2013

Three Essays on Mutual Fund Ratings

Wee Seng NG
Singapore Management University, weeseng.ng.2007@smu.edu.sg

THREE ESSAYS ON
MUTUAL FUND RATINGS

NG WEE SENG

SINGAPORE MANAGEMENT UNIVERSITY

2013

Three Essays on Mutual Fund Ratings

by

Ng Wee Seng

Submitted to Lee Kong Chian School of Business in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Business (Finance)

Dissertation Committee:

Dr. Jeremy Goh (Supervisor/Chair)
Associate Professor of Finance
Singapore Management University

Dr. Aurobindo Ghosh (Co-supervisor)
Senior Lecturer of Finance
Singapore Management University

Dr. Jerry Cao
Assistant Professor of Finance
Singapore Management University

Dr. Quoc-Anh Do
Associate Professor of Economics
Sciences Po

Singapore Management University
2013

Abstract

Three Essays on Mutual Fund Ratings

Ng Wee Seng

The incessant growth of the mutual fund industry has made the task of selecting mutual funds an increasingly challenging one. Unsophisticated investors turn to low-cost and readily available ratings to guide their investment decisions. Unsurprisingly, mutual fund ratings are hugely popular and influential. Anecdotal evidence and academic findings both suggest that investors gravitate towards top-rated funds.

Rating is a double-edged sword. Although the use of rating simplifies the otherwise onerous job of evaluating mutual fund performance, it can lead to adverse consequences. Investors who invest only in top-rated funds are inadvertently assuming that good ratings indicate good future performance. However, some academic studies have called into question the predictive ability of mutual fund ratings. Furthermore, the efficacy of a rating depends on its persistence over time. The twin questions of whether ratings possess predictive power and whether ratings are persistent are the main catalysts for this study.

This dissertation comprises three essays on mutual fund ratings. The first examines the relation between the Morningstar stewardship grade and various fund characteristics, such as size, expense ratio and fund manager's tenure, that are known

to be determinants of fund performance. With a data set spanning a period that covers both the recent financial crisis and the year in which a major revamp of the stewardship grade methodology was implemented, I further investigate whether financial crisis or methodology change could have any impact on the results.

In the second essay, I model the Morningstar star rating as a continuous-time Markov process and use the estimated transition probabilities to study the rating dynamics for different types of mutual funds and for funds having different corporate governance ratings given by the Morningstar stewardship grades. Overall, persistence is weak. However, among funds with a good initial star rating, those with a good stewardship grade exhibit a higher probability of having their rating maintained or upgraded, and a lower probability of having their rating downgraded. Results of this kind provide credence to the notion that corporate governance matters in performance.

In the third essay (co-authored with Jeremy Goh and Aurobindo Ghosh), we perform both ranked portfolio tests and predictive panel regressions to corroborate the dependence of risk-adjusted return (four-factor alpha or star rating) on corporate governance score (the stewardship grade) while controlling for fund-specific characteristics. We also propose the use of an objective corporate governance score based on principal component analysis in both static and dynamic fixed-effects regression models. Our results reveal that corporate governance scores do predict performance, thereby reaffirming the economic value of corporate governance to mutual fund investors.

TABLE OF CONTENT

ACKNOWLEDGEMENTS

The completion of this dissertation would not have been possible without the unconditional support, professional advice and constructive criticisms from my dissertation committee, Dr. Jeremy Goh, Dr. Aurobindo Ghosh, Dr. Jerry Cao and Dr. Quoc-Anh Do. I thank all of them wholeheartedly. I am especially indebted to my supervisors and co-authors, Dr. Jeremy Goh and Dr. Aurobindo Ghosh, for equipping me with useful research and programming skills that are instrumental in bringing my research project to fruition.

I am extremely fortunate to have full moral support from my parents and siblings who understand my predicament and believe in me. Their love and encouragement provide me with the mental strength I need to overcome difficulties whenever they arise.

It is my pleasure to recognize the administrative support from Miss Shelly Li and Miss Jes Ong from the Lee Kong Chian School of Business and all the staff of Sim Kee Boon Institute for Financial Economics.

Finally, I would like to show my heartfelt appreciation to teachers, friends and colleagues from the Singapore Management University and the National University of Singapore who have made a positive difference in my life.

Chapter 1 The Determinants of Morningstar Stewardship Grades

1.1 Introduction

The rapid growth of the mutual industry has precipitated a proliferation of mutual fund rating agencies. Readily available mutual fund ratings, such as those created by Lipper and Morningstar, provide mutual fund investors and financial advisers with a low-cost and convenient tool for screening mutual funds. First published in 1985, the Morningstar's star rating is probably the most prevalent and influential mutual fund rating.

The revelation of high-profile mutual fund scandals in 2002 – 2003 had not only triggered numerous academic publications examining various aspects of mutual fund corporate governance, but also caught the attention of Morningstar, which launched the *Fiduciary Grades* (renamed the *Stewardship Grade* in 2005) in 2004. Unlike the star rating which is based on a fund's past performance, this new rating is purported to evaluate a fund based on its standard of corporate governance. The stewardship grade (ranging from F(worst) to A(best)) is calculated as the aggregate score of five components – corporate culture, board quality, manager incentives, fees and regulatory history. For a detailed description of the stewardship grade methodology, we refer the readers to Morningstar (2007).

Both anecdotal evidence and academic studies suggest that mutual fund ratings play an influential role in investor's decision-making process. Indeed, the recent work of Del Guercio and Tkac (2008) and Wellman and Zhou (2007), who separately examine the relation between changes in fund ratings and fund flows, produce strong evidence that mutual fund investors buy funds with good ratings

and sell those with poor ratings. Specifically, the former find that funds that receive the top (5-star) Morningstar rating attract, on average, 53% abnormal inflows of money. The latter run a parallel study on the Morningstar stewardship grade and draw a similar conclusion.

While there has been quite a number of academic studies aimed at dissecting the star rating (Blume, 1998, Khorana and Nelling, 1998, Blake and Morey, 2000, Morey and Gottesman, 2006), publications on the stewardship grade are relatively scarce. Given the potential influence that the stewardship grade has on mutual fund investors' investment decisions, a comprehensive study on its efficacy is warranted. This paper aims to explore the determinants of the stewardship grade in order to better understand how this grade is related to various fund characteristics, such as fund age, fund size and fund manager's tenure, that are in turn known to be relevant to fund performance[1]. With a data set spanning a period that covers both the recent financial crisis and the year in which a major methodological revamp of the stewardship grade was implemented, we further investigate whether financial crisis or methodology change could have any impact on the results.

In examining the relation between stewardship grade and contemporaneous fund characteristics, we find that funds with a better stewardship grade are more seasoned, have a larger asset under management, operate at lower expenses and are managed by more experienced fund managers. Using multinomial logit regressions, we find statistically and economically strong evidence that stewardship grade exhibits a negative relation to prior year turnover

[1] We explore two other important issues – persistence of ratings and predictive power of ratings in two separate papers, Goh, Ghosh and Ng (2013) and Ghosh and Ng (2013).

ratio and prior year expense ratio and a positive relation to prior year star rating and prior year fund size.

The rest of this paper is organized as follows. Section 2 gives a description of the data used. Section 3 presents and discusses the empirical results. Section 4 concludes the paper.

1.2 Data

From Morningstar Direct, we obtain monthly ratings data, including the 3-year, 5-year, 10-year ratings (whenever available[2]), overall star ratings and stewardship grades with all stewardship grade components (corporate culture (CC), board quality (BQ), fees score (FS), managerial incentive (MI) and regulatory history (RH). We further collect monthly data on important fund information such as average and longest manager tenure, over the period November 2004 through May 2011. For simplicity and for subsequent references, We 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. Using the Morningstar's 'US Broad Asset Class', we divide the samples into four groups, namely 'balanced funds', 'bond funds' ('municipal bond' or 'taxable bond'), 'international stock funds' and 'U.S. stock funds'.

In order to examine the relation between important fund characteristics and stewardship grade, we merge the Morningstar data with data on quarterly portfolio turnover ratio, quarterly expense ratio and monthly total net asset, from the Centre for Research in Securities Prices (CRSP) Survivorship Bias Free

[2] 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 (2009).

Mutual Fund database. We include only funds whose fund identifiers from Morningstar (identifier = 'Ticker') and CRSP (Fund Identifier = 'Nasdaq') match.

1.3 Empirical Results

Frequency Distributions and Descriptive Statistics

Table 1 Panels B to E display the frequency distributions of star ratings and stewardship grade in the form of a two-way contingency tables for funds in each of the following categories : 'balanced funds', 'bond funds' ('municipal bond' or 'taxable bond'), 'international stock funds' and 'U.S. stock funds' categorized under Morningstar's 'US Broad Asset Class' classification. We also report the figures for the combined sample in Panel A.

We observe that across all fund categories, only a small percentage of funds receive the best or worst star rating and stewardship grade. Approximately 1% to 4% (respectively 6% to 10%) of the sample funds are awarded the worst (respectively best) stewardship grade. The corresponding figures for star rating are 2% to 4% for 1-star rating and 10% to 17% for 5-star rating. The percentage of equity (US stock or international stock) funds that receive the worst stewardship grade is apparently lower than that for bond or balanced funds. It is also noticeable that the proportion of funds with the best ratings outnumber that for the worst rating. In addition, most (about one-third) of the funds receive the middle or second-best rating('B' or 'C' for stewardship grade and '3-star' or '4-star' for star rating).

In 2007, Morningstar implemented the following methodology changes to the stewardship grade:

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 is made mandatory.

3. Regulatory history score is changed from a scale of 0 to 2 to -2 to 0.

In view of the above changes, we divide the sample into two subsets corresponding to two non-overlapping periods: Period 1 covering November 2004 through January 2007 and period 2 covering February 2007 through May 2011. Here, we assume that the methodology changes took effect from the first month of 2007 and will thus be reflected earliest in the February data. It is interesting to find out how, if any, the frequency distributions of the ratings will be affected by these events.

Table 1 Panel F reports the frequency distributions of the star rating, stewardship grade and each of the five stewardship components for every fund category. With the exception of bond funds, all other categories register a sharp (about 50%) decline in the proportion of funds that receive the best board quality grade from period 1 to period 2. By a much smaller margin, the proportion of funds with best regulatory history grade declines significantly as well.

As for corporate culture, a notable decrease in the percentage of top grades is found in the 'bond' and 'international stock' categories. These observations suggest that changes in the rating methodology do have an impact on the distribution of scores for the two stewardship components – board quality and corporate culture - that these changes are targeting.

A further examination of the table reveals that the percentage of funds getting 'C' for stewardship grade surged after the methodology changes. This applies to all fund categories. For example, we see an increase from 26% to 49% in bond funds and an increase from 31% to 46% in U.S. stock funds.

Average Star Rating of Portfolios Formed by Stewardship Grade

We first investigate whether a fund with a good stewardship rating at the point of investing will also be ranked highly in the star rating going forward. Table 2 reports the quarterly average rating of funds which receive an initial rating of 5-star over a 72-month post-rating period. The sample funds are divided into five portfolios categorized by a fund's an initial stewardship grade. A graphical representation of the results is displayed in Figure 1.

For the initial three-year post-rating period, funds with the best stewardship grade have the highest average star rating. However, we do not observe a strict monotonic relation between initial stewardship grade and average post-rating star rating. Over a longer time period, the average rating of each portfolio tends to converge to an average rating of 3 to 3.5. But by and large, funds with good stewardship grade ('B' or 'A') do, on average, have a higher average star rating than those with average or poor ('C', 'D' or 'F') stewardship grade. Although not reported, we perform a paired sample t-test with Newey-West adjusted standard errors using the time series of differences in average ratings between good stewardship funds and poor stewardship funds. The results are statistically significant.

Difference in Fund Characteristics

To examine the relation between mutual fund ratings and various important fund characteristics such as fund size, fund age, turnover ratio and expense ratio, we compile in Table 3 the descriptive statistics of important fund characteristics. Specifically, the fund characteristics used in this study include expense ratio, turnover ratio, monthly absolute fund flow calculated 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), natural logarithm of monthly total net asset, natural logarithm of fund age (in months) and average manager tenure (in years).

In Panel A, we classify the sample funds according to their stewardship grade (SG). In Panel B, we group the sample funds by their star rating (SR). We then compute, for each portfolio, the mean and standard deviation of each of the fund characteristics. The figures reported are the time series averages of the mean and standard deviations for each fund variable.

From both Panels A and B, we observe that, across all four fund categories, funds with a good stewardship grade (of at least a 'B') or a good star rating (at least 4-star) have lower expense ratio, lower turnover ratio (except for bond funds where the average turnover ratio of 'B'–rated fund is higher than that of 'F'–rated funds) and larger total net asset compared with funds with a grade of C or below. For all the fund categories, funds with a 5-star rating have the highest average manager tenure. With the exception of bond funds, funds with the best ('A') stewardship grade are associated with the highest average manager tenure. However, we find no notable difference for fund age.

As for fund flows, the figures in Panel B suggest that 4-star and 5-star funds attract investment money much more than those with a poorer rating. In fact, for this particular sample, funds with a rating below 4-star suffer from fund outflows, as indicated by the negative signs. It is also worth noting that across all (except bond) fund categories, funds with the best 5-star rating record the highest average fund inflows. The same conclusion applies to the relation between stewardship grade and fund flow for bond funds, balanced funds and international stock funds – funds with a 'B' or 'A' grade register higher positive fund inflows

than funds with poorer grades. However, U.S. funds register negative fund flows regardless of their stewardship grade.

