not distinguishable from a *I(1)* or non-stationary process.[7]

In 2007, Morningstar implemented the following methodology changes to the Stewardship Grades:

1. The weighting on Corporate Culture is increased from 2 to 4 (out of 10)
2. The requirement that independent directors make up 75% of the board be mandatory.
3. Regulatory history score is changed from a scale of 0- 2 to -2 to 0.

[7] These panel unit root tests on monthly data however are based on the assumption that the Stewardship Grade is updated as soon as there are any changes in the governance structure. As these grades are followed by Morningstar analysts and a report written at least once every year, we cannot be certain of this hypothesis (Lutton et. al., 2011). We have ongoing research where we pursue the persistence of the Stewardship Grade and its implication on the performance grade relationship (Ng and Ghosh, 2012)

We refer the reader to Lutton et. al. (2011) for more details. In view of the above changes, we repeat every regression by restricting the sample to data that corresponds to the period January 2007 through May 2011.

We acknowledge the fact that monthly Stewardship Grades might not be updated regularly. There is a significant chance that any changes in ratings are probably related to the time at which Morningstar team evaluates the component Stewardship scores from both direct and indirect sources (Lutton et al., 2011). We also observe a strong persistence of Stewardship Grades vis-à-vis the performance measures. We intend to address this issue in this paper and a follow-up work on persistence (Ng and Ghosh, 2012).

4 Fund Performance: Out of Sample Predictability and Stewardship Grades

4.1 Out-of-sample Predictability

Panels A and B of Table 3 elucidate the significant correlation and strong feedback loop (or Granger causality) that exists between Stewardship Grade (and its components) and Star Rating, it remains to be seen how of this translates into out of sample and long term fund performance measured by the weighted risk adjusted returns i.e. the Star-rating. Table 4 illustrates the difference in out-of-sample performance of portfolio of funds in different groups ranked by their ratings. Applying robust t-tests with Newey-West (HAC) standard errors we can infer that the Morningstar Star Rating has a significant and positive relation with risk adjusted returns i.e., the one year four-factor alpha. In contrast, when ranked by the Morningstar Stewardship Grade in the overall sample period of November 2004-May 2010 (Table 4 Panel A, Columns SG and SR_SG) the constructed portfolio

shows a statistically insignificant but negative on the one year four-factor alpha. This unsurprisingly, must have contributed to the conventional wisdom that overall corporate governance of mutual funds does not make a significant contribution to fund performance in a positive way. Having said that, after the change of methodology in calculating the Stewardship grade was introduced in 2007 (Table 4 Panel B), there was an economically significant positive effect between the top rated and the bottom rated funds.

The overall lack of statistical significance is robust to using the First Principal Component (FPC) score in place of Stewardship Grades (Table 4, Panels A and B, column FPC). When ranked by the both Star rating and FPC score (SR*FPC), we find that even the one year four-factor alpha is positive and economically significant comparing portfolio that are top third rated and bottom rated funds (Table 4 Panel A and B, column SR*FPC). In summary, using Carhart (1997) one-year four factor alpha, we find overall there is some evidence (statistical or economic) that a corporate governance score (even one that is objective and data-driven like FPC) based highly rated portfolios would be able to consistently outperform a lower rated one without controlling for other factors.

The results though are significantly stronger when we use a different more long-term measure of fund performance like the raw star rating score. Portfolios ranked in top third of corporate governance score during the ranking period outperform the middle and bottom third both statistically and economically significant manner. For the period of November 2004 to May 2010, the top third portfolios ranked by previous months Stewardship grade outperformed the middle third by between 0.45, while beat the bottom third by 0.75 percent. The corresponding outperformance when ranked by the objective data-driven FPC score is 0.3 between top and second tercile, and 0.58 between top and

bottom tercile portfolios. These results are not just robust but stronger in the subsample from January 2007-May 2010.

Overall, we can infer that when combined with the Morningstar star ratings, the standard Stewardship Grade and the proposed FPC score seem to be doing a good job in predicting the difference in out-of-sample four-factor alpha between the top and lower tercile portfolios ranked by their corresponding ratings.. However, a stronger and a more consistent result is that corporate governance scores by themselves seem to be more discerning with highest tercile rated portfolios on average outperforming lower rated terciles for the long term weighted risk adjusted return like the star rating.

<Insert Table 4>

4.2 Predictive Panel Regressions of Short term Performance and Stewardship

Predictive performance analysis in Subsection 4.1 indicates that Morningstar Star Rating does have a strong impact on the out-of-sample predictability of the Carhart (1997) four-factor alpha when separately grouped by Stewardship grades and the First Principal Component (FPC) scores. Furthermore, as we observed in Subsection 4.1, the FPC score also plays a significant role in determination of the out-of-sample performance in terms of both the four-factor alpha and the Star Rating. We would first perform predictive panel regressions on the various performance measures including the four-factor alpha, a monthly or yearly performance measure, followed by the Star Rating which is a weighted long term risk-adjusted performance measure.

We perform different specifications of the predictive panel data regression models for US Domestic Equity Mutual Funds. Standard fixed effect model for the yearly data (collected in December, 2005-2011) assuming *strict exogeneity* of the regressors would result in inconsistent estimators (Wooldridge, 2010). The insignificance of the SG and FPC scores in the models might be attributed to the violation of the strict exogeneity assumption in these specifications (Table excluded here).

To address the problem of *endogeneity* we use two stage least squares estimators in the predictive panel regression setting in Table 5 Panel A (Specification EN). The instruments used for the static two stage least squares specification includes previous period's values of the average tenure of the manager, the longest tenure of the manager, the turnover ratio, the expense ratio and the absolute flow variable. These instruments are correlated with the endogenous regressors, Stewardship or FPC scores. Furthermore, these instruments only affect the four-factor alpha or the star rating score (i.e., the dependent variable) through the explanatory variables satisfying the *exclusion restriction* (or *exogeneity*) for valid instruments. We have also included the explanatory variables lagged values of size and age as instruments to ensure that the necessary rank condition is satisfied.

From Table 5 model (SR_EN), we observe that the lagged Star Rating does continue to have a positive impact on the four-factor alpha and size has a negative and significant effect, consistent with findings in Berk and Green (2004) and Bebchuk et. al. (2009). Compared to the preceding models, we find that the Stewardship score has a significantly positive relation with four-factor alpha in specification (SG_EN). In addition, we find once again that fund size has a significantly negative effect on performance. Similar results transpire when we replace Stewardship score with the First Principal

Component in model (FPC_EN). Thus addressing the inherent endogeneity in the model does bring out the effectiveness of corporate governance measures in determining the risk-adjusted performance of US Equity Mutual Funds.