The above analysis is qualitative. For a quantitative treatment, we perform the t-test with Newey-West robust standard errors on the difference in fund characteristics between any two rating-groups. Table 4 reports the results. As before, we divide the sample into four groups based on Morningstar's 'US Broad Category' Classification – 'balanced funds', 'bond funds' ('municipal bond' or 'taxable bond'), 'international stock funds' and 'U.S. stock funds'. All the sample funds included in this study have received monthly stewardship grades (including each of the five stewardship components) and the star ratings over the period November 2004 through May 2011.

In Panel A, we divide the sample funds into five portfolios according to their stewardship grade. In Panel B, we group the sample funds by their Morningstar star rating. For each month over the period November 2004 (month 1) through May 2011 (month 79), the sample funds are ranked by one or both of SG and SR. Funds with SR (respectively SG) = 5 are placed in the top (group 3) SR (respectively SG) group. Funds with SR (respectively SG) < = 2 are placed in the bottom (group 1) (respectively SG) group. The remaining finds are placed in the middle (group 2) (respectively SG) group. The difference in mean fund variable between any two groups is calculated.

In order to investigate whether the results will be affected by the methodology changes to the stewardship grade rating system implemented in 2007 or the financial crisis that occurred during the period August 2008 to March 2009[3]), we divide the sample period into two sub-periods in two different ways:

[3] Based on the Chicago Board Options Exchange Market Volatility Index (VIX)

(I) November 2004 to January 2007 (Evaluation Period A) and February 2007 to May 2007 (Evaluation Period B) for methodology changes;

(II) November 2004 to August 2008 (Evaluation Period C) and April 2009 to May 2011 (Evaluation Period D) for financial crisis.

Over each sub-period, we perform t-test with Newey-West adjusted standard errors on the differences in group means between any two non-overlapping sub-periods (A versus B for (I) and C versus D for (II)). The symbols 2_1, 3_2 and 3_1 denote the difference in mean performance measures between the middle and bottom, top and middle and top and bottom groups respectively.

From Panel A where funds are ranked by their stewardship grades, age is found to be significantly and positively associated with stewardship grade for balanced funds but negatively associated with stewardship grade for bond and equity funds during evaluation periods A, B and C. This conclusion, however changes somewhat for evaluation period D (post-crisis) where both age and stewardship grade have a significantly positive association for the two equity fund groups. Size and manger tenure are two other fund characteristics that generally exhibit a positive relation to stewardship grade. With the exception of a few cases (for example, fund size, 3_2 for international stock funds, evaluation period A, and manager tenure, 2_1 for balanced funds, evaluation period B), this differences in group mean values are positive and highly significant in most cases.

Turnover ratio and expense ratio are two factors that are documented in the literature to have a negative impact on future fund performance (Elton et al. 1993). It turns out that these factors are also negatively related to stewardship grade. For almost all cases, funds with a better stewardship grade tend to have a lower expense ratio and a lower turnover ratio. These relations are especially

pronounced for the post-crisis period (evaluation period D) in which the differences are significantly negative at 1% level in all specifications. The relation between stewardship grade and fund flow is not as straightforward though. The signs of the difference in fund flows vary across cases and the differences are insignificant in many cases.

For the sake of comparison, we rank funds by their star rating and repeat the above analysis. While size and manager tenure are, in most cases, positively related to star rating, age does not yield a similar result. With the exception of bond funds, funds in all other categories exhibit an inverse relation between age and stewardship grade. Specifically, funds with a better rating are younger on the average, regardless of the evaluation periods used. The opposite is true for balanced funds where better-rated funds are more seasoned. The corresponding results for equity funds depend on the evaluation periods examined. Age is negatively associated with star rating when the evaluation periods are A and C, but positive for periods B and D. This demonstrates that methodology changes to the rating system and market conditions do affect the overall results on the relation between age and stewardship rating. Unsurprisingly, funds with better ratings generally have lower average expense ratio and turnover ratio, though there are exceptional cases (for example, the difference 3_2 for U.S. stock funds is significantly positive).

Results of Logistic Regressions

Khorana and Nelling (1998) apply multinomial probit regressions to seek determinants of the star rating. They find that funds with good ratings are associated with larger fund size, longer manager tenure and lower expense ratio.

In the same spirit, we examine the relation between stewardship grade and lagged (one-year) fund characteristics, such as age and size, using multinomial logit regressions in which the dependent variable is the Morningstar's stewardship grade.

We run the regression separately for each of the four fund categories – 'balanced funds', 'bond funds' ('municipal bond' or 'taxable bond'), 'international stock funds' and 'U.S. stock funds'. For every fund category, we estimate two regression models, each comprising two specifications in which different sets of regressors are included. In both specifications, the independent variables include prior-year logarithm of age (in months), prior-year logarithm of fund size (in millions), prior-year turnover ratio, prior-year expense ratio, prior-year fund flow, prior-year average manager tenure (in years). In specification (2), prior-year raw score of the star rating is included as an additional explanatory variable.

For the first regression model whose results are reported in Table 5 Panel A, we run yearly regressions using the December samples (2004 to 2010). We include year dummies to control for the year effect. In Table 5 Panel B, we run Fama and MacBeth regressions using monthly data from month 1(November 2004) through month 79 (May 2011). In both yearly and monthly regressions, the results indicate that prior-year star rating is significantly positive for balanced funds and U.S. stock funds, but not for bond funds and international stock funds. We also find that fund age has a significantly positive coefficient for balanced funds in both specifications (1) and (2). For the other three fund types, age has a negative impact on stewardship grade. In addition, fund size is also positive for all fund

types in both specifications (1) and (2) while manager tenure has a positive relation with stewardship grade for all fund types except balanced funds.

Since fee score is one of the five stewardship components that determine the stewardship grade, it is not surprising that the coefficient on expense ratio is negative across the board, and is statistically significant in almost all cases. The same applies to turnover ratio, which again comes as no surprise since a high portfolio turnover is likely to be viewed unfavorably by the Morningstar analysts who determine the stewardship grade. These results provide some evidence that the stewardship grade has effectively incorporated these two negative aspects of fund management.

Finally, using yearly regressions, the relation between fund flows is mostly insignificant. After controlling for past star rating, fund flow is significant only for US stock funds and its coefficient is negative. When Fama and MacBeth regressions are used, the coefficients of fund flow are significantly positive for bond, balanced and international stock funds but negative for US funds under both specifications (1) and (2). These results suggest that the relation between past fund characteristics and stewardship grade varies with the type of funds examined. The same applies to the relation between past star rating and stewardship grade.

4 Conclusions

In this paper, we seek to gain a better understanding of the Morningstar stewardship grade by examining how this corporate governance rating is related to several important fund characteristics. This study can be regarded as an extension of the work by Korana and Nelling (1998) who analyze the star rating. In assessing contemporaneous relation using ranked portfolio tests, we find that

funds with good stewardship grades are those managed by more experienced managers, incur low expenses and possess a large asset base. In an attempt to investigate whether the recent financial crisis and the methodology changes to the stewardship grade have any significant impact on our results, we report that while the results may vary across the different evaluation periods, the relation between stewardship grade and each of the variables: expense ratio, size and turnover ratio is robust to the sub-period used.

In seeking the determinants of the stewardship grade, we employ multinomial logit regressions using both yearly and monthly data to ensure robustness. We find that controlling for past star rating, stewardship grade is positively related to prior-year age for balanced funds but negatively related to prior-year age for balanced and U.S. stock funds. The relation between stewardship grade and prior-year size is statistically significant and positive for all fund categories. In line with the inverse relation between fund performance and manager replacement documented in Khorana (1996), manager tenure exhibits a positive predictive relation with stewardship grade for bond, international stock and U.S. stock funds, but the relation is negative for balanced funds. The heterogeneity of the results across different fund types underpins the importance of taking the mutual fund types into consideration in conducting such analyses.

Consistent with the results from the ranked portfolio tests, we find that turnover ratio and expense ratio are both significantly negative for stewardship grade across all fund types and in both yearly and monthly regressions. These results attest the ability of stewardship grade in capturing negative factors, such as high expenses and fund managers' excessive trading activities, that are known to adversely affect fund performance.

Table 1

Frequency Distributions of Morningstar Star Ratings and Stewardship Grades

This table reports the percentage of funds that receive each of the 25 mutual fund rating pairs formed by the star rating (abbrev. SR) and the stewardship grade (abbrev SG) based on a pooled sample of monthly ratings data over the period November 2004 – May 2011. Sample funds are divided into four groups categorized by Morningstar's 'US Broad Asset' classification - 'balanced funds', 'bond funds' ('municipal bond' or 'taxable bond'), 'international stock funds' and 'U.S. stock funds. Panel A displays the results for the entire sample. Panel B to E give the frequency tables for each of the four subsamples. The numbers in bold indicate the overall percentage. Numbers within ( ) and { } are respectively the row (for specific SR rating) and column(for specific SG rating) percentages. Numbers in Panel F displays the frequency distribution of each rating over the five fund categories for two sub periods: (1) before methodology change (November 2004 – January 2007) and (2) after methodology change 2007 – May 2011.

Panel A (Entire Sample Nov 2004 – May 2011)

		SG					
		1	**2**	**3**	**4**	**5**	
SR	**1**	51 **0.09** (2.76) {4.07}	284 **0.48** (15.37) {4.12}	1050 **1.76** (56.82) {4.47}	413 **0.69** (22.35) {1.82}	50 **0.08** (2.71) {0.95}	1848 **3.1**
	2	290 **0.49** (2.95) {23.13}	1763 **2.96** (17.94) {25.59}	4985 **8.37** (50.73) {21.24}	2276 **3.82** (23.16) {10.01}	513 **0.86** (5.22) {9.79}	9827 **16.49**
	3	502 **0.84** (2.34) {40.03}	2913 **4.89** (13.60) {42.28}	9059 **15.2** (42.28) {38.60}	7472 **12.54** (34.87) {32.86}	1480 **2.48** (6.91) {28.23}	21426 **35.95**
	4	305 **0.51** (1.61) {24.32}	1436 **2.41** (7.59) {20.84}	6240 **10.47** (33.00) {26.59}	8772 **14.72** (46.39) {38.58}	2158 **3.62** (11.41) {41.17}	18911 **31.73**
	5	106 **0.18** (1.40) {8.45}	493 **0.83** (6.50) {7.16}	2135 **3.58** (28.16) {9.10}	3806 **6.39** (50.20) {16.74}	1041 **1.75** (13.73) {19.86}	7581 **12.72**
	Total	1254	6889	23469	22739	5242	59593
	(%)	**2.1**	**11.56**	**39.38**	**38.16**	**8.8**	**100**

SR	Panel B Balanced Funds						Panel C Bond Funds					
	SG						SG					
	1	**2**	**3**	**4**	**5**	**Total**	**1**	**2**	**3**	**4**	**5**	**Total**
	5	66	95	6	24	196	28	38	257	30	5	358
1	**0.09**	**1.17**	**1.68**	**0.11**	**0.43**	**3.47**	**0.19**	**0.26**	**1.77**	**0.21**	**0.03**	**2.46**
	(2.55)	(33.67)	(48.47)	(3.06)	(12.24)		(7.82)	(10.61)	(71.79)	(8.38)	(1.40)	
	{2.86}	{7.89}	{5.37}	{0.26}	{4.25}		{4.79}	{2.03}	{4.24}	{0.58}	{0.58}	
	77	252	537	178	5	1049	125	330	867	374	53	1749
2	**1.36**	**4.46**	**9.51**	**3.15**	**0.09**	**18.58**	**0.86**	**2.27**	**5.97**	**2.57**	**0.36**	**12.04**
	(7.34)	(24.02)	(51.19)	(16.97)	(0.48)		(7.15)	(18.87)	(49.57)	(21.38)	(3.03)	
	{44.00}	{30.11}	{30.37}	{7.74}	{0.88}		{21.37}	{17.61}	{14.30}	{7.26}	{6.18}	
	90	346	508	720	122	1786	219	795	2192	1130	283	4619
3	**1.59**	**6.13**	**9**	**12.75**	**2.16**	**31.64**	**1.51**	**5.47**	**15.09**	**7.78**	**1.95**	**31.79**
	(5.04)	(19.37)	(28.44)	(40.31)	(6.83)		(4.74)	(17.21)	(47.46)	(24.46)	(6.13)	
	{51.43}	{41.34}	{28.73}	{31.30}	{21.59}		{37.44}	{42.42}	{36.15}	{21.95}	{32.98}	
	3	137	462	1031	182	1815	138	481	2136	2303	340	5398
4	**0.05**	**2.43**	**8.18**	**18.26**	**3.22**	**32.15**	**0.95**	**3.31**	**14.7**	**15.85**	**2.34**	**37.16**
	(0.17)	(7.55)	(25.45)	(56.80)	(10.03)		(2.56)	(8.91)	(39.57)	(42.66)	(6.30)	
	{1.71}	{16.37}	{26.13}	{44.83}	{32.21}		{23.59}	{25.67}	{35.23}	{44.74}	{39.63}	
	0	36	166	365	232	799	75	230	611	1311	177	2404
5	**0**	**0.64**	**2.94**	**6.47**	**4.11**	**14.15**	**0.52**	**1.58**	**4.21**	**9.02**	**1.22**	**16.55**
	(0.00)	(4.51)	(20.78)	(45.68)	(29.04)		(3.12)	(9.57)	(25.42)	(54.53)	(7.36)	
	{0.00}	{4.30}	{9.39}	{15.87}	{41.06}		{12.82}	{12.27}	{10.08}	{25.47}	{20.63}	
Total	175	837	1768	2300	565	5645	585	1874	6063	5148	858	14528
	3.1	**14.83**	**31.32**	**40.74**	**10.01**	**100**	**4.03**	**12.9**	**41.73**	**35.44**	**5.91**	**100**