With a predictive panel data model, we can exploit the dynamic behavior through possible inter-relationship with the lagged four-factor alpha as a covariate. However, given only 6 years of data after accommodating for the year lost for constructing the four-factor alpha, the results were expected to be weak at best. Even then, we find that in the overall model (SGA_DY) after controlling for the lagged dependent variable or lagged four-factor alpha, lagged value of the fund size continues to exhibit a significantly negative relation with performance. This result holds across all the regression specifications analyzed. Surprisingly, we also find that Corporate Culture too plays an economically significant negative role in predicting performance. One possible explanation for this minor anomaly is that after controlling for past performance (in terms of lagged alpha), Corporate Culture seems to create possible managerial entrenchment and generate a negative effect (Brown, Harlow and Starks, 1996, Ding and Wermers, 2009, Bebchuck and Cohen, 2004, Bebchuk et al., 2009). This finding is also highlighted by a significant negative effect on fund size (Berk and Green, 2004). We also observe that lagged four-factor alpha tends to have a significant negative impact on future alpha after controlling for other covariates. This effect is economically significant in all four dynamic models. Such a negative effect of previous period's risk adjusted performance can signal possible lack of short term persistence and possibly "window dressing" activity of mutual funds with poor past returns.. Finally, we conclude from models (SG_DY) and (FPC_DY) that although economically significant

neither the Stewardship score nor FPC score have a statistically significant positive impact on performance after controlling for lagged four-factor alpha.

Since the methodology for Stewardship Grade was revamped substantially in 2007, we re-estimate all the regression models using data over the period on and after 2007. Incorporating endogeneity (Models *_EN), results in Panel B are qualitatively similar as those in Panel A for the full sample. One of the main differences is that previous period's size although still economically negative, is insignificant statistically. As before in the static models (*_EN) accommodating for endogeneity, both Stewardship and FPC score turns out to be significant in determining the four factor risk adjusted returns.

<Insert Table 5>

4.3 Predictive Panel Regressions of Long term Performance and Stewardship

The main objective of pursuing good governance is to ensure a long-term sustainable performance. This necessitates the search of an appropriate measure of performance. While the four-factor alpha suffices to be a short term risk adjusted measure of mutual fund performance, its single (monthly or yearly) horizon precludes it from being a viable measure for long term performance. There are a few reasons for this premise. First, an accurate evaluation of the four-factor alpha substantially reduces the data series, particularly for a yearly data in which only a few years of Stewardship Grades are availability. Second, extant literature established that good mutual fund performance (or "hot hands") is not very persistent (Hendricks et al., 1993, Goetzman et al., 1994, Brown et al., 1995). Consequently, using a yearly measure generates a "bounce" which might

deviate from longer run objectives. Third, it is not clear how risk adjusted returns of different time horizons may be combined into a consolidated long-term performance measure, making it a challenge to reach a consensus on the use of such a measure. Finally, published ratings data from sources like Morningstar are more readily available to and trusted by individual investors than model-based risk adjusted returns. Taking all of these into account, a weighted measure of risk adjusted returns of different durations like the Star Rating have gained tremendous popularity among both academics and practitioners (Blake and Morey, 2000, Wellman and Zhou, 2007, Del Guercio and Tkac, 2008).

With a long-term investment objective in mind, we prefer to analyze the raw Star Rating measure with respect to a corporate governance score and other control variables in a predictive panel data setting. To address the possible endogeneity issue that can make the estimated coefficients inconsistent, we use two stage last squares on more parsimonious models described in Table 6 Panel A specifications (EN). Instruments used are lagged values of stewardship scores, average and longest manager tenure, log(age), expense ratio, log(size), turnover ratio and fund flows. The included endogenous regressor variables are Stewardship scores components (Model SGA_EN), the Stewardship score itself (Model SG_EN) or the FPC score (Model FPC_EN). In model (SGA_EN) we observe, none of the stewardship components has statistically significant effect on Star Rating which might be related to possible multicollinearity. We also observe that size and turnover ratio have a positive but insignificant impact on Star Rating. In the Model (SG_EN), lagged size has a significant negative impact on Star Rating or weighted long- term performance consistent with the findings of Berk and Green (2004).

In Model SG_EN (and FPC_EN), both the Stewardship score (and FPC score) and turnover ratio have a positive and significant impact on Star Rating. However, lagged size has statistically significant negative coefficient for the full sample. We can reconcile the somewhat counter-intuitive result on turnover by noting that Star Rating is a long-term measure of past performance which might not be affected by recent active portfolio management. Besides, the relationship between portfolio turnover and fund performance has been a controversial issue. For example, both Carhart (1997) and Malkiel (1995) document a negative association between fund performance and turnover. But results from Grinblatt and Titman (1994) and Wermers (2000) report a positive relation between performance and turnover, thus suggesting that active trading can be positive for fund performance (for an international perspective, see Rao, 2010).

In the dynamic panel data model for yearly data (Table 6 Panel A Specification DY), we find past Star Rating play a significant positive role in all models (SGA_DY, SG_DY and FPC_DY). We also observe that size play a significant negative role, while turnover plays a negative economically significant role after adjusting for past star rating. Both the Stewardship score in model (SG_DY) and the FPC score in model (FPC_DY) are significantly and positively associated with Star Rating when we control for past Star Ratings. Considering that we are using only six years of data, this result further corroborates our view of the inherent long- term relationship between Star Rating and corporate governance of the mutual funds. Results in Table 6 Panel B Specifications EN, which are based on data taken on or after 2007, are qualitatively the same as those reported in Panel A on the whole sample.

As annual reports and financial statements are released once a year, we cannot expect the components of Stewardship Grades to change more frequently than that. However, as we have a wide cross section and different funds have different dates of release of financial statements and quarterly updates, we can assume that some variation in the monthly data on Stewardship Grades exists despite its persistence. With the variation of the Star Rating per month, and its dependence on the current Stewardship scores, it is worthwhile to explore the structural dependence of the two measures in the monthly panel. Furthermore, due to the availability of a longer monthly series, our analysis can also focus on co-evolution of the two processes controlling for other factors.

To address non-exogeneity of the explanatory variables we employ the two stage least squares technique for panels with instruments given by twice lagged dependent variable and lagged values of average manager tenure, longest manager tenure, turnover ratio, expense ratio, log of size and fund flows. In Model (SGA_EN) we find that only Fee Scores has negative and significant coefficient, while other components are positive and significant in explaining star rating. Turnover ratio is positive and significant, and size has also has a positive impact. Interestingly, when we replace individual stewardship component grades with the Stewardship score (or FPC score), the coefficient of size turns negative and significant.

One exciting part about the monthly data is that the number of observations on time domain is sufficiently large to do a complete analysis of the time series in the dynamic panel context. For the Model (SGA_DY), in the presence of the lagged dependent variable star rating, all coefficients of the stewardship components were significant, although Board Quality, Corporate Culture and Fee Score turned out to be negative in the full sample. We

also find that although lagged size is positive, lagged age has a strong negative impact on Star Rating consistently for all three dynamic models. In addition, we also find that being in financial crisis year (2008) was significantly negative for the Star Rating. In examining the shorter sample series from January 2007 to May 2011 (Table 6 Panel D) using static, endogenous and dynamic models, we find similar results.

In the data series with the revised methodology for Stewardship scores, our proposed objective data-driven First Principal Component (FPC) score reduces the dimensionality problem and shows a strong positive significance in models with the more subjective Stewardship score (Table 6 Panel D). We also find a consistently positive effect of turnover, size, average manager tenure and absolute flow, and a negative effect of expense ratio.

In our naïve models where we treat the fund specific variables as exogenous, we are assuming that these variables have a direct impact on the dependent variable: Star Rating. However, we can always argue that these variables are affecting the Star Rating through some other variables like the Stewardship. Hence, it might be more meaningful to include variables that are associated with stability (manager tenure), cost of running the fund (expense ratio) and reputational impact (fund flow) as instruments on direct variables like fund size and turnover ratio. Our two stage least squares on the subsample after 2007 shows significant negative impact of size (consistent with Berk and Green, 2004) while maintaining a positive impact on the Stewardship variable (models FPC_EN and SG_EN). However, turnover appears to assert a positive impact on Star Rating when composite Stewardship variable (SG or FPC) rather than individual components is used.

With the monthly panel from 2007, we can apply the Arrelano and Bond (1991) methodology to evaluate the effect of differences in the Stewardship components and composite indices on the Star Rating without facing a dimensionality problem caused by a short time dimension. We see from Table 6, Panel D model (SGA_DY) that controlling for lagged raw Star Rating score, turnover, size and age, the entire set of stewardship components are significant although BQ, CC, FS and MI have negative effects. Lagged age seems to have a negative significant impact on Star Rating controlling for the difference of funds fixed effects.

Considering that Stewardship components are persistent, we proceed to explore whether changes in Stewardship could be more informative. It turns out that with the exception of Corporate Culture and Managerial Incentive, the first order difference of all stewardship component scores have a negative and significant impact, controlling for past rating, size, age and turnover (Model DSGA_DY). Similar analysis with our proposed FPC score reveals that while the lagged FPC as expected has a positive and significant impact on Star Rating, the lagged first difference of FPC has a negative significant effect after controlling for lagged Star Rating, turnover, age and size (Models FPC_DY and DFPC_DY). We further observe that other than turnover ratio which continues having positive effect, variables like size, age and the indicator for the crisis period all have a significant negative impact on Star Rating. These results are more or less corroborated in the results based on models (SG_DY) and (DSG_DY).

We reckon that due to strong persistence of the corporate governance structure, any changes in these ratings are taken to be highly informative to the investing public and the effect gets reflected heavily in long term investment while controlling for past Star Rating.

Hence, as opposed to the stock variable reflected in the stewardship score itself, changes in the score might communicate past problems in management and induce a negative overreaction.

To check for persistence in the Star Rating, we use two lags of the Star Rating (Model FPC_DY2). Our results suggest that the Star Rating has a lasting effect. As expected, size, age and crisis period have significant negative impact on Star Rating while turnover ratio shows a positive relation.

5 Conclusions

According to the 2012 Investment Company Factbook, ownership of mutual funds by U.S. households hit 44% in year 2011 compared with less than 6% two decades ago. As of the end of 2011, the number of mutual funds in the U.S. market exceeded 19000 while the number of mutual funds available worldwide was close to 73000. Given the multitudes of funds that small institutional investors and retail investors have to choose from, they often rely on mutual fund ratings like the Morningstar Star Rating to guide their investment decision. This throws us back to the long standing problem of evaluating future unobserved performance based on past performance, a practice which can be detrimental to the long-term financial well-being of investors.

The objective of this paper is simple, and really two-fold. First, we evaluate the predictability of performance after controlling for other factors. Second, we comprehensively uncover the relationship of performance, both short-term (like a risk adjusted performance measure like the Carhart's four factor alpha) and long-term (like

Morningstar Star Rating), with some non-return-based performance measures related to the specific funds.

This paper, to the best of our knowledge, is the first rigorous attempt to examine two popular and influential strands of research on mutual fund ratings – Morningstar's Star Rating and Stewardship Grade – in a comprehensive and econometrically robust manner. To examine the predictive power of the ratings, we conduct a ranked portfolio test and predictive panel regressions for both monthly and annual data.

Our investigations lead to several key findings. First, all our empirical results unequivocally indicate that a good Star Rating is associated with good one-year post-rating risk-adjusted return. Second, we further show that after adjusting for endogeneity using a two stage least squares approach, we find a strong and unmistakable link between Stewardship score or our proposed FPC score, and separately for both short term (four factor alpha) and long term performance (Star Rating) measures. Third, using a dynamic panel model, we evaluate how a corporate governance score such as the Stewardship Grade, is still strongly and positively significant in the presence of past Star Ratings. This substantiates the link claimed between governance and performance for mutual funds.

Our research shows that Morningstar Star Stewardship score (or better still, the proposed FPC score) provides a powerful tool for fund selection, just like the Star Ratings. This study helps provide investors with some useful insights into the relation of two seemingly unrelated ratings.

Reference

Baltagi, B. H. “Econometric Analysis of Panel Data.” Fort Edition, New York, John Wiley and Sons Ltd (2008).

Baker, M., and Wurgler, J. “Investor Sentiment and the Cross-Section of Stock Returns.” The Journal of Finance, 61 (2006), 1645–1680.

Bebchuk, L. A. The Case against Board Veto in Corporate Takeovers. University of Chicago Law Review 69 (2002), 973–1035.

Bebchuk, L.A., and A. Cohen. “The Costs of Entrenched Boards.” Journal of Financial Economics, 78 (2005), 409–433.

Bebchuk L.; Cohen, A.; and Ferrell, A. “What Matters in Corporate Governance?” The Review of Financial Studies, 22 (2009), 783–827.

Bebchuk, L. A., and Weisbach, M. S. “The State of Corporate Governance Research.” The Review of Financial Studies, 23 (2010), 939–961.

Berk, J. B. and R. Green. “Mutual Fund Flows and Performance in Rational Markets.” Journal of Political Economy, 112 (2004), 1269–1295.

Blake, C., and Morey, M., “Morningstar Ratings and Mutual Fund Performance.” Journal of Financial and Quantitative Analysis, 35(2000), 451–483.

Brown K. C.; Harlow W. V.; and Starks L. T. “Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry.” The Journal of Finance 51 (1996), 85–110.

Brown, S. J., and Goetzmann, W. N. “Performance Persistence.” The Journal of Finance, 50 (1995), 679–698.

Carhart, M. “On persistence in mutual fund performance.” Journal of Finance, 51(1997), 57–82.

Casavecchia, L., and Tooman, A. “Managerial Herding, Investor’s Sensitivity and the Role of Mutual Fund Internal Governance.” Working Paper, University of Technology, Sydney and ANZ Banking Group (2012).

Chen, C. R., and Huang, Y. “Mutual Fund Governance and Performance: A Quantile Regression Analysis of Morningstar’s Stewardship Grade.” Corporate Governance: An International Review, 19 (2011), 311–333.

Del Guercio, D., and Tkac P. A. “Star power: The effect of Morningstar Ratings on Mutual Fund Flows.” Journal of Financial and Quantitative Analysis, 43 (2008), 907– 936.

Ding, B., and Wermers, R. “Mutual Fund Performance and Governance Structure: The Role of Portfolio Manager and Boards of Directors, Working Paper, University of Maryland (2009).

Elton, E. J.; Gruber M. J.; and Blake C. R. “The Persistence of Risk-Adjusted Mutual Fund Performance.” Journal of Business, 69 (1996), 133–157.