		Panel D International Stock Funds						Panel D U.S. Stock Funds					
		SG						SG					
		1	2	3	4	5	Total	1	2	3	4	5	Total
		8	28	123	72	12	243	10	152	575	305	9	1051
SR	**1**	**0.08**	**0.28**	**1.24**	**0.73**	**0.12**	**2.46**	**0.03**	**0.51**	**1.95**	**1.03**	**0.03**	**3.56**
		(3.29)	(11.52)	(50.62)	(29.63)	(4.94)		(0.95)	(14.46)	(54.71)	(29.02)	(0.86)	
		{5.41}	{2.12}	{3.33}	{1.86}	{1.40}		{2.89}	{5.31}	{4.81}	{2.67}	{0.30}	
		26	399	917	311	158	1811	62	782	2664	1413	297	5218
	2	**0.26**	**4.04**	**9.28**	**3.15**	**1.6**	**18.33**	**0.21**	**2.65**	**9.02**	**4.78**	**1.01**	**17.67**
		(1.44)	(22.03)	(50.64)	(17.17)	(8.72)		(1.19)	(14.99)	(51.05)	(27.08)	(5.69)	
		{17.57}	{30.27}	{24.84}	{8.04}	{18.41}		{17.92}	{27.34}	{22.30}	{12.37}	{10.03}	
		43	560	1334	1450	217	3604	150	1212	5025	4172	858	11417
	3	**0.44**	**5.67**	**13.5**	**14.67**	**2.2**	**36.47**	**0.51**	**4.1**	**17.01**	**14.12**	**2.9**	**38.65**
		(1.19)	(15.54)	(37.01)	(40.23)	(6.02)		(1.31)	(10.62)	(44.01)	(36.54)	(7.52)	
		{29.05}	{42.49}	{36.14}	{37.50}	{25.29}		{43.35}	{42.38}	{42.06}	{36.52}	{28.98}	
		62	243	914	1437	353	3009	102	575	2728	4001	1283	8689
	4	**0.63**	**2.46**	**9.25**	**14.54**	**3.57**	**30.45**	**0.35**	**1.95**	**9.24**	**13.55**	**4.34**	**29.42**
		(2.06)	(8.08)	(30.38)	(47.76)	(11.73)		(1.17)	(6.62)	(31.40)	(46.05)	(14.77)	
		{41.89}	{18.44}	{24.76}	{37.16}	{41.14}		{29.48}	{20.10}	{22.83}	{35.02}	{43.33}	
		9	88	403	597	118	1215	22	139	955	1533	514	3163
	5	**0.09**	**0.89**	**4.08**	**6.04**	**1.19**	**12.3**	**0.07**	**0.47**	**3.23**	**5.19**	**1.74**	**10.71**
		(0.74)	(7.24)	(33.17)	(49.14)	(9.71)		(0.70)	(4.39)	(30.19)	(48.47)	(16.25)	
		{6.08}	{6.68}	{10.92}	{15.44}	{13.75}		{6.36}	{4.86}	{7.99}	{13.42}	{17.36}	
	Total	148	1318	3691	3867	858	9882	346	2860	11947	11424	2961	29538
		1.5	**13.34**	**37.35**	**39.13**	**8.68**	**100**	**1.17**	**9.68**	**40.45**	**38.68**	**10.02**	**100**

Panel F (Frequency Distribution of Individual Rating over Two Sub-period)

	November 2004 to January 2007						February 2007 to May 2011					
Balanced Funds												
rating	1	2	3	4	5	N	1	2	3	4	5	N
bq	0.00	4.59	12.79	70.63	11.99	73	0.00	4.11	30.74	59.50	5.66	71
cc	0.15	8.48	19.66	46.04	25.67	73	1.63	11.01	30.51	31.35	25.52	71
fs	3.63	4.36	22.63	22.24	47.13	73	13.41	1.14	11.31	25.62	48.51	71
mi	16.30	26.23	38.58	10.68	8.21	73	2.52	24.71	36.21	22.89	13.68	71
ri	3.62	5.67	9.91	5.71	75.08	73	0.54	5.53	15.49	14.45	63.99	71
sg	3.01	6.26	19.87	65.23	5.64	73	3.18	19.44	37.28	27.76	12.35	71
sr	0.80	14.86	32.72	36.72	14.90	112	5.36	21.71	31.08	30.05	11.80	109
Bond Funds												
rating	1	2	3	4	5	N	1	2	3	4	5	N
bq	0.03	4.65	14.93	75.29	5.11	175	0.00	4.16	36.92	52.11	6.81	189
cc	1.36	12.72	23.52	35.74	26.67	175	1.07	10.54	33.89	36.41	18.10	189
fs	7.84	3.74	18.54	13.79	56.10	175	10.85	0.94	13.75	25.50	48.97	189
mi	16.40	26.58	36.30	12.45	8.28	175	3.58	25.86	44.80	18.15	7.62	189
rh	6.98	7.09	13.16	3.16	69.60	175	0.56	4.41	18.20	16.53	60.29	189
sg	4.86	10.13	25.85	51.32	7.85	175	3.76	14.14	49.43	27.75	4.92	189
sr	2.65	11.04	30.95	36.85	18.52	316	4.17	15.45	32.08	33.19	15.11	314
International Stock Funds												
rating	1	2	3	4	5	N	1	2	3	4	5	N
bq	0.00	1.97	16.16	58.63	23.23	131	0.00	0.12	31.05	55.71	13.12	122
cc	0.08	3.02	29.53	31.39	35.97	131	1.19	14.32	29.88	32.60	22.01	122
fs	4.87	4.15	19.15	18.82	53.00	131	13.16	0.85	18.79	26.12	41.08	122
mi	14.20	20.72	44.46	7.89	12.73	131	2.40	26.06	37.62	17.59	16.33	122
rh	0.29	7.23	17.92	4.72	69.85	131	0.06	2.76	19.47	16.70	61.00	122
sg	0.00	4.82	29.22	54.46	11.51	131	2.38	18.31	41.96	30.23	7.12	122
sr	2.85	18.34	31.93	32.19	14.69	213	3.36	17.85	38.13	28.64	12.03	213

U.S. Stock Funds

rating	1	2	3	4	5	N	1	2	3	4	5	N
bq	0.24	5.60	16.54	56.75	20.87	416	0.02	0.76	41.73	45.29	12.20	353
cc	0.58	5.94	25.88	39.95	27.65	416	0.56	10.24	30.04	36.50	22.66	353
fs	7.77	7.53	21.21	19.02	44.48	416	11.32	1.51	16.97	25.69	44.51	353
mi	14.00	19.28	34.64	17.60	14.48	416	4.33	19.42	34.75	22.05	19.45	353
rh	2.30	7.12	12.06	5.21	73.31	416	0.21	1.56	14.06	17.01	67.16	353
sg	1.20	6.94	31.45	49.31	11.10	416	1.22	11.51	46.30	31.73	9.24	353
sr	2.43	16.30	37.16	31.50	12.60	625	4.34	18.66	39.03	28.26	9.71	622

Table 2
Average Star Rating Over Time

This table reports the average star ratings, computed at quarterly intervals, of five mutual fund portfolios formed by the initial (month 1, November 2004) stewardship grade (SG) (1 to 5, where 1 = 'A', 2 = 'B' and so on), of the sample funds. Only funds with a full series of monthly star rating months over the subsequent 78 months (December 2004 through May 2011) are included.

Average Quarterly star rating

	SG				
Month	**1**	**2**	**3**	**4**	**5**
4	3.41	3.27	3.15	3.69	3.90
7	3.36	3.27	3.08	3.70	3.85
10	3.36	3.27	3.12	3.68	3.82
13	3.18	3.13	3.06	3.67	3.80
16	3.23	3.15	3.00	3.66	3.79
19	3.23	3.02	3.00	3.65	3.76
22	3.23	3.06	2.99	3.56	3.65
25	3.18	2.98	3.00	3.61	3.73
28	3.32	3.00	2.96	3.59	3.59
31	3.23	3.04	3.08	3.61	3.58
34	3.27	2.98	3.02	3.60	3.61
37	3.32	3.02	2.99	3.57	3.54
40	3.36	2.96	3.02	3.56	3.45
43	3.27	2.96	3.05	3.57	3.44
46	3.36	2.98	3.04	3.57	3.41
49	3.14	2.92	3.04	3.41	3.20
52	3.23	2.94	3.02	3.38	3.15
55	3.18	2.85	3.04	3.41	3.18
58	3.14	2.94	3.05	3.40	3.24
61	3.05	2.96	3.05	3.42	3.21
64	3.14	3.02	3.08	3.39	3.17
67	3.23	2.92	3.05	3.38	3.25
70	3.23	2.85	3.06	3.37	3.23
73	3.36	2.92	3.03	3.36	3.30
76	3.27	3.00	3.00	3.39	3.20
79	3.32	2.92	3.01	3.36	3.23
N	22	48	220	409	71

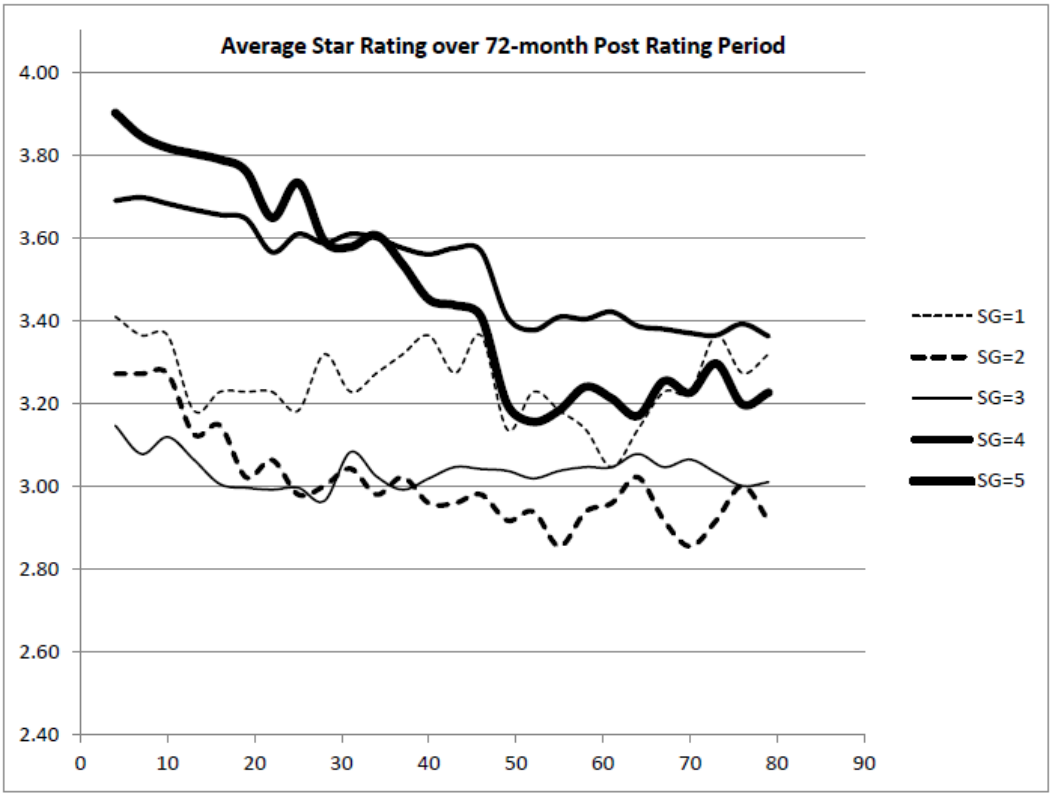


Figure 1. Graphs of Average Star Rating Over Time. This graph plots the portfolio average star rating of the five portfolios formed by a fund's initial (month 1, November 2004) stewardship grade (SG) (1 to 5, where 1 = 'A', 2 = 'B' and so on) over time, for a period of 78 months.

Table 3
Descriptive Statistics of Fund-specific Variables

This table presents descriptive statistics of important fund characteristics considered in this study. The sample funds consist of funds classified as 'balanced funds', 'bond funds' ('municipal bond' or 'taxable bond'), 'international stock funds' or 'U.S. stock funds under Morningstar's 'US Broad Category' Classification) that receive monthly stewardship grades (abbrev. SG) (including each of the five stewardship components) and the star ratings (Abbrev SR) over the period November 2004 – May 2011. Fund characteristics include expense ratio and turnover ratio reported in the CRSP mutual fund database, monthly absolute fund flow calculated 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), natural logarithm of monthly total net asset, natural logarithm of fund age (in months) and average manager tenure. In Panel A, funds are further grouped into five portfolios according to a fund's SG. We compute, for each portfolio, the mean and standard deviation of each of the fund characteristics. The figures reported are the time series averages of the monthly average values. In Panel B, funds are grouped by their SR.