Fama, E., and French, K. R. “The Common Risk Factors in the Returns on Stocks and Bonds.” Journal of Financial Economics, 33 (1993), 3–56.

Gil-Bazo, J., and Ruiz-Verdu, P. “The Relation between Price and Performance in Mutual Fund Industry.” The Journal of Finance, 64 (2009), 2153–2183.

Goetzmann, W. N., and Ibbotson, R. G. “Do Winners Repeat? Patterns in Mutual Fund Performance.” Journal of Portfolio Management, 20 (1994), 9–18.

Gompers, P. A.; Ishii, J.; and Metrick, A. “Corporate Governance and Equity Prices.” Quarterly Journal of Economics, 118 (2003), 107–155.

Gottesman, A., and Morey, M. “Mutual Fund Corporate Governance and Performance.” Review of Financial Economics, 21 (2012), 69 – 81.

Grinblatt, M., and Titman, S. “Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings.” Journal of Business, 62 (1989), 394–415.

Grinblatt, M., and Titman, T. “The Persistence of Mutual Fund Performance.” Journal of Finance, 42 (1992), 1977–1984.

Hendricks, D.; Patel, J.; and Zeckhauser, R. “Hot Hands in Mutual Funds: Short-run Persistence of Performance 1974-88.” Journal of Finance, 48 (1993), 93–130.

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

Jensen, M. C. “Risk, the pricing of capital assets, and evaluation of investment Portfolios.” Journal of Business, 42 (1969), 167–247.

Kurniawan, M.; How, J.; and Verhoeven, P. “Monitoring Style Drift: Evidence from Equity Funds.” Working Paper, Queensland University of Technology (2012).

Lai, S., Tiwari, A., and Zhang, J. “Mutual Fund Flows, Performance Persistence, and Board Quality.” Working Paper, Singapore Management University (2009).

Li, D.; Moshirian, F.; Pham, P.K.; and Zein, J. "When Financial Institutions Are Large Shareholders: The Role of Macro Corporate Governance Environments." The Journal of Finance, 59 (2006), 2975–3007

Lutton, L. P.; Rushkewicz, K.; Liu, K.; and Ling, X. "Mutual fund Stewardship Grade Research Paper. Morningstar Inc. (2011). URL: http://corporate.morningstar.com/us/documents/MethodologyDocuments/StewardshipGradeMethodology.pdf

Malkiel, B. G., 1995. "Returns from Investing in Equity Mutual Funds 1971–1991." Journal of Finance, 50 (1995), 549–572.

Morey, R. M., and Gottesman, A. "Morningstar Mutual Fund Ratings Redux." The Journal of Investment Consulting, 8 (2006), 25–37.

Morningstar Inc. "Fact sheet: The Morningstar Stewardship Grade for funds (2007)." URL: http://corporate.morningstar.com/us/documents/stewardshipgradefunds/StewGradeMethodology_06-18-07.pdf

Navone, M. "Investor's Distraction and Strategic Reprising Decisions." Journal of Banking and Finance, 26(2011), 1291–1303.

Rao, D. N. "Portfolio Turnover and its Effect on Performance of Equity-Oriented Mutual Fund Schemes: An Empirical Study in the Indian Context." Working Paper. Kings University (2010).

Roe, M. J. "Political and Legal Restraints on Ownership and Control of Public Corporations." Journal of Financial Economics, 27(1990), 7–42.

Shleifer, A. and Vishny, R.W. "A Survey of Corporate Governance." The Journal of Finance, 52 (1997), 737–783.

Smith, A. The Wealth of Nations, Cannan Edition, New York: Modern Library (1937).

Wellman, J., and Zhou, J. "Corporate Governance and Mutual Fund Performance: A First Look at the Morningstar Stewardship Grades." Working Paper, Binghamton University (2007).

Wermers, R. "Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transaction Costs and Expenses." Journal of Finance, 55 (2000), 1655–1695.

Wooldridge, J.M. "Econometric Analysis of Cross Section and Panel Data." MIT Press, Cambridge, MA (2010)

Zhou, J.; Ng, L.; and Wang, Q. "Are Better Governed Funds Better Monitors?" Journal of Corporate Finance, Forthcoming.

TABLE 1

Frequency Distribution of Morningstar Ratings and Stewardship Grades

Panel A. Frequency Table of Morningstar Ratings for January Sample of Year 2004 – 2010
This panel reports the percentage of funds that receive the various Morningstar Star Ratings (1-star (Worst) to 5-star (Best)) awarded in the month of January for year 2005 - 2011. Numbers in ( ) indicate percentages.

	Star Rating					
Year	1-star	2-star	3-star	4-star	5-star	N
2005	16 (1.93)	113 (13.68)	285 (34.5)	276 (33.41)	136 (16.46)	826
2006	19 (2.04)	155 (16.66)	328 (35.26)	302 (32.47)	126 (13.54)	930
2007	34 (3.24)	177 (16.87)	376 (35.84)	330 (31.45)	132 (12.58)	1049
2008	30 (3.13)	185 (19.35)	340 (35.56)	292 (30.54)	109 (11.4)	956
2009	57 (6.58)	177 (20.46)	326 (37.68)	208 (24.04)	97 (11.21)	865
2010	44 (5.45)	169 (20.94)	294 (36.43)	222 (27.5)	78 (9.66)	807
2011	58 (6.7)	162 (18.72)	326 (37.68)	224 (25.89)	95 (10.98)	865

Panel B. Frequency Table of Stewardship Grades for January sample of Year 2004 – 2010
This panel reports the percentage of funds that receive the various Stewardship Grades (F (Worst) to A (Best)) awarded in the month of January for year 2005 - 2011. Numbers in ( ) indicate percentages.

	Stewardship Grade					
Year	F	D	C	B	A	N
2005	30 (3.63)	90 (10.89)	230 (27.84)	408 (49.39)	68 (8.23)	826
2006	19 (2.04)	80 (8.6)	285 (30.64)	459 (49.35)	87 (9.35)	930
2007	10 (0.95)	86 (8.19)	348 (33.17)	498 (47.47)	107 (10.2)	1049
2008	30 (3.13)	195 (20.39)	441 (46.12)	230 (24.05)	60 (6.27)	956
2009	32 (3.69)	157 (18.15)	427 (49.36)	196 (22.65)	53 (6.12)	865
2010	25 (3.09)	118 (14.62)	391 (48.45)	200 (24.78)	73 (9.04)	807
2011	7 (0.8)	142 (16.41)	407 (47.05)	238 (27.51)	71 (8.2)	865

Panel C. Two-way Frequency Table of Stewardship Grades and Star Ratings for January Sample of Year 2004 – 2010

This panel reports the two-way frequencies for Stewardship Grades and Star Ratings received by funds as of month of January for year 2005, 2007, 2009, and 2011. Numbers in ( ) indicate percentages. (We omit results for 2006, 2008 and 2010 due to conserve space)