Panel A (Funds Ranked by SG)

		Balanced Funds			Bond Funds			International Stock Funds (less Specialty Funds)			U.S. Stock Funds		
	SG	Mean	Standard Deviation	N	Mean	Standard Deviation	N	Mean	Standard Deviation	N	Mean	Standard Deviation	N
log (age in mth)	1	5.4311	0.5712	2	5.7274	0.4367	7	5.3612	0.1996	3	5.5446	0.4699	4
	2	5.3177	0.6211	11	5.5483	0.3911	24	5.1011	0.3608	17	5.3487	0.5348	36
	3	5.6434	0.6028	22	5.5193	0.3591	77	5.1616	0.4129	47	5.3420	0.5933	151
	4	5.4696	0.5030	29	5.4548	0.3645	65	5.1804	0.4199	49	5.3328	0.5825	145
	5	6.0538	0.7524	7	5.3856	0.3804	11	5.2496	0.7648	11	5.4158	0.5661	37
Expense Ratio	1	0.0109	0.0021	2	0.0088	0.0025	7	0.0163	0.0024	3	0.0148	0.0041	4
	2	0.0106	0.0047	11	0.0089	0.0028	24	0.0149	0.0028	17	0.0130	0.0041	36
	3	0.0093	0.0037	22	0.0079	0.0026	77	0.0131	0.0033	47	0.0113	0.0031	151
	4	0.0056	0.0040	29	0.0052	0.0028	65	0.0100	0.0046	49	0.0090	0.0041	145
	5	0.0070	0.0029	7	0.0053	0.0020	11	0.0107	0.0024	11	0.0095	0.0029	37

Flow	1	-15.8013	16.4118	2	2.6245	23.6006	7	-1.8727	6.1690	3	-7.3567	13.3041	4
	2	6.4112	29.3113	11	-1.2957	33.9580	24	-3.5021	19.7823	17	-5.2035	38.2812	36
	3	1.1845	101.6828	22	-3.0935	89.9930	76	-16.1445	106.0025	46	-10.3115	65.4906	150
	4	4.8793	152.9344	29	16.1339	213.0763	64	3.9284	308.3049	48	-29.1147	207.9870	141
	5	8.1863	712.5463	7	13.1604	583.3033	11	22.4351	1506.4040	11	-14.9674	109.7163	37
Avg Manager Tenure	1	10.3716	2.9757	2	9.6223	6.2016	7	7.7875	3.6773	3	5.9292	4.8849	4
	2	5.9681	3.2333	11	6.6196	4.3032	24	5.4225	3.9679	17	6.0465	4.1294	36
	3	7.0225	6.3172	22	7.6898	5.4911	77	5.4967	3.7554	47	6.2335	4.4078	151
	4	6.0138	4.5002	26	7.7513	5.8135	65	6.6026	4.0385	49	8.0780	5.7141	145
	5	8.3759	5.7010	7	11.2299	5.8113	11	8.4536	3.0685	11	12.1496	5.9260	37
Turnover Ratio	1	0.6642	0.6206	2	0.7342	0.7992	7	0.9236	0.5219	3	0.7866	0.3821	4
	2	0.7961	0.6226	11	1.2640	1.4353	24	0.8657	0.4397	17	0.8241	0.6063	36
	3	0.7682	0.5909	22	1.0853	1.6355	77	0.7214	0.6002	47	0.8107	0.6093	151
	4	0.4652	0.3962	29	0.8941	1.4188	65	0.4934	0.3602	49	0.5869	0.5219	145
	5	0.3229	0.1386	7	0.5545	0.4582	11	0.3928	0.3321	11	0.4220	0.3790	37
log(size in mil.)	1	6.3736	1.1158	2	6.0375	0.7737	7	5.5505	0.2499	3	6.4250	1.3303	4
	2	6.2256	1.4285	11	6.2706	1.0078	24	6.4155	1.0142	17	6.1922	1.2118	36
	3	7.0444	1.5599	22	6.8948	1.1074	77	6.7196	1.5289	47	6.6048	1.3112	151
	4	8.1381	1.4234	29	7.5401	1.2809	65	7.6997	1.5706	49	7.5657	1.4885	144
	5	8.6485	1.7946	7	7.3870	1.3758	11	7.3932	2.3079	11	7.4495	1.5140	37

Panel B (Funds Ranked by SR)

		Balanced Funds			Bond Funds			International Stock Funds			U.S. Stock Funds		
	SR	Mean	Standard Deviation	N	Mean	Standard Deviation	N	Mean	Standard Deviation	N	Mean	Standard Deviation	N
log (age in mth)	1	5.5203	0.6665	5	5.1627	0.6280	11	4.8636	0.4572	7	5.1893	0.6279	23
	2	5.3112	0.5278	21	5.4412	0.4904	44	5.1067	0.4229	38	5.2883	0.5830	111
	3	5.4049	0.6020	35	5.4565	0.4262	100	5.1028	0.4134	77	5.2643	0.5465	239
	4	5.3658	0.6173	36	5.3726	0.3972	108	5.0556	0.4257	64	5.1984	0.5546	183
	5	5.4595	0.7168	14	5.2899	0.4165	51	4.9708	0.3723	28	5.0946	0.5343	67
Expense Ratio	1	0.0116	0.0039	5	0.0106	0.0032	11	0.0146	0.0034	7	0.0128	0.0039	23
	2	0.0114	0.0044	21	0.0096	0.0035	44	0.0141	0.0042	38	0.0123	0.0043	111
	3	0.0082	0.0046	35	0.0082	0.0032	100	0.0131	0.0049	77	0.0107	0.0043	239
	4	0.0068	0.0046	36	0.0063	0.0028	108	0.0122	0.0047	64	0.0105	0.0038	183
	5	0.0072	0.0035	14	0.0058	0.0028	51	0.0118	0.0036	28	0.0111	0.0037	67
Flow	1	-11.4311	18.5122	5	-4.3910	47.2009	11	-11.3377	45.0375	7	-21.5236	57.4639	23
	2	-7.2207	19.4738	21	-7.2507	54.9681	43	-13.9025	68.8446	38	-19.8871	79.7406	110
	3	-14.4030	275.6932	34	-4.2154	73.4253	99	-8.1950	241.3322	76	-18.5372	138.4470	236
	4	7.3398	232.5347	35	6.9956	237.6626	107	3.3115	361.7599	63	-3.3110	132.7013	179
	5	32.3327	196.3583	14	31.7786	227.4810	50	40.9732	411.0880	27	14.6665	181.1819	65
Avg Manager Tenure	1	5.1476	2.4149	5	6.3816	4.4186	11	4.7332	2.7212	7	4.9780	3.5034	23
	2	5.0868	2.8159	21	6.5745	5.2905	44	4.9261	3.6043	38	6.3243	4.6196	111
	3	6.6434	4.1938	34	7.6934	5.4357	100	6.1606	4.0024	77	6.9588	5.0103	239
	4	6.6028	4.7008	33	7.7562	5.2151	108	7.0631	3.7881	64	8.6428	5.6819	183
	5	9.4109	7.9325	14	8.2942	6.0427	51	6.7994	3.8878	28	9.4937	5.6121	67

Turnover Ratio												
	1	0.8407	0.5093	5	0.9767	0.8666	11	0.8003	0.5443	7	0.8405	0.5396	23
	2	0.8753	0.5823	21	1.3324	1.6069	44	0.8638	0.8581	38	0.8448	0.6223	111
	3	0.7054	0.6784	35	0.9305	1.3198	100	0.6489	0.5578	77	0.7525	0.6090	239
	4	0.5376	0.4957	36	0.9238	1.3405	108	0.5559	0.4423	64	0.6611	0.5590	183
	5	0.5040	0.3376	14	1.5359	2.8312	51	0.5422	0.4248	28	0.6223	0.6201	67
log(size in mil.)	1	6.0504	1.2887	5	5.9416	1.1171	11	5.3419	1.7430	7	5.6534	1.7790	23
	2	5.9155	1.3986	21	6.1776	1.3882	44	5.9159	1.5556	38	6.0062	1.5902	111
	3	6.8384	1.6408	35	6.5036	1.3705	100	6.4200	1.7769	77	6.4997	1.6071	239
	4	7.4388	1.6541	36	6.9839	1.2341	108	7.0611	1.6495	64	6.8303	1.5620	182
	5	7.9020	1.6662	14	7.1919	1.3780	51	7.5226	1.6215	27	7.1397	1.5042	66

Table 4
Statistical Tests for Difference in Mean Fund Characteristics

This table reports results of statistical tests for difference in mean fund characteristics between the top and bottom rating groups. The sample consists of funds classified as 'balanced funds', 'bond funds' ('municipal bond' or 'taxable bond'), 'international stock funds' and 'U.S. stock funds' under Morningstar's 'US Broad Category' Classification) that receive monthly stewardship grades (abbrev. SG) (including each of the five stewardship components) and the star ratings (Abbrev SR) over the period November 2004 – May 2011. Fund characteristics include expense ratio and turnover ratio reported in the CRSP mutual fund database, monthly absolute fund flow calculated 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), natural logarithm of monthly total net asset, natural logarithm of fund age (in months) and average manager tenure. In Panel A, funds are further grouped into five portfolios according to a fund's SG. In Panel B, funds are grouped by their SR. For each month over the period November 2004 (month 1) to May 2011 (month 79), sample funds are ranked by one or both of the Morningstar stewardship grades (abbrev. SG) and/or Morningstar star ratings (abbrev SR). Funds with SR (respectively SG) = 5 are placed in the top group(group 3) SR (respectively SG) group. Funds with SR (respectively SG) < = 2 are placed in the bottom group (group 1). The remaining finds are placed in the middle group (group 2).The difference in mean fund variable for funds in any two groups is computed. In order to account for the methodology changes in mutual fund rating system in 2007 as well as the period of high market volatility (August 2008 – March 2009), the sample period is divided into two sub-periods in two ways: (1) Before methodology change (November 2004 - January 2007); after methodology change (February 2007 - May 2007) and (2) Before financial crisis (Nov 2004 - August 2008); after financial crisis (April 2009 – May 2011). For each of these sub-periods, I perform the Newey-West t-test on the differences between any two rating groups. The symbols 2_1, 3_2 and 3_1 denote the difference in mean performance measures between the middle and bottom, top and middle and top and bottom groups respectively. Numbers in parentheses are the Newey-West t-test (4 lags) standard errors. The p-values are in typed bold-face.

Panel A (Fund Ranked by SG)

Evaluation Period A : Before Methodology Change Sample Nov 2004 - Jan 2007

	Balanced Funds			Bond Funds			International Stock Funds			U.S. Stock Funds		
Variable	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**
log (age in mth)	0.1278	0.6422	0.7700	-0.1193	-0.1752	-0.2945	0.1034	-0.2806	-0.1772	-0.1629	-0.1088	-0.2717
	0.18	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Expense Ratio	-0.0043	0.0010	-0.0033	-0.0021	-0.0011	-0.0032	-0.0033	-0.0001	-0.0035	-0.0027	0.0001	-0.0026
	< 0.01	**0.09**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**0.11**	**< 0.01**	**< 0.01**	**0.40**	**< 0.01**
Flow	8.0769	72.7089	80.7857	2.9169	-2.1035	0.8134	29.5295	49.0827	78.6122	-13.4491	23.9261	10.4770
	0.25	**0.01**	**0.02**	**0.60**	**0.66**	**0.93**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**0.04**

Avg Manager Tenure	-1.4651 **< 0.01**	2.9734 **< 0.01**	1.5082 **< 0.01**	0.8702 **< 0.01**	2.1477 **< 0.01**	3.0179 **< 0.01**	0.3660 **< 0.01**	2.7834 **< 0.01**	3.1494 **< 0.01**	0.7664 **< 0.01**	4.9989 **< 0.01**	5.7654 **< 0.01**
Turnover Ratio	0.0194 **0.70**	-0.4004 **< 0.01**	-0.3810 **< 0.01**	-0.0435 **0.22**	-0.0732 **0.27**	-0.1167 **0.09**	-0.0650 **< 0.01**	-0.1153 **0.02**	-0.1802 **< 0.01**	-0.0631 **< 0.01**	-0.0768 **< 0.01**	-0.1399 **< 0.01**
log(size in mil.)	1.3157 **< 0.01**	0.7861 **< 0.01**	2.1018 **< 0.01**	1.0448 **< 0.01**	-0.4629 **< 0.01**	0.5819 **< 0.01**	0.2534 **< 0.01**	0.0487 **0.53**	0.3021 **< 0.01**	0.5875 **< 0.01**	-0.3223 **< 0.01**	0.2652 **0.02**