			Star Rating				
Year	N	Stewardship Grade	1-star	2-star	3-star	4-star	5-star
2005	826	F	1 (0.12)	3 (0.36)	18 (2.17)	5 (0.6)	3 (0.36)
		D	0	16 (1.93)	43 (5.2)	20 (2.42)	8 (0.96)
		C	8 (0.96)	52 (6.29)	93 (11.25)	60 (7.26)	17 (2.05)
		B	4 (0.48)	37 (4.47)	116 (14.04)	165 (19.97)	86 (10.41)
		A	0	5 (0.6)	15 (1.81)	26 (3.14)	22 (2.66)
2007	1049	F	0	6 (0.57)	1 (0.09)	3 (0.28)	0
		D	5 (0.47)	24 (2.28)	38 (3.62)	17 (1.62)	2 (0.19)
		C	17 (1.62)	91 (8.67)	136 (12.96)	74 (7.05)	30 (2.85)
		B	11 (1.04)	46 (4.38)	167 (15.91)	192 (18.3)	82 (7.81)
		A	1 (0.09)	10 (0.95)	34 (3.24)	44 (4.19)	18 (1.71)
2009	865	F	4 (0.46)	7 (0.8)	12 (1.38)	7 (0.8)	2 (0.23)
		D	16 (1.84)	44 (5.08)	64 (7.39)	25 (2.89)	8 (0.92)
		C	25 (2.89)	93 (10.75)	164 (18.95)	99 (11.44)	46 (5.31)
		B	11 (1.27)	23 (2.65)	67 (7.74)	65 (7.51)	30 (3.46)
		A	1 (0.11)	10 (1.15)	19 (2.19)	12 (1.38)	11 (1.27)
2011	865	F	0	1 (0.11)	2 (0.23)	1 (0.11)	3 (0.34)
		D	19 (2.19)	35 (4.04)	49 (5.66)	25 (2.89)	14 (1.61)
		C	29 (3.35)	87 (10.05)	160 (18.49)	97 (11.21)	34 (3.93)
		B	10 (1.15)	32 (3.69)	87 (10.05)	74 (8.55)	35 (4.04)
		A	0	7 (0.8)	28 (3.23)	27 (3.12)	9 (1.04)

Table 2
Descriptive Statistics of Fund Variables

This table presents descriptive statistics of important fund variables used in this study. The sample consisting of funds classified as 'U.S. Stock' , 'International Stock' , 'Specialty', 'Bond' ('Taxable Bond' or 'Municipal Bond') and 'Balanced' funds under Morningstar's 'US Broad Category' Classification) receive Stewardship Grades (abbrev. SG) (including each of the five stewardship components) and the Star Ratings (Abbrev SR) in December 2005, 2006, 2007, 2008, 2009 and 2010 November 2004 – May 2011. We report the time series averages of the cross-sectional mean, standard deviation and number of funds.

	Balanced Funds			Bond Funds			International Stock Funds			Specialty Funds			U.S. Stock Funds		
Variables	Mean	Standard Deviation	N	Mean	Standard Deviation	N	Mean	Standard Deviation	N	Mean	Standard Deviation	N	Mean	Standard Deviation	N
log(age in mth)	5.3683	0.6290	110	5.3795	0.4482	315	5.0461	0.4293	213	5.2276	0.4730	13	5.2097	0.5725	624
Expense Ratio	0.0083	0.0049	110	0.0074	0.0034	315	0.0129	0.0046	213	0.0123	0.0055	13	0.0111	0.0042	624
Absolute Fund Flows (in mil)	12.2943	340.0523	110	3.0699	277.3657	315	25.6994	543.2236	212	-2.5658	28.7420	13	-7.3041	229.1272	621
Average Manager Tenure	6.5834	4.9476	107	7.5845	5.4973	315	6.2438	3.9193	213	7.9215	4.5291	13	7.5199	5.3072	624
Turnover ratio	0.6645	0.6046	110	1.0801	1.7108	315	0.6601	0.7053	213	0.6154	0.6746	13	0.7303	0.6040	624
log(size in mil)	6.9439	1.7365	110	6.6896	1.3813	315	6.6019	1.7866	213	6.4433	1.0002	13	6.5347	1.6467	624

Table 3

Correlation and Granger Causality Relation between Star Rating and Stewardship Grades

Panel A Correlation of Star Rating with Stewardship Grade and its Components

This table reports the correlation between the Star Rating and the Stewardship Grade (SG) and each of the Stewardship Grade components - Corporate Culture (CC), Board Quality (BQ) , Manager Incentives (MI), Fees Score (FS) and Regulatory History (RH) for the January sample of 2005 , 2007, 2009 and 2011 (we omit results for 2006, 2008 and 2010 to conserve space)

	January 2005 (N=826)	January 2007 (N=1049)	January 2009 (N=865)	January 2011 (N=865)
SG	0.27183	0.28052	0.19043	0.17891
	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)
BQ	0.23148	0.25559	0.04808	0.0048
	(< 0.01)	(< 0.01)	(0.1578)	(0.8879)
FS	0.17368	0.1345	0.05052	-0.01846
	(< 0.01)	(< 0.01)	(0.1377)	(0.5877)
MI	0.04744	0.01121	0.11455	0.20896
	(0.1732)	(0.7170)	(< 0.01)	(< 0.01)
CC	0.2898	0.30522	0.17222	0.21621
	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)
RH	0.19754	0.22287	0.16411	-0.01057
	(< 0.01)	(< 0.01)	(< 0.01)	(0.7564)

Panel B Pairwise Granger Causality Test on raw scores of Star Rating (SR) and Stewardship Grade (SG)

This table reports the F-statistics and p-value (in parentheses) of Pairwise Granger causality test (lag length 2) between Star Ratings and Stewardship Grades or individual stewardship component grade using monthly time series data from the January 2005 – January 2011.

Variable	SR GC Variable	Variable GC SR
SG	10.0908***	9.5923***
	(< 0.01)	(< 0.01)
BQ	0.97223	4.91146***
	(0.3782)	(< 0.01)
CC	18.2667***	6.67852***
	(< 0.01)	(< 0.01)
FS	1.63902	2.19359
	(0.1942)	(0.1115)
MI	13.5747***	0.51167
	(< 0.01)	(0.5995)
RH	2.57021	8.99095***
	(0.0765)	(< 0.01)

Table 4. Ranked Portfolio Tests

This table reports results of the statistical tests for difference in mean monthly one-year out-of-sample performance measure (one-year four factor alpha or star rating raw score) between two rating groups. For each month over the period November 2004 (month 1) to May 2010 in Panel A and January 2007 to May 2010 in Panel B, sample funds are ranked by one or both of the Morningstar Stewardship Grades (abbrev. SG) and/or Morningstar Star Ratings (abbrev SR) and the difference in mean one-year out-of-sample four-factor alpha for funds in two rating groups is observed. A Newey-West adjusted t-test is performed on the time series of differences. Funds in the top SGSR group are those in both top SR and top SG groups. Similarly, funds in the bottom SG_SR group are those in both bottom SR and bottom SG groups. The remaining funds are placed in the middle SG_SR group. For the First Principal Component (FPC) of the Stewardship Grade factors, funds in the top, middle and bottom group (approximately 30% in each of the top and bottom groups and the remaining 40% in the middle group) are ranked 3, 2 and 1 respectively. symbols 3_2 and 3_1 denote the difference in mean performance measures between the top and middle and between top and bottom groups respectively. Numbers in parentheses are the Newey-West adjusted t-test (4 lags) standard errors. The symbols *, ** and *** denote respectively significance at the 10%, 5% and 1% level.

Panel A. Ranking Period: November 2004 – May 2010

	Out-of-sample Performance Measure						
	One-year four factor Alpha					**Star Rating Raw Score**	
Difference	SR	SG	SR_SG	FPC	SR_FPC	SG	FPC
3_2	0.0248*	-0.0233	-0.0686	0.0139	0.0213	0.4509***	0.3071***
	(0.0131)	(0.0157)	(0.0496)	(0.0180)	(0.0294)	(0.0357)	(0.0249)
3_1	0.0946***	-0.0094	-0.0445	-0.0056	0.0566	0.7490***	0.5887***
	(0.0239)	(0.0305)	(0.0473)	(0.0151)	(0.0361)	(0.0366)	(0.404)

Panel A. Ranking Period: January 2007 – May 2010

	Out-of-sample Performance Measure						
	One-year four factor Alpha					**Star Rating Raw Score**	
Difference	SR	SG	SR_SG	FPC	SR_FPC	SG	FPC
3_2	0.0324*	0.0017	0.0157	0.0354	0.0624**	0.4103***	0.2823***
	(0.0162)	(0.0202)	(0.0458)	(0.0253)	(0.0305)	(0.0452)	(0.0335)
3_1	0.1313***	0.0089	0.0216	-0.0040	0.0774*	0.7140***	0.4863***
	(0.0241)	(0.0455)	(0.0435)	(0.0241)	(0.0441)	(0.0513)	(0.0290)

Table 5
Regressions of Risk-adjusted Returns on Mutual Fund Ratings

We report estimates of regressions to examine the extent to which Morningstar's stewardship grades and/or star ratings predict future return using yearly data from December 2005 (respectively December 2007) through December 2010 in Panel A (respectively Panel B). The regression specification is

$$S_{it} = \mathbf{b}^T \mathbf{Rating}_{it} + \mathbf{c}^T \mathbf{x} + e_{it}.$$

S_{it} (in %) is the one-year Carhart's four-factor alpha. $\mathbf{Rating}_{it}$ is a vector of variables which include one or more of the following mutual fund ratings variables: lagged raw score of star rating (SR), lagged raw score of stewardship grade (SG), lagged raw score of the five stewardship components – corporate culture (CC), board quality (BQ) , manager incentives (MI), fees Score (FS) and regulatory history (RH), and the First Principal Component (FPC) score derived from the stewardship component scores via principal component analysis. **x** is a vector of control variables known to be related to mutual fund performance. Control variables in **x** include lagged expense ratio and turnover ratio obtained from the CRSP mutual fund database, lagged fund flow, lagged logarithm of fund total net asset (in millions), lagged logarithm of fund age (in months) and lagged four factor alpha. We estimate three different models, each with various specifications involving a different set of independent variables.

(Model Specification EN)	Two-stage least squares regression model. The instrumental variables used here include prior(one-year) values of the following variables : average manager tenure, longest manager tenure. fund flows, log (age), expense ratio, log(size) and turnover ratio.
(Model Specification DY)	Dynamic panel model. Instrumental variables used in various specifications are as follows. For specification (SGA_DY): prior one-year and two-year values of S_{it}, prior one-year values of each of: average manager tenure, longest manager tenure. Fund flows, log (age), expense ratio, log(size) and turnover ratio; for specifications (SR_DY) , (SG_DY) and (FPC_DY): prior one-year values of each of: average manager tenure, longest manager tenure. Fund flows, log (age), expense ratio, log(size) and turnover ratio

Panel A (December 2004, 2005, 2006, 2007, 2008, 2009 and 2010 US Equity Funds)

	(Model Spec EN)			(Model Spec DYN)			
Explanatory Variables	(SR_EN)	(SG_EN)	(FPC_EN)	(SGA_DY)	(SR_DY)	(SG_DY)	(FPC_DY)
Intercept	-0.1418**	-0.5565**	-0.5233**				
	(0.0983)	(0.2220)	(0.2175)				
lagged BQ				-5.0945			
				(4.0169)			
lagged CC				-2.051			
				(3.756)			
lagged FS				1.5513			
				(2.3830)			
lagged MI				-1.5619			
				(1.416)			
lagged RH				-1.6243			
				(1.6255)			
Lagged SR raw score	0.0617				0.0511*		
	(0.0402)				(0.0880)		
lagged SG raw score		0.1134**				0.4091	
		(0.0489)				(0.3401)	
lagged FPC			0.2547**				0.2160
			(0.1134)				(1.2094)
lagged FF alpha				-0.3509***	-0.7581*	-0.8813**	-0.7421**
				(0.0957)	(0.2418)	(0.3573)	(0.3554)
lagged turnover ratio	-0.0011	-0.0011	-0.0017	0.0176	-0.0004	-0.0036	-0.0020
	(0.0009)	(0.0007)	(0.0012)	(0.0139)	(0.0007)	(0.0023)	(0.0040)
lagged size	-0.0147**	-0.0388**	-0.0367**	-0.5873***	-0.3291***	-0.2007	-0.2982
	(0.0068)	(0.0193)	(0.0188)	(0.1547)	(0.0651)	(0.2049)	(0.2307)
lagged age							
Fixed Time Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Fund Effect	No	No	No	No	No	No	No
Fixed Fund Effect (Difference)	No	No	No	Yes	Yes	Yes	Yes
N	626	531	531	500	500	500	500

Panel B (December 2007, 2008, 2009 and 2010 Samples)

	(Specification EN)		
	(SR_EN)	(SG_EN)	(FPC_EN)
Intercept	-0.1675	-0.6342*	-0.5233*
	(0.1351)	(0.3237)	(0.2175)
lagged BQ			
lagged CC			
lagged FS			
lagged MI			
lagged RH			
Lagged SR raw score	0.0628		
	(0.0589)		
lagged SG raw score		0.1355*	
		(0.0746)	
lagged FPC			0.2547**
			(0.1134)
lagged turnover ratio	-0.0008	-0.0010	-0.0017
	(0.0009)	(0.0011)	(0.0012)
lagged size	-0.0117	-0.0497	-0.0367
	(0.0098)	(0.0304)	(0.0188)
lagged age			
Fixed Time Effect	Yes	Yes	Yes
Fixed Fund Effect	No	No	No
Fixed Fund Effect (Difference)	No	No	No
N	626	525	525

Table 6
Regressions of Morningstar Star Ratings on Overall/Component Stewardship Grades

We report estimates of regressions to examine the relation between Morningstar star rating and prior period stewardship grades or stewardship component grades. In Panel A (respectively B), yearly data from the December samples spanning 2005 (respectively 2007) through 2010 is used. In Panel C (respectively D), monthly data from November 2004 (respectively January 2007) through May 2011 is used. The regression specification is

$$SR_{it} = \mathbf{b}^{T}\,\mathbf{Rating}_{it} + \mathbf{c}^{T}\,\mathbf{x} + e_{it}.$$

SR_{it} is the Morningstar star rating of fund *i* at year *t*.. $\mathbf{Rating}_{it}$ is a vector of variables that are one or a combination of the following mutual fund ratings variables: lagged raw score of star rating (SR), lagged raw score of stewardship grade (SG), lagged raw score of the five stewardship components – corporate culture (CC), board quality (BQ), manager incentives (MI), fees Score (FS) and regulatory history (RH), and the First Principal Component (FPC) score derived from the stewardship component scores via principal component analysis. **x** is a vector of control variables known to be related to mutual fund performance. Control variables in **x** include lagged turnover ratio obtained from the CRSP mutual fund database, lagged logarithm of fund total net asset (in millions), lagged average and longest manager tenure, lagged logarithm of fund age (in months) and lagged absolute fund flow. We estimate three different models, each with various specifications involving a different set of independent variables.