Evaluation Period B : After Methodology Change Sample February 2008 – May 2011

	Balanced Funds			**Bond Funds**			**International Stock Funds**			**U.S. Stock Funds**		
Variable	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**
log (age in mth)	0.1556 **< 0.01**	0.5099 **< 0.01**	0.6655 **< 0.01**	-0.0955 **< 0.01**	-0.0576 **< 0.01**	-0.1531 **< 0.01**	0.0291 **< 0.01**	0.2756 **< 0.01**	0.3046 **0.00**	0.0386 **0.01**	0.1838 **< 0.01**	0.2223 **< 0.01**
Expense Ratio	-0.0030 **< 0.01**	-0.0009 **< 0.01**	-0.0039 **< 0.01**	-0.0026 **< 0.01**	-0.0013 **< 0.01**	-0.0039 **< 0.01**	-0.0035 **< 0.01**	-0.0012 **< 0.01**	-0.0048 **< 0.01**	-0.0032 **< 0.01**	-0.0010 **< 0.01**	-0.0042 **< 0.01**
Flow	-6.6581 **0.46**	-24.7127 **0.60**	-31.3708 **0.51**	4.7042 **0.59**	14.1221 **0.75**	18.8262 **0.68**	-13.6432 **0.18**	12.7033 **0.94**	-0.9399 **1.00**	-12.7991 **< 0.01**	-6.5641 **0.29**	-19.3632 **0.01**
Avg Manager Tenure	0.2762 **0.13**	1.8275 **< 0.01**	2.1037 **< 0.01**	0.0650 **0.66**	4.4209 **< 0.01**	4.4860 **< 0.01**	0.5346 **0.03**	2.2099 **< 0.01**	2.7445 **< 0.01**	1.1389 **< 0.01**	5.1780 **< 0.01**	6.3169 **< 0.01**
Turnover Ratio	-0.2982 **< 0.01**	-0.2488 **< 0.01**	-0.5470 **< 0.01**	-0.1653 **< 0.01**	-0.6356 **< 0.01**	-0.8009 **< 0.01**	-0.3155 **< 0.01**	-0.2952 **< 0.01**	-0.6108 **< 0.01**	-0.1245 **< 0.01**	-0.4131 **< 0.01**	-0.5376 **< 0.01**
log(size in mil.)	1.3951 **< 0.01**	1.2718 **< 0.01**	2.6668 **< 0.01**	0.8744 **< 0.01**	0.5568 **< 0.01**	1.4312 **< 0.01**	1.1006 **< 0.01**	0.2783 **0.34**	1.3789 **< 0.01**	0.9673 **< 0.01**	0.7718 **< 0.01**	1.7391 **< 0.01**

Evaluation Period C : Sample (Before Crisis) Nov 2004 – August 2008

	Balanced Funds			Bond Funds			International Stock Funds			U.S. Stock Funds		
Variable	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**
log (age in mth)	0.1629	0.6018	0.7647	-0.1135	-0.1428	-0.2563	0.0382	-0.0757	-0.0375	-0.0724	-0.0241	-0.0966
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**0.44**	**0.70**	**0.08**	**0.51**	**0.19**
Expense Ratio	-0.0037	0.0000	-0.0037	-0.0023	-0.0013	-0.0036	-0.0034	-0.0005	-0.0040	-0.0029	-0.0003	-0.0032
	< 0.01	**0.99**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**0.12**	**< 0.01**
Flow	9.5500	55.2231	64.7731	7.6444	6.6154	14.2598	15.9766	74.0101	89.9867	-11.7728	17.5677	5.7949
	0.09	**0.25**	**0.20**	**0.06**	**0.79**	**0.57**	**0.03**	**0.66**	**0.60**	**< 0.01**	**< 0.01**	**0.17**
Avg Manager Tenure	-0.9109	2.2588	1.3479	0.4481	2.8561	3.3043	0.5472	2.7988	3.3460	0.7986	4.9818	5.7804
	< 0.01	**< 0.01**	**< 0.01**	**0.03**	**< 0.01**	**< 0.01**	**0.07**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Turnover Ratio	-0.0195	-0.3337	-0.3532	-0.0728	-0.2065	-0.2793	-0.1640	-0.1161	-0.2802	-0.0575	-0.1645	-0.2220
	0.69	**< 0.01**	**< 0.01**	**0.03**	**0.00**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
log(size in mil.)	1.3716	0.9705	2.3421	0.9789	-0.0957	0.8832	0.6143	-0.2448	0.3695	0.7228	-0.0020	0.7209
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**0.59**	**< 0.01**	**< 0.01**	**0.13**	**< 0.01**	**< 0.01**	**0.99**	**< 0.01**

Evaluation Period D : Sample (After Crisis) April 2009 – May 2011

	Balanced Funds			Bond Funds			International Stock Funds			U.S. Stock Funds		
Variable	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**
log (age in mth)	0.1240	0.4980	0.6221	-0.0898	-0.0417	-0.1315	0.0844	0.2911	0.3756	0.0116	0.2395	0.2511
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**0.54**	**< 0.01**	**< 0.01**
Expense Ratio	-0.0029	-0.0006	-0.0035	-0.0025	-0.0011	-0.0036	-0.0034	-0.0015	-0.0048	-0.0032	-0.0012	-0.0044
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Flow	-3.7633	-72.4752	-76.2385	4.2498	23.0942	27.3440	-12.6954	-27.4089	-40.1044	-12.7071	-20.0429	-32.7500

	0.49	**0.15**	**0.14**	**0.78**	**0.76**	**0.73**	**0.36**	**0.83**	**0.75**	**< 0.01**	**< 0.01**	**< 0.01**
Avg Manager Tenure	0.7214	2.1907	2.9121	0.2176	4.6248	4.8424	0.4960	1.9278	2.4238	1.3504	5.1680	6.5184
	< 0.01	**< 0.01**	**< 0.01**	**0.41**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Turnover Ratio	-0.4130	-0.2676	-0.6807	-0.2116	-0.7677	-0.9793	-0.3488	-0.4026	-0.7514	-0.1906	-0.4831	-0.6737
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
log(size in mil.)	1.3053	1.4081	2.7135	0.8470	0.7337	1.5806	1.0739	1.0793	2.1531	1.0379	1.0164	2.0544
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**

Panel B Fund Ranked by SR

Evaluation Period A : Before Methodology Change Sample Nov 2004 - Jan 2007

	Balanced Funds			**Bond Funds**			**International Stock Funds**			**US Stock Funds**		
Variable	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**
log (age in mth)	0.0633	0.0463	0.1096	0.1418	-0.1509	-0.0091	-0.0002	-0.2381	-0.2383	-0.1094	-0.1589	-0.2683
	0.14	**0.14**	**0.04**	**< 0.01**	**< 0.01**	**0.68**	**0.98**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Expense Ratio	-0.0041	-0.0015	-0.0056	-0.0027	-0.0013	-0.0040	-0.0023	-0.0001	-0.0024	-0.0014	0.0003	-0.0010
	< 0.01	**0.00**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**0.74**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Flow	24.1770	35.5620	59.7390	3.2969	18.2854	21.5823	30.2584	68.4279	98.6863	13.0846	29.1764	42.2610
	< 0.01	**0.02**	**< 0.01**	**0.46**	**0.00**	**0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Avg Manager Tenure	1.1934	2.8170	4.0104	0.9147	0.0217	0.9364	1.4506	0.5539	2.0046	1.5876	1.6147	3.2023
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**0.91**	**< 0.01**	**< 0.01**	**0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**

Turnover Ratio	-0.1192	-0.2595	-0.3788	-0.4604	0.2625	-0.1979	-0.1351	-0.0010	-0.1361	-0.0792	-0.1149	-0.1941
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**0.02**	**< 0.01**	**0.93**	**< 0.01**	**0.00**	**< 0.01**	**< 0.01**
log(size in mil.)	1.1203	0.7275	1.8478	0.4399	0.4013	0.8412	1.2656	0.5051	1.7707	0.5690	0.3553	0.9242
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**

Evaluation Period B : After Methodology Change Sample February 2008 – May 2011

	Balanced Funds			**Bond Funds**			**International Stock Funds**			**U.S. Stock Funds**		
Variable	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**
log (age in mth)	0.0422	0.0876	0.1298	-0.0458	-0.1077	-0.1535	0.0139	-0.0423	-0.0283	-0.0044	-0.1284	-0.1328
	0.03	**< 0.01**	**< 0.01**	**0.02**	**< 0.01**	**< 0.01**	**0.22**	**0.04**	**0.23**	**0.87**	**< 0.01**	**< 0.01**
Expense Ratio	-0.0039	0.0003	-0.0036	-0.0026	-0.0014	-0.0040	-0.0011	-0.0012	-0.0023	-0.0020	0.0005	-0.0015
	< 0.01	**0.32**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Flow	-4.8881	35.5768	30.6887	10.6398	36.0013	46.6411	2.0182	30.9270	32.9452	6.2608	25.0829	31.3436
	0.54	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**0.79**	**0.21**	**0.20**	**0.04**	**< 0.01**	**< 0.01**
Avg Manager Tenure	1.7292	2.7656	4.4948	1.3979	0.8509	2.2488	1.7717	0.0756	1.8473	1.5921	1.9126	3.5047
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**0.83**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Turnover Ratio	-0.3101	-0.0421	-0.3522	-0.2758	0.7906	0.5148	-0.2989	-0.1006	-0.3995	-0.1533	-0.0789	-0.2322
	< 0.01	**0.15**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**0.05**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Log(size in mil.)	1.2696	0.7732	2.0429	0.7215	0.4599	1.1814	0.6622	0.9809	1.6431	0.7650	0.5725	1.3375
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**

Evaluation Period C : Sample Before Crisis Nov 2004 – August 2008

	Balanced Funds			Bond Funds			International Stock Funds			U.S. Stock Funds		
Variable	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**
log (age in mth)	0.0366	0.0498	0.0864	0.0715	-0.1306	-0.0591	0.0035	-0.1892	-0.1857	-0.1093	-0.1565	-0.2658
	0.21443	**0.06758**	**0.01036**	**0.04506**	**< 0.01**	**0.03356**	**0.67027**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Expense Ratio	0.00	-0.0010	-0.0056	-0.0028	-0.0013	-0.0042	-0.0019	-0.0006	-0.0025	-0.0016	0.0003	-0.0013
	< 0.01	**0.01909**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**0.01803**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Flow	20.3939	31.8176	52.2115	8.1463	26.3018	34.4481	24.4963	61.9975	86.4938	12.7270	33.8393	46.5663
	< 0.01	**< 0.01**	**< 0.01**	**0.01984**	**< 0.01**	**< 0.01**	**< 0.01**	**0.03083**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Avg Manager Tenure	1.2731	2.4965	3.7696	1.1763	0.1216	1.2979	1.4234	0.9497	2.3731	1.5439	1.5377	3.0817
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**0.33245**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Turnover Ratio	-0.1873	-0.1772	-0.3645	-0.3870	0.2915	-0.0956	-0.0997	0.0393	-0.0604	-0.0527	-0.0938	-0.1464
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**0.17526**	**< 0.01**	**0.06123**	**0.07363**	**< 0.01**	**< 0.01**	**< 0.01**
Log(size in mil.)	1.1491	0.8846	2.0337	0.5567	0.4505	1.0072	1.1387	0.6060	1.7448	0.6189	0.4742	1.0931
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**

Evaluation Period D : Sample (After Crisis) April 2009 – May 2011

	Balanced Funds			Bond Funds			International Stock Funds			U.S. Stock Funds		
Variable	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**	**2_1**	**3_2**	**3_1**
log (age in mth)	0.0792	0.1164	0.1956	-0.0547	-0.1190	-0.1737	0.0227	0.0139	0.0366	0.0623	-0.1095	-0.0472
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**0.25**	**0.13**	**0.05**	**< 0.01**	**< 0.01**	**0.01**
Expense Ratio	-0.0029	0.0004	-0.0026	-0.0024	-0.0014	-0.0038	-0.0011	-0.0011	-0.0022	-0.0019	0.0006	-0.0013
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**

Flow	-8.5094	35.3938	26.8844	11.3184	47.9027	59.2211	-2.6933	37.0109	34.3177	1.5640	19.9292	21.4932
	0.39	**0.03**	**0.11**	**0.03**	**< 0.01**	**< 0.01**	**0.79**	**< 0.01**	**< 0.01**	**0.74**	**0.06**	**0.05**
Avg Manager Tenure	2.0859	2.9201	5.0059	1.3154	1.3966	2.7120	1.9870	-0.9038	1.0832	1.6317	2.2088	3.8405
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Turnover Ratio	-0.3323	-0.0472	-0.3795	-0.2893	1.1686	0.8793	-0.5049	-0.2339	-0.7388	-0.2513	-0.0948	-0.3461
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**
Log(size in mil.)	1.3349	0.7184	2.0533	0.7416	0.4711	1.2127	0.4810	1.1068	1.5878	0.8296	0.5432	1.3729
	< 0.01	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**	**< 0.01**

Table 5

Determinants of stewardship grades using Logistic Regressions

This table reports results of multinomial logit regressions in which the dependent variable is a ordinal variable given by the Morningstar's stewardship grade (which takes integral values of 1 to 5 – '1' corresponds to a 'F' grade and '5' corresponds to a 'A' grade) and the independent variables are some or all of the following variables: lagged one-year value of the following variables: raw scores of Morningstar star rating, raw scores of stewardship grade, fund age (in months), logarithm of fund size (in millions), turnover ratio, expense ratio, fund flow, average manager tenure. Year dummy variables are also included in regressions using yearly data. In Panel A, we run regressions using the December sample, spanning from 2004 through 2010. The reported numbers not in parentheses are the estimated regression coefficients. The numbers in parentheses are the p-values. In Panel B, we run Fama & MacBeth (1974) regressions month by month, from month 14 through month 79. The numbers not in parentheses are the time series average of the estimated regression coefficients. The numbers in parentheses are the Newey-West adjusted p-values. *, ** and *** denote significance at 10%, 5% and 1% respectively.