<u>Yearly Regression (Panel A and B)</u>

(Model Specification EN)	Two-stage least squares regression model. For (SGA_EN), the instrumental variables used here include prior one-year average manager tenure, longest manager tenure, fund flows, log (age), expense ratio, log(size), turnover ratio, expense ratio and all five stewardship component grades (BQ, CC, FS, MI and RH). For (SG_EN), instruments used are prior one-year average manager tenure, longest manager tenure, fund flows, log (age), turnover ratio and expense ratio
(Model Specification DY)	Dynamic Panel model. Instrumental variables used in various specifications are as follows. For all specifications, instruments used are prior one-year and two-year raw scores of SR and prior one-year average manager tenure, longest manager tenure, fund flows, log (age), log(size) , turnover ratio and expense ratio

Monthly Regression Panel C and D

(Model Specification EN) Two-stage least squares regression model. For (SGA_EN), the instrumental variables used here include prior one-month values of the following variables: first principal component of stewardship grades, average manager tenure, longest manager tenure, log (age), expense ratio, log(size) and turnover ratio. For (FPC_EN) and (SG_EN), instruments used are prior one-month values of the following variables: average manager tenure, longest manager tenure, log (age), expense ratio, log(size) , turnover ratio and fund flows.

(Model Specification DYN) Dynamic Panel model

Panel C Instrumental variables used in various specifications are as follows. For specification (SGA_DY): prior one-month, prior one- and two-month raw star rating raw scores, prior one-month values of the following variables: board quality score, corporate culture score, fee score, manager incentive score, regulatory history score, average manager tenure, longest manager tenure, expense ratio, log(size) , turnover ratio and fund flows. and For (SG_DY): prior one- and two-month raw star rating raw scores, prior one-month values of the following variables: average manager tenure, longest manager tenure, expense ratio, log(size) , turnover ratio and fund flows. For specification (FPC_DY): prior one- and two-month raw star rating raw scores, prior one-month values of the following variables: average manager tenure, longest manager tenure, expense ratio, log(size) , turnover ratio , fund flows and a dummy that takes a value of 1 if the time period is in or after January 2007.

Panel D Instrumental variables used in specification (FPC_DY) are prior one- and two-month raw star rating raw scores, prior one-month values of the following variables: average manager tenure, longest manager tenure, expense ratio, log(size) , turnover ratio , fund flows and a dummy that takes a value of 1 if the time period is in or after January 2007. Instrumental variables used in all other specifications are prior one- and two-month raw star rating raw scores, prior one-month values of the following variables: average manager tenure, longest manager tenure, expense ratio, log(size) , turnover ratio and fund flows

Table 6
Regressions of Morningstar Star Ratings on Overall/Component Stewardship Grades

Panel A (December 2004, 2005, 2006, 2007, 2008, 2009 and 2010 Samples)

	(Specification EN)			(Specification DY)		
	(SGA_EN)	(SG_EN)	(FPC_EN)	(SGA_DY)	(SG_DY)	(FPC_DY)
Intercept	-2.3316	-1.4026***	-1.2679***			
	(9.4294)	(0.1599)	(0.1639)			
lagged BQ	-6.1152			1.5233		
	(8.5596)			(1.5346)		
lagged CC	-3.1755			-0.5023		
	(7.9687)			(0.9719)		
lagged FS	0.2201			1.6888		
	(2.8163)			(1.3333)		
lagged MI	1.2526			-0.1882		
	(2.0506)			(0.8381)		
lagged RH	8,9256			0.8795		
	(13.6769)			(0.6528)		
lagged SG raw score		0.7440***			0.7020**	
		(0.0489)			(0.3014)	
lagged FPC			1.7034***			1.7134***
			(0.1036)			(0.5547)
lagged SR raw score				0.1980***	0.4947***	0.4447***
				(0.0715)	(0.0544)	(0.0520)
lagged Turnover ratio	0.0109	0.0087**	0.0040	-0.0055	-0.0024	-0.0047*
	(0.0354)	(0.00394)	(0.0043)	(0.00522)	(0.0025)	(0.0027)
lagged size	0.2388	-0.0913***	-0.0803***	-0.1940**	-0.2969***	-0.2561***
	(0.1843)	(0.0332)	(0.0250)	(0.0859)	(0.0654)	(0.0615)
lagged avg manager tenure						
lagged longest manager tenure						
lagged absolute flows						
lagged age						
Fixed Time Effect	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Fund Effect	No	No	No	No	No	No
Fixed Fund Effect (Difference)	No	No	No	Yes	Yes	Yes
N	536	536	536	508	508	508

Panel B (December 2007, 2008, 2009 and 2010 Samples)

	(Specification EN)			(Specification DY)		
	(SGA_EN)	(SG_EN)	(FPC_EN)	(SGA_DY)	(SG_DY)	(FPC_DY)
intercept	1.1733	-1.4776***	-1.344***			
	(3.8057)	(0.2432)	(0.1106)			
lagged BQ	-2.8588			3.5513*		
	(4.6754)			(1.9048)		
lagged CC	0.1352			-2.1204		
	(2.4047)			(1.7410)		
lagged FS	-0.5677			2.3248		
	(0.6533)			(2.5091)		
lagged MI	0.6922			2.2919		
	(0.5085)			(1.1419)		
lagged RH	2.6490			1.6128		
	(5.1095)			(1.7169)		
lagged SG raw score		0.7894***			1.2171***	
		(0.0717)			(0.4386)	
lagged FPC			1.7953***			2.8346***
			(0.1061)			(0.8100)
lagged SR raw score				0.1574	0.4355***	0.3427***
				(0.1062)	(0.0763)	(0.0675)
lagged turnover ratio	0.0076	0.0032	-0.0015	-0.0082	-0.0057	-0.0093**
	(0.0225)	(0.0028)	(0.0042)	(0.0071)	(0.0036)	(0.0042)
lagged size	0.1655	-0.1388***	-0.1120***	-0.2375	-0.3617***	-0.2996***
	(0.1517)	(0.0382)	(0.0288)	(0.1544)	(0.0919)	(0.0816)
lagged avg manager tenure						
lagged longest manager tenure						
Absolute flows						
Age						
Fixed Time Effect	No	No	No	Yes	Yes	Yes
Fixed Fund Effect	No	Yes	Yes	No	No	No
Fixed Fund Effect (Difference)	No	No	No	Yes	Yes	Yes
N	527	527	527	501	501	501

Panel C (Monthly Samples From November 2004 to May 2011)