Panel A (Yearly Data, December 2004 – December 2010)

	Balanced Funds		Bond Funds		International Stock Funds		U.S. Stock Funds	
Independent Variable	**(1)**	**(2)**	**(1)**	**(2)**	**(1)**	**(2)**	**(1)**	**(2)**
Prior-year SR Raw Score		0.6452***		-0.0791		0.1585*		0.3393***
		(< 0.01)		(0.3173)		(0.0645)		(< 0.01)
Prior-year log(age)	0.4725***	0.5772***	-0.8861***	-0.9047***	-0.2124	-0.1621	-0.4241***	-0.300***
	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(0.1846)	(0.3170)	(< 0.01)	(< 0.01)
Prior-year log(size)	0.3953***	0.2883***	0.2892***	0.2930***	0.0963*	0.0695	0.1987***	0.1360***
	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(0.0945)	(0.2396)	(< 0.01)	(< 0.01)
Prior-year Turnover Ratio	-0.6329***	-0.5625***	-0.2203***	-0.2232***	-0.6943***	-0.6923***	-0.5134***	-0.480***
	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)
Prior-year Expense Ratio	-113.1***	-108.0***	-293.0***	-304.8***	-197.4***	-194.2***	-130.5***	-134.9***
	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)

Prior-year Fund Flow	-0.0010*	-0.0008	0.0003***	0.0004	-0.0003	-0.0004	-0.0002	-0.0001*
	(0.0681)	(0.1448)	(< 0.01)	(0.4746)	(0.5864)	(0.4651)	(0.3113)	(0.0502)
Prior-year Mgr Tenure	-0.0355*	-0.0583***	0.0261**	0.0267**	0.1026***	0.0972***	0.0867***	0.0791***
	(0.0798)	(< 0.01)	(0.0203)	(0.0177)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)
D_2005	0.8247**	0.4749	0.8596***	0.8841***	0.8820***	0.8551***	0.4795***	0.4085***
	(0.0225)	(0.1985)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)
D_2006	0.7068**	0.5560	0.7068***	0.7231***	0.7742***	0.7597***	0.5057***	0.4460***
	(0.0431)	(0.1155)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)
D_2007	-0.4998	-0.6659*	-0.6295***	-0.6294***	-1.0195***	-1.0125***	-0.6059***	-0.654***
	(0.1528)	(0.0601)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)
D_2008	-0.7792	-0.8952***	-0.7980***	-0.7951***	-1.0551***	-1.0464***	-0.5966***	-0.649***
	(0.0217)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)
D_2009	-0.6070***	-0.6291*	-0.6652***	-0.6703***	-0.6226**	-0.6458**	-0.0719	-0.1175
	(0.0740)	(0.0666)	(< 0.01)	(< 0.01)	(0.0187)	(0.0149)	(0.6389)	(0.4449)
Sample Size	371	371	986	986	678	678	2043	2043
Pseudo R^2	0.3637	0.4161	0.3549	0.3555	(0.3462)	0.3495	0.2595	0.2764

Panel B (Monthly Data November 2005 – May 2011)

	Balanced Funds		Bond Funds		International Stock Funds		U.S. Stock Funds	
Independent Variable	**(1)**	**(2)**	**(1)**	**(2)**	**(1)**	**(2)**	**(1)**	**(2)**
Prior-year star rating Raw Score		0.8296***		-0.0483		0.0512		0.3391***
		(< 0.01)		(0.2857)		(0.5200)		(< 0.01)
Prior-year Age	0.5506***	0.7194***	-0.9353***	-0.9508***	0.0802***	0.1765	-0.3733***	-0.2509***
	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(0.4584)	(< 0.01)	(< 0.01)
Prior-year log(size)	0.3614***	0.1880***	0.3581***	0.3540***	0.1841	0.0824	0.2177	0.1516
	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(0.4245)	(< 0.01)	(< 0.01)	(< 0.01)
Prior-year Turnover Ratio	-0.4720***	-0.4887***	-0.2107***	-0.2099***	-0.6615***	-0.6212***	-0.4663***	-0.4299***
	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)
Prior-year Expense Ratio	-155.91***	-139.21***	-309.03***	-318.69***	-208.98***	-211.26***	-141.19***	-144.991***
	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)
Prior-year Fund Flow	0.0010**	0.0009**	0.0009***	0.0008***	0.0010***	0.0010***	-0.0002	-0.0006**
	(0.0374)	(0.0246)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)	(0.1726)	(< 0.01)
Prior-year Avg Manager Tenure	-0.0411***	-0.0740***	0.0161**	0.0169**	0.1083***	0.1044***	0.0885***	0.0823***
	(< 0.01)	(< 0.01)	(0.0453)	(0.0469)	(< 0.01)	(< 0.01)	(< 0.01)	(< 0.01)

Chapter 2 Corporate Governance and Persistence of Mutual Fund Ratings: A Markov Chain Approach

2.1 Introduction

The mutual fund Industry is burgeoning. According to the Investment Company Fact Book 2012, as of the end of year 2012, the total asset under management worldwide exceeded 26 trillions U.S. dollars and the total number of mutual funds available worldwide surpassed 73,000. To mutual fund investors, one implication of these statistics is the increasingly tedious and costly task of screening and selecting funds. For those who seek a simple and affordable solution to their conundrum, mutual fund ratings such as those provided by Morningstar and Lipper are obvious choices.

Debuted in 1985 and revamped in 2002, the Morningstar star ratings are probably the most popular and influential mutual find rating[1]. Star ratings (five possible grades with 1-star the worst and 5-star the best) are based on a fund's relative standing within its own peer category in terms of some measure of its past risk-adjusted returns. Funds in the top 10- percentile and the next 22.5-percentile are graded the best 5-star rating and second-best 4-star rating respectively. Funds in the middle 35-percentile and the next 22.5-percentile are given a 3-star rating and a 2-star rating respectively. Those in the bottom 10-percentile receive the worst 1-star rating.

In response to public outcry over the mutual fund scandals in 2002 – 2003 which precipitated a series of prosecutions and regulatory reforms, Morningstar

[1] See Morningstar Factsheet (Morningstar, 2008a)for the methodology of the star rating

launched in 2004 a new mutual fund governance evaluation system to complement the well-established star rating. This new rating, first named the *Fiduciary Grade* and subsequently renamed the *Stewardship Grade* in year 2005, is derived from five corporate governance scores based on the following five factors - (i) Board Quality (0 to 2 points) (ii) Corporate Culture (0 to 4 points) (iii) Fees (0 to 2 points) (iv) Manager Incentives (0 to 2 points) (v) Regulatory Issues (-2 to 0 points). [2]

Mutual fund ratings are popular for many reasons. Investors, especially the less experienced and hence less informed ones, regard them as an endorsement of a fund's quality. The top rating recognized with the prestigious Morningstar's 5-star icon is often used as an advertising tool to entice investors. Indeed, one can find ample evidence suggesting that investors gravitate towards top-rated funds. Some reports claim *"... almost 90% of the new money that flowed into stock funds last year went to funds with four-star and five-star ratings*" – (Wall Street Journal, April 5, 1996) and "*… Even in 2008's brutal market, when the other star-rated funds saw net outflows ranging from $111 billion for 3-star funds to $14 billion for 4-star funds, 5-star funds enjoyed $67.5 billion in net inflows*' – (Wall Street Journal, 1 June, 2010). On the academic front, two recent studies (Del Guercio and Tkac, 2008, Wellman and Zhou, 2007) document that the funds that receive the top rating for the first time attract very high abnormal fund inflows over a six-month post-rating period. Their studies also demonstrate that both rating upgrades and downgrades have significant impact on fund flows.

Given the influence mutual fund ratings have on investors' financial decision-making, one natural question to explore is whether the "chasing the star" strategy is

[2] See Morningstar Factsheet for Stewardship Grade (2007)

sound? To answer this question, one has to examine at least two crucial issues. The first concerns the predictive nature of ratings. It is important to find out whether ratings possess any ability to predict future performance and if so, to what extent and over what time horizon. The second issue to investigate is the persistence of ratings. More specifically, one should find out to what extent top-rated funds are able to maintain its rating and for how long a period of time. Our separate study (Ghosh, Goh and Ng, 2012) among others has explored the first question using a more structural approach. One objective of this current paper is to focus on the second question and test empirically the degree of persistence of the Morningstar star rating.

Surprisingly, the amount of academic publications that examine mutual fund ratings pale in comparison with those that focus on other aspects of mutual funds such as fund performance and fund managers' skills. In particular, academic study on persistence of mutual fund ratings are difficult to come by. Khorana and Nelling (1998) are among the earliest to examine the persistence of mutual fund ratings. Their findings indicate that the Morningstar star ratings exhibit a high degree of persistence over the period December 1992 – June 1995. However, Warshawsky et al. (2000) find that their results could be plagued by survivorship bias. The latter examine the degree of persistence of the top two ratings (four-star and five-star) by tracking the percentage of funds that can retain their top rating over a one-year observation window over the period 1997 – 1998. Their results reveal that more than 50% of these top-rated funds suffer a rating downgrade, and the degree of persistence varies with the age of funds: older funds show a higher degree of persistence than the younger ones.

Blake and Morey (2000) use a data set covering the period 1992 – 1997 to evaluate the predictive ability of the star ratings. They find little evidence to support the notion that highest-rated (5-star) funds can outperform funds with a lower rating. However, their findings do suggest that poor ratings indicate weak future performance. Morey and Gottesman (2006) re-examine the issue of predictability based on Morningstar's new rating methodology. Analyzing the first batch of funds rated under the new methodology as of June 2005 and tracking their performance over the period July 2002 – June 2005, they find that monotonically, high-rated funds exhibit better post-rating performance than their low-rated counter-parts. For example, two-star funds outperform one-star funds over a three-year post-evaluation period. These results provide evidence that the revamped system does possess some predictive power[3].

In a separate but related line of research, many authors examine the persistence of mutual fund performance. Earlier work by Grinblatt and Titman (1992), Hendricks et al. (1993), Goetzmann and Ibbotson (1994), Brown and Goetzmann(1995) and Malkiel (1995) all provide some evidence that persistence of mutual fund performance exists. Carhart (1997) demonstrates that short-term persistence disappears once the momentum factor of Jegadeesh and Titman (1993) is included in the regression model, thus suggesting that top performing funds generate superior returns simply by holding stocks that have recently performed well. Subsequent studies that employ the Carhart's four-factor model for performance produce mixed results. Bollen and Busse (2005) find evidence of persistence over a

[3] These studies do not directly address the persistence of ratings, but rather their predictive relation to future performance

short three-month horizon. Using bootstrap analysis, Kosowski et al. (2006) document persistence among growth-oriented funds but find no evidence of persistence among income-oriented funds. Huij and Verbeek (2007) apply Bayesian techniques to examine short-run persistence and report that persistence is most pronounced among the relatively younger small cap funds and growth funds.

In this paper, we model the dynamics of mutual fund ratings as a continuous-time Markov process and use the estimated transition probabilities to examine the degree of persistence of ratings. The use of Markov chains to model the evolution of financial ratings can be dated back to the seminal work of Jarrow, Lando and Turnbull (1997) in the context of credit ratings. However, the concept of applying the same approach to mutual fund ratings has only been brought to light recently by the publication of Garnier and Pujol (2007), followed by the related work of Duret et al. (2008) and Herei et al. (2010). Using time-homogeneous discrete-time Markov chains, Garnier and Pujol (2007) study persistence of mutual fund ratings provided by Morningstar and Standard and Poors[4] based on a sample of mutual funds distributed in France over the period 2000 – 2005.

Duret et al. (2008) employ continuous-time Markov chains under the assumption of time-homogeneity to re-examine the sample funds analysed by Garnier and Pujol (2007). Both studies contend that persistence of ratings is weak over a long horizon of three years, although a certain degree of short-term persistence exists. Using monthly Morningstar star ratings of funds distributed in Europe over the period 2000 – 2009, Hereil, Mitaine and Moussavi (2010) propose a new measure of persistence based on their estimated transition probabilities. This measure

[4] Morningstar acquired Standard & Poors' fund data business, including fund ratings, in 2007

corresponds to the time during which the probability of having no rating changes exceeds the probability of a rating migration. Focusing on the top rating, they find that the average duration time of the 5-star rating is approximately 5 months.

None of the above studies addresses the important issue of time homogeneity. Although Hereil et al. (2010) acknowledge that the assumption of time homogeneity is probably tenuous, they conduct no further examination on its validity. Time homogeneity of transition probabilities is not an innocuous assumption and warrants verification. In the literature on the use of Markov models for credit ratings, there has been mounting evidence against time homogeneity. For example, both Nickell et al. (2000) and Bangia et al. (2002) demonstrate that transition probabilities of credit ratings vary with the business cycles. In the light of this, we consider a range of issues central to the estimation of transition probabilities. With the aid of standard statistical tests of Anderson and Goodman (1957), we first assess the validity of time-homogeneity in a robust manner, using data that covers different time periods and over different time horizons. Having derived strong statistical evidence that the Morningstar star ratings is not a time-homogeneous process, we estimate the transition probabilities based on the non-parametric Aalen-Johansen estimator (Anderson, Hansen and Keiding 1991, Gill and Johansen 1990) over a 72-month period.