Dependent variable: star rating Raw Scores (SRRaw) Sample : Dec 2004 to May 2011

	(SGA_EN)	(FPC_EN)	(SG_EN)	(SGA_DY)	(FPCDcrisis_DY)	(FPC_DY)	(SG_DY)
intercept	1.2182***	-1.8985***	-1.6111***				
	(0.0174)	(0.0900)	(0.0480)				
lagged BQ	0.3079***			-0.0129***			
	(0.0157)			(0.0001)			
lagged CC	0.1799***			-0.0287***			
	(0.0160)			(0.0008)			
lagged FS	-0.1336***			-0.0593***			
	(0.0080)			(0.0035)			
lagged MI	0.1687***			0.0189***			
	(0.0120)			(0.0003)			
lagged RH	0.1815***			0.0911***			
	(0.0115)			(0.0008)			
lagged SG raw score		1.8644***	0.7531***				0.0013***
		(0.0444)	(0.0137)				(1.32E-06)
lagged FPC					0.0023***	-0.0055	
					(1.61E-06)	(0.0384)	
lagged SR raw				0.6930***	0.6933***	0.6936***	0.6883***
				(0.0001)	(1.61E-05)	(0.0006)	(8.69E-06)
Lagged turn ratio	0.0041***	0.0042***	0.0088***	0.0005***	0.0007***	0.0011	0.0005***
	(0.0006)	(0.0009)	(0.0009)	(4.41E-05)	(7.68E-07)	(0.0023)	(1.45E-06)
lagged size	0.1405***	-0.0668***	-0.0729***	0.0552**	0.0348***	0.0377#*	0.0587***
	(0.0038)	(0.0076)	(0.0077)	(0.0003)	(1.25E-05)	(0.0178)	(1.52E-06)
Lagged age				-0.5264***	-0.4127***	-0.2080	-0.4293***
				(0.0017)	(8.80E-05)	(0.5198)	(1.18E-05)
D_crisis					-0.0597***	-0.0413	
					(0.0011)	(0.0204)	
FPC*D_crisis					0.0042***		
					(3.12E-05)		
Fixed Time Effect	Yes	Yes	Yes		No	Yes	Yes
Fixed Fund Effect	No	No	No		No	No	No
Fixed Fund Effect (Difference)	No	No	No		Yes	No	No
N	625	625	625		588	588	588

Panel D (Monthly Samples From January 2007 to May 2011)

	Specification EN		
	(SGA_EN)	(FPC_EN)	(SG_EN)
intercept	1.2202***	-2.0980****	-1.5869***
	(0.0330)	(0.0952)	(0.0697)
lagged BQ	0.2408***		
	(0.0203)		
lagged CC	0.1391***		
	(0.0181)		
lagged FS	-0.1750***		
	(0.0049)		
lagged MI	0.2125***		
	(0.0146)		
lagged RH	0.1793***		
	(0.0188)		
lagged SG raw score			0.7654***
			(0.0179)
lagged FPC		1.9864****	
		(0.0472)	
lagged turnover ratio	0.0038***	0.0010	0.0060***
	(0.0007)	(0.0008)	(0.0008)
lagged size	0.1550***	-0.0935****	-0.1004***
	(0.0029)	(0.0089)	(0.0090)
lagged expense ratio			
lagged avg manager tenure			
lagged absolute flows			
Fixed Time Effect	Yes	Yes	Yes
Fixed Fund Effect	No	No	No
Fixed Fund Effect (Difference)	No	No	No
N	607	607	607

Panel D (contd.)

Dependent Variable : star rating Raw Scores. Sample: Jan 2007 to May 2011

	(Specification DY)						
	(SGA_DY)	(DSGA_DY)	(DFPC_DY)	(DSG_DY)	(FPC_DY)	(SG_DY)	(FPC_DY2)
lagged BQ	-0.0413***						
	(0.0003)						
lagged CC	0.0047***						
	(0.0006)						
lagged FS	-0.0257***						
	(0.0002)						
lagged MI	-0.0136***						
	(0.0003)						
lagged RH	0.1059***						
	(0.0009)						
lagged SG raw score						0.0095***	
						(3.74E-05)	
lagged FPC					0.0275****		0.0182***
					(0.0026)		(0.0073)
lagged SR raw score	0.7433***	0.7459***	0.7434***	0.7457***	0.7441****	0.7441***	0.6749***
	(7.60E-06)	(0.0001)	(1.10E-05)	(2.03E-05)	(0.0005)	(6.90E-05)	(0.0013)
lagged SR raw score (lag 2)							0.1448***
							(0.0002)
Lagged First Difference of BQ		-0.0268***					
		(0.0003)					
Lagged First Difference of CC		0.0254***					
		(0.0013)					
Lagged First Difference of FS		-0.0477***					
		(0.0002)					
Lagged First Difference of MI		0.0069***					
		(0.0002)					
Lagged First Difference of RH		-0.0879*					
		(0.0008)					
Lagged First Difference of FPC			-0.0655****				
			(4.75E-05)				
Lagged First Difference of SG raw score				-0.0271***			
				(2.46E-06)			
lagged turnover ratio	0.0006***	0.0003***	0.0005	0.0003***	0.0008****	0.0006***	0.0013***
	(1.29E-07)	(6.50E-06)	(3.59E-06)	(2.47E-07)	(0.0001)	(3.61E-07)	(0.0004)
lagged size	0.0384***	0.0402***	-0.0013***	0.0406***	0.0096****	0.0430***	-0.0368**
	(0.0001)	(0.0002)	(3.41E-05)	(5.47E-06)	(0.0019)	(1.04E-05)	(0.0006)
lagged age	-0.5558***	-0.3897***	-0.4220***	-0.3911***	-0.4531****	-0.4412***	-0.3814***
	(0.0015)	(0.0006)	(6.72E-05)	(8.51E-05)	(0.0062)	(2.40E-05)	(0.0171)
D_crisis(-1)			-0.0430*		-0.0505****		-0.0431***
			(0.0001)		(0.0023)		(0.0226)
Fixed Time Effect	No	No	No	No	No	No	No
Fixed Fund Effect	No	No	No	No	No	No	No
Fixed Fund Effect (Difference)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	560	558	558	558	560	560	560

Appendix: Principal Component of Stewardship Components

In our empirical analyses, the FPC (First Principal Component) is calculated from a monthly sample of funds (December 2004 – December 2010. Specially, we calculate the first principal component (FPC) based on each of the following seven monthly data sets: Dec 2004 – Nov 2005, Dec 2004 – Nov 2006, …., Dec 2004 – Nov 2010.

Monthly Samples (U.S. Stock Funds)

SG Component	Dec 04 – Nov 05	Dec 04 – Nov 06	Dec 04 – Nov 07	Dec 04 – Nov 08	Dec 04 – Nov 09	Dec 04 – Nov 10
BQ	0.50	0.50	0.50	0.46	0.43	0.40
CC	0.63	0.64	0.64	0.65	0.66	0.66
FS	0.29	0.28	0.30	0.29	0.31	0.32
MI	0.10	0.09	0.08	0.08	0.09	0.11
RH	0.51	0.51	0.50	0.53	0.53	0.53