With the estimated transition probabilities over a long time horizon, we proceed to conduct a study on the dynamics of the Morningstar star ratings for different types of mutual funds. Keeping the interests of an individual investor in mind, we seek to answer the questions: How likely a 5-star rated mutual fund can

retain its rating in the future and once rated, what is the duration of a 5-star rating. It is also worthwhile to investigate how likely a 4-star fund can be upgraded to a 5-star rating over a specific time horizon Essentially, the transition probability that a rating remains unchanged between two observation months measures the strength of persistence of the rating, while the transition probability corresponding to a change of rating from a 5-star to a lower rank indicates the likelihood of a downgrade.

Given that the star rating is calculated from a long-term (at least three years) past risk-adjusted return (see Morningstar, 2009]), our paper contributes to the literature by examining long-term performance persistence of mutual fund within the Markov chain framework, using the star rating as the performance measure.

Having both the star ratings and the stewardship grade, we are also in a position to investigate whether corporate governance plays a role in persistence and transitions of ratings. For example, by dividing funds with a 5-star rating into groups based on a fund's stewardship grade, we can examine whether the degree of rating persistence is influenced by the standard of corporate governance. For this purpose, we divide the sample funds into categories based on a fund's stewardship grade and analyze the differences in transition probabilities across different stewardship grade groups. It turns out that funds with better governance generally exhibit a lower chance of rating downgrade and a higher chance of rating upgrade. Results of this kind provide some credence to the widely accepted notion that corporate governance plays a positive and significant role in protecting the interests of the shareholders i.e. the mutual fund investors.

The scope of this paper can be broadened to the use of Markov chains in examining the persistence of mutual fund performance, extending the commonly-used "Winner-Loser" methodology adopted by many authors, such as Goetzmann and Ibbotson (1994), Brown and Goetzmann (1995) and Malkiel (1995). One obvious advantage that the Markov chain approach has over the WL method is that one can examine many more states (five in the case of Morningstar Fund ratings) than a simple dichotomy.

The outline of this paper is as follows. In Section 2, we give an overview of the theory of discrete-time and continuous-time Markov chains that is relevant to this study and discuss related issues such as the statistical tests for time homogeneity, estimation of transition matrices and derivation of some useful persistence measures from the estimated transition probabilities. Section 3 describes the data set and presents the empirical findings. We conclude this paper and offer suggestions for further research in Section 4.

2.2 Markov Chain Models for Mutual Fund Ratings

Markov Chains

A typical mutual fund rating system produces a time series of ordinal variables (ratings) $\{R_1, R_2, ..., R_K\}$ for which $R_i < R_j$ if $i < j$. These ratings are typically based on some measure of past performance. For the case of Morningstar star rating system, risk-adjusted returns based on expected utility theory are used to rank mutual funds within specific peer groups on a monthly basis. The end-product is a set of 5 grades (1-star, 2-star, 3-star, 4-star and 5-star).

A stochastic process $\{R(t): t = 0, 1, 2, ...\}$ is said to be a discrete-time Markov chain on a finite set S = {1, 2, .., *K*} if for all integers $t_0, t_1, t_2, ..., t_n$ with

$$t_n < t_{n-1} < ... < t_1 < t_0 < t,$$

the following *Markov property* holds

$$\begin{aligned} &\mathrm{P}\left(R_t = j \mid R_{t_0} = i, R_{t_1} = i_1, R_{t_2} = i_2, .., R_{t_n} = i_n\right) \\ &= \mathrm{P}\left(R_t = j \mid R_{t_0} = i\right) := p_{ij}(t_0, t) \end{aligned}$$

The set S is known as the state space of the Markov chain and the probability $p_{ij}(t_0, t)$ is called the $(t - t_0)$ – step transition probability for transition from state *i* at time t_0 to state *j* at time *t*. The $K \times K$ square matrix

$$\mathbf{P}\left((t_0, t)\right) = \left(p_{ij}(t_0, t)\right)$$

is known as the transition matrix for the interval [t_0 , *t*]. By the definition of transition matrices, $0 \le p_{ij}(t_0, t) \le 1$ and each row sum, $\sum_{j=1}^{K} p_{ij}(t_0, t)$ of the transition matrix equals to unity.

A discrete-time Markov chain is said to be time-homogeneous if the transition probability $p_{ij}(t_0, t)$ depends only on the duration $t - t_0$ over which transition of states occurs. In this case, $p_{ij}(t_0, t)$ can be written as $p_{ij}(t - t_0)$ and the corresponding transition matrix is given by

$$\mathbf{P}\left(t - t_0\right) = \left(p_{ij}(t - t_0)\right)$$

It follows that for any time-homogeneous Markov chain and for any integers *n* > *m*,

$$\mathbf{P}(n) = \mathbf{P}(m) \times \mathbf{P}(n - m)$$

and

$$\mathbf{P}(n) = (\mathbf{P}(1))^{n}.$$

In this paper, we consider monthly Morningstar ratings for which the state space is $S = \{1, 2, 3, 4, 5\}$. For example, $\{R_t = i\}$ is the event that a fund has a i – star rating at month t .

When the full rating histories including the exact transition times are known, the continuous-time Markov chains offer significant advantages. The evolution of continuous-time Makov chains is described via the so-called transition intensities $\mathbf{\Lambda}(t) = (\lambda_{ij}(t))$, $1 \le i, j \le K$, given by

$$\lambda_{ij}(t) = \lim_{h \to 0+} \frac{\mathrm{P}(R_{t+h} = j \mid R_{t_0} = i)}{h}$$

Transition matrices can then be derived from the above quantities. For a mathematical exposition, see the Appendix in Lando and Skødeberg (2002).

Estimation of Transition Probabilities

As we shall demonstrate in Section 3, our statistical tests reject the hypothesis of time-homogeneity for our data. Hence, it suffices to discuss in this section the estimator for non-homogenous Markov chains.

For a non-homogenous case Markov chain, the rating process is characterized by a transition matrix **P**(*s*, *t*) over the period (*s*, *t*). Following Aalen and Johansen (1978), Gill and Johansen(1990), the transition probabilities over a one-month period $[m, m+1]$ can be estimated using the following Aalen-Johansen (or the so-called product-limit) estimator:

$$\hat{\mathbf{P}}(m, m+1) = (\mathbf{I} + \hat{\mathbf{A}}(m)),$$

where **I** is the identity and $\hat{\mathbf{A}}(m)$ is given by

$$\hat{\mathbf{A}}(m)=\begin{pmatrix} -\frac{\tilde{n}_1(m)}{n_1(m)} & \frac{n_{12}(m)}{n_1(m)} & \frac{n_{13}(m)}{n_1(m)} & \frac{\tilde{n}_{14}(m)}{n_1(m)} & \frac{n_{15}(m)}{n_1(m)} \\ \frac{n_{21}(m)}{n_2(m)} & -\frac{\tilde{n}_2(m)}{n_2(m)} & \frac{n_{23}(m)}{n_2(m)} & \cdots & \frac{n_{25}(m)}{n_2(m)} \\ \cdots & \cdots & -\frac{\tilde{n}_3(m)}{n_3(m)} & \cdots & \cdots \\ \cdots & \cdots & \cdots & -\frac{\tilde{n}_4(m)}{n_4(m)} & \cdots \\ \frac{n_{51}(m)}{n_5(m)} & -\frac{\tilde{n}_{52}(m)}{n_5(m)} & \cdots & \cdots & -\frac{\tilde{n}_5(m)}{n_5(m)} \end{pmatrix}$$

Here, $\tilde{n}_i(m)$ counts the number of rating transitions away from state *i* from month *m* to month (*m* + 1), $n_{ij}(m)$ counts the number of transitions from state *i* at month *m* to state *j* at month (*m* + 1) and $n_i(m)$ denotes the total number of mutual funds with a rating of *i* at month *m*.

It is easy to see that each row of $\hat{\mathbf{A}}(m)$ sums to zero. A *k*-month transition estimator, $\hat{\mathbf{P}}(0,k)$ over the period [0, *k*] can then be obtained via matrix multiplication of the corresponding one-month transition matrices,

$$\hat{\mathbf{P}}(0,k)=\prod_{m=0}^{k-1}\left(\mathbf{I}+\hat{\mathbf{A}}(m)\right).$$

One obvious advantage of the above estimator is that it utilizes all available information on rating transitions in the data set including the exact time (month) at which rating changes take place.

Probabilities of Rating Persistence and Rating Change

Hereil et al. (2010) propose the use of following function to derive measure of rating persistence

$$d_i(t) = \mathrm{P}\left(R_t = j \middle| R_0 = i\right) - \mathrm{P}\left(R_t \neq i \middle| R_0 = i\right) \\ = 2\mathrm{P}\left(R_t = j \middle| R_0 = i\right) - 1 \quad ,$$

for each $i = 1, 2, .., K$. A positive $d_i(t)$ indicates a higher chance of maintaining an initial of *i*-star. Hereil et al. (2010) also suggest that the following quantity be used to measure how resilient a particular rating is

$$\tau_i = \inf\left\{t > 0 : d_i(t) \leq 0\right\}.$$

Clearly, τ_i represents the earliest time at which migration of state *i* is more likely to occur than not.

For different pairs of ratings, *i* and *j*, the time series of the transition probabilities $\left\{ p_{ij}(0,t) : t = 1, 2, 3... \right\}$ can provide useful insights on the evolution of mutual fund ratings. For the case when *i* = *j*, one can use these transition probabilities $p_{ii}(0,t)$ to measure the persistence of the rating *i* since $p_{ii}(0,t)$ gives the odds that a fund retains its initial rating of *i* after *t* months. For *j* > *i*, the probabilities $p_{ij}(0,t)$ indicate the chance of a rating upgrade from the initial rating of *i* to a higher rating of *j*. Of particular interest is the case when *i* = 4 and *j* = 5, since investors are interested in how likely a 4-star fund can be upgraded to 5-star in subsequent months. Similarly, for *j* < *i*, the probabilities $p_{ij}(0,t)$ represent the chance of a rating downgrade from *i* to *j* . The case when *i* = 5 and *j* = 4 may be of special significance to investors for a similar reason.

2.3 Data and Empirical Analysis

Data Description

The data covers 79 months of rating (Morningstar star ratings and stewardship grades) history from November 2004 through May 2011. There are altogether 1803 mutual funds, all of which are classified under the Morningstar's 'US Broad Asset Class' as balanced funds, bond funds ('municipal bond' or 'taxable bond'), 'international stock funds' or 'U.S. stock' fund. For simplicity, we shall enumerate the months as month 1 – month 79.

Ratings Statistics and Dynamics

We conduct a preliminary analysis of rating persistence by tracking the time series average star rating of funds with the same initial rating over a post-ranking period of duration up to 72 months. Specifically, we sort the sample funds into five portfolios based on their initial (month 1, November 2004) star rating (SR) (1-star to 5- star). Table 1 displays the average star rating of each portfolio computed at quarterly intervals over the next 78 months. It is obvious that each time series of average rating exhibits a long-term "mean-reversion" pattern with the average rating for each portfolio stabilizing at an average of about 2.5 to 3.5, thus suggesting that the level of rating persistence is low over a long (3-year to 6-year) time horizon.

Time Homogeneity Test

Following Anderson and Goodman (1957), we divide a test period of duration 13 months into 12 non-overlapping subintervals $[t-1, t]$, $t = 1, 2, .., 12$ and apply the cohort method on each of these sub-intervals to obtain an estimate $\hat{p}_{ij}(t-1,t)$ for the transition probability $p_{ij}(t-1,t)$, and also on the entire 12-month time period to obtain the estimate $\hat{p}_{ij}(0,12)$ for each fixed i, $i = 1, 2,3\ldots, 5$.

For testing the null hypothesis that $p_{ij}(t-1,t)$ does not depend on t, given a specific i, the following test statistic has an asymptotic chi-squared distribution with 44 degrees of freedom[5]

$$\chi_i^2 = \sum_{t,j} \frac{n_i(t)\left(\hat{p}_{i,j}(t) - \hat{p}_{i,j}\right)^2}{\hat{p}_{i,j}},$$

where we have used the notations in Appendix A. For testing the joint hypotheses that $p_{ij}(t-1,t)$ is independent of t for all possible i, test statistics

$$\sum_i \chi_i^2 = \sum_i \sum_{t,j} \frac{n_i(t)\left(\hat{p}_{i,j}(t) - \hat{p}_{i,j}\right)^2}{\hat{p}_{i,j}}$$

has an asymptotic chi-squared distribution with 220 degrees of freedom.

We perform the above test using five non-overlapping intervals of length 12 months, and repeat the test using four test periods each of length 36 months. For the latter, we divide each test period into three non-overlapping sub-periods of length 12 months. Table 2 reports the results of these tests.

Overall, there is overwhelming evidence against time-homogeneity. The hypothesis of time-homogeneity is rejected in all cases. Although not reported, we

[5] For K states and T transition periods, the number of degree of freedom is $(K-1)(T-1)$

have also conducted the test based on four subsamples formed by dividing the sample into four groups based on a fund's 'Morningstar U.S. Broad Asset' category: 'U.S. Stock' , 'International Stock' , 'Bond' ('Taxable Bond' or 'Municipal Bond') and 'Balanced'. We reach the same conclusion. It follows that the Aalen-Johansen estimator is the appropriate estimator for the transition matrices.

Persistence Probability and Probability of Rating Upgrade or Downgrade

With the aid of the estimated transition probabilities, we proceed to investigate whether the degree of rating persistence /upgrade /downgrade will be differentiated by the type of funds being examined. For this purpose, we split the sample into four groups according to the four categories - 'U.S. stock', 'international stock' , 'bond' ('taxable bond' or 'municipal bond') and 'balanced' funds under Morningstar's 'US Broad Category' classification. Figure 1 displays the graphs of selected transition probabilities against time. More precisely, we consider transition among the three states: 3-star, 4-star and 5-star funds. Of particular interest would be rating upgrade from 4-star to 5-star and rating downgrade form 5-star to 4-star. On the whole, we find no obvious evidence of heterogeneity in the transition probabilities among the four types of funds. The only notable exception is the case of a rating upgrade from 3-star to 5-star where bond funds outperform the other groups.

We proceed to investigate whether corporate governance makes a difference to the above results. To this end, we use a fund's stewardship grade at month 1 (November 2004 data) to partition the sample funds into two SG (stewardship grade) groups. Specifically, funds with SG = 'A' are placed in the 'Good SG' group while

funds with SG = 'D' or 'F' are placed in the 'Poor SG' group. We then compute the estimated transition matrices from each of these subsamples. Figure 2 presents the graphs of estimated transition probabilities $\hat{p}_{i,j}(0,t)$ versus time, where $i, j = 3, 4, 5$ and $t = 1, 2, ...,73$ denotes month with month 1 = November 2004.

These graphs provide strong evidence on the positive relation between stewardship grade and mutual fund rating changes. For rating upgrades (3-star to 4-star or 4-star to 5-star), funds with a good initial stewardship grade exhibit a higher transition probability (hence, higher chance of upgrading) than funds with a poor initial stewardship grade over the entire 72-month observation period. We can draw a similar conclusion by examining rating downgrades (4-star to 3-star or 5-star to 4-star). Specifically, funds with a good initial stewardship grade have a lower transition probability (hence, lower chance of downgrading) than funds with a poor initial stewardship grade over the entire 72-month observation period. On the flip side, as depicted in the bottom right panel of figure 2, there is no convincing evidence that funds with better stewardship grade have a better chance of keeping the top 5-star rating over a short time horizon of one to two years, though over a longer time span, 5-star funds with better stewardship grades do show a higher chance of maintaining the top rating.

To complement the above qualitative analyses with quantitative results, we perform both Wilcoxon signed rank test and paired-sample *t* – tests with Newey-West adjusted standard errors on the time series of differences in transition probabilities. More specifically, letting $p_{i,j}^{(\text{good})}(t)$ denote the transition probability, Prob(Rating

after *t* month = *j*-star | Initial Rating at month 1 = *i*-star and initial stewardship grade = 'A') and $p_{i,j}^{(\text{poor})}(t)$ denote the transition probability, Prob(Rating after *t* month = *j*-star | Initial Rating at month 1 = *i*-star and initial stewardship grade = 'D' or 'F'), we test the hypothesis that the difference $p_{ij}^{(\text{good})}(t) - p_{ij}^{(\text{poor})}(t)$ is significantly positive for upgrades ($j > i$) and significantly negative for downgrades ($j < i$). Table 3 presents the results.

Following Hereil et al. (2008), we compute the quantity $d_i(s,t) = 2p_{ii}(s,t) - 1$ for each initial rating *i* and for three different initial months *s* (namely, month 1 (November 2004), month 13(November 2005), month 25 (November 2006)), and plot the graphs of $d_i(s,t)$ against time *t* in Figure 3 based on the entire sample. Two salient observations can be made from these graphs. First, the persistence of each rating, as measured by the quantity

$$\tau_i = \inf\{t > 0 : d_i(t) \le 0\},$$

(graphically being the time at which the graph of $d_i(s,t)$ crosses the horizontal axis) is very low for all ratings, invariably for less than a year. Second, 3–star rating exhibits the highest degree of persistence while the best (5–star) and the worst (1–star) rating show the lowest.

Finally, as a robustness check for our earlier results on the relation between stewardship grade and rating transition probability, we re-construct the graphs in Figure 2 for four different starting months : month 1 (November 2004), month 13, month 25 and month 37 over a period of 36 months. To conserve space, we display in Figure 4 only the graphs for rating transitions from 4-star rating. We find that despite

non-homogeneity, the relation between stewardship grade and the probability of rating changes are robust to the choice of the starting month.

Summing up, the above results suggest that persistence of the star rating is weak regardless of the type of funds assessed. In particular, the top-rated funds that are most favored by investors have a high chance (probability > 0.4) of suffering a rating downgrade. On a positive note, corporate governance appears to have a positive impact on rating persistence and rating changes, suggesting that investors should consider not just past performance, but also the standard of corporate governance when selecting mutual funds.

2.4 Conclusions

A typical empirical method of examining the persistence of mutual fund performance is the ranked-portfolio test in which during a so-called formation period, the sample funds are first sorted and placed in two or more groups based on some performance measures (such as the Sharpe ratio and Carhart's four factor alpha). Using the same performance metric, the mean return of each portfolio is then computed during a subsequent period known, as the evaluation period, and the portfolios are re-sorted based on the mean portfolio return. The strength of the correlation between the formation-period ranking and evaluation-period ranking is a measure of performance persistence. In particular, in the studies by Goetzmann and Ibbotson (1994), Brown and Goetzmann (1995) and Malkiel (1995), a fund is categorized as a winner (loser) if its performance is above (below) the median fund return. This gives rise to a 2 by 2 contingency table of winners and losers in the two ranking periods. Standard statistical tests of two-way tables based on log odds ratio

can then be applied to test the significance of occurrences of 'Winner-Winner' ('Loser-Loser') which indicates a high degree of persistence of good (poor) performance.

In this paper, we model mutual fund rating as a continuous-time Markov chain and use the transition probabilities to examine the persistence of ratings. This method can be viewed as a generalization of the above "Winner-Loser" dichotomy. One issue that arises from the use of a Markov process is the choice of estimators for the transition matrices which depends on whether the true transition probabilities are time-homogeneous. Following standard statistical procedures in Anderson and Goodman (1957), we conduct asymptotic chi-square tests of time-dependence using different time periods and durations to ensure robustness of our results. It turns out that the null hypothesis of time-homogeneity is rejected in all the tests, leading us to conclude that the Aalen-Johnsen estimator is the appropriate tool for estimating transition probabilities.

Using the estimated 72-month transition matrices for both the entire sample as well as for subsamples based either on a fund's category (bond, balanced, international stock or U.S. stock) or on a fund's corporate governance rating (stewardship grade), we examine the strength of rating persistence as well as the likelihood of rating upgrades or downgrades over time and investigate whether the results vary across different fund categories or across different corporate governance groups. Results obtained suggest that while persistence of rating is feeble, funds with a good stewardship grade tend to have a greater chance of maintaining their top star rating (4-star) or improving their rating (4-star to 5-star), and a smaller chance of

experiencing a rating downgrade. These results provide some support for the contention that corporate governance and performance go hand in hand.

The results on the persistence of ratings warrant caution on the use of ratings. The lack of strong persistence of the 5-star rating over a short horizon underscores the need for investors to consider criteria besides ratings when selecting funds. The evidence of the positive role that good corporate governance plays in rating persistence is an indication that investors would be better off if they take into consideration both the star rating and the stewardship grade when making investment decisions.

We would like to point out that many issues surrounding the persistence of ratings have not been addressed in this paper but shall be addressed in future research. For example, one can seek to determine factors that can affect the degree of persistence of ratings. A simple approach to this problem would be to sort the sample funds by specific fund characteristics, such as managerial experience and turnover ratio, that can potentially affect the rating dynamics, and form equal-size subsamples (e.g. quintiles). One can then compare the transition matrices estimated from the subsamples using well-known metrics that quantify and measure the difference between two transition matrices (see e.g. Jafry and Schuermann, 2004, Truck and Rachev, 2011).

Another notable omission from this paper is the impact that fund exit from and fund entry into the rating system has on the empirical results. One possible extension of our model is one that includes an additional state, besides the five states that represent the ratings, to incorporate the events of fund entry or fund exit. If we

assign the value 6 to this state, then a transition from state 6 to 5 represents the initiation of the best rating.

Finally, it has been documented in a number of publications (Bangia et al. 2002, Lando and Skodeberg, 2002) that credit ratings of firms exhibit a so-called rating drift, a phenomenon whereby firms that have been downgraded (upgraded) before are likely to be downgraded (upgraded) in the next period. To investigate the phenomenon of rating drift, one has to separately estimate three conditional transition probabilities: one conditioned on upgrading in the prior period, one conditioned on downgrading in the prior period and one conditioned on no rating change in the prior period. It would be interesting to see how the issue of rating momentum can be addressed in the context of mutual fund ratings.

Table 1

Average Star Rating Over Time

This table reports the average star rating over time of mutual fund portfolios formed by a fund's initial rating. The sample funds are sorted into five portfolios by their initial (month 1, November 2004) star rating (SR) (1-star to 5-star). The cross-sectional subsequent average star rating of each portfolio, computed at quarterly intervals over the next 78 months, are displayed.

	Initial Star Rating				
	1	2	3	4	5
Month					
4	1.24	2.11	3.02	3.87	4.77
7	1.32	2.14	3.02	3.83	4.74
10	1.71	2.18	3.03	3.81	4.67
13	1.85	2.21	2.99	3.76	4.65
16	1.88	2.31	2.99	3.71	4.56
19	1.74	2.28	3.00	3.72	4.57
22	1.62	2.28	2.95	3.69	4.48
25	1.85	2.32	2.98	3.66	4.47
28	1.88	2.41	2.98	3.62	4.43
31	2.03	2.51	3.07	3.63	4.36
34	2.26	2.59	3.05	3.58	4.28
37	2.21	2.62	3.04	3.58	4.23
40	2.35	2.60	3.05	3.56	4.19
43	2.47	2.66	3.04	3.56	4.17
46	2.44	2.66	3.04	3.57	4.07
49	2.44	2.69	3.01	3.43	3.78
52	2.50	2.69	3.03	3.39	3.78
55	2.50	2.64	3.00	3.43	3.82
58	2.65	2.63	2.99	3.43	3.77
61	2.59	2.60	2.99	3.42	3.79
64	2.50	2.58	2.98	3.40	3.77
67	2.53	2.61	2.98	3.39	3.76
70	2.59	2.61	2.98	3.37	3.80
73	2.50	2.66	2.96	3.36	3.80
76	2.44	2.69	2.95	3.37	3.76
79	2.35	2.78	2.97	3.34	3.73
N	34	170	423	462	225

Table 2

Statistical Tests of Time Homogeneity

This table reports results of statistical tests of time homogeneity of Morningstar star rating using five test periods, each of length 12 months, for each possible rating (i =1 for one-star to i = 5 for 5-star). Each period is divided into T = 12 monthly in (Panel A) and T = 3 yearly (in Panel B) test periods. *, **, and ** indicate significance (against the null hypothesis) at the 10, 5 and 1% level of significance respectively.

Panel A (*T* = 12 sub-periods of duration 1 month)

	Test for individual i		Joint Test
Test Period	*i*	Test Statistics	Test Statistics
month 1 - 13	1	100.3***	1134.93***
	2	284.85***	
	3	354.78***	
	4	310.93***	
	5	84.07***	
month 13 - 25	1	173.99***	908.45***
	2	248.9***	
	3	122.17***	
	4	275.11***	
	5	137.29***	
month 25 - 37	1	126.26***	1458.35***
	2	291.37***	
	3	481.29***	
	4	350.09***	
	5	209.33***	
month 37 - 48	1	292.65***	3930.53***
	2	1035.17***	
	3	1279.24***	
	4	946.02***	
	5	377.46***	
month 49 - 61	1	103.07***	981.53***
	2	216.23***	
	3	311.09***	
	4	243.74***	
	5	107.4***	
month 61 - 73	1	57.21***	757.34***
	2	114.33***	
	3	235.87***	
	4	234.34***	
	5	115.58***	

Panel B (T = 3 sub-periods of duration 12 months

	Test for individual i		Joint Test
Test Period	*i*	Test Statistics	Test Statistics
month 1 -37	1	35.23***	361.71***
	2	66.58***	
	3	84.18***	
	4	123.35***	
	5	52.37***	
month 13 -49	1	363.39***	2904.40***
	2	631.26***	
	3	793.04***	
	4	642.68***	
	5	474.04***	
month 25 - 61	1	387.01***	2600.91***
	2	509.97***	
	3	663.13***	
	4	607.76***	
	5	439.04***	
month 37 - 73	1	663.98***	4399.98***
	2	974.19***	
	3	1142.99***	
	4	982.69***	
	5	636.13***	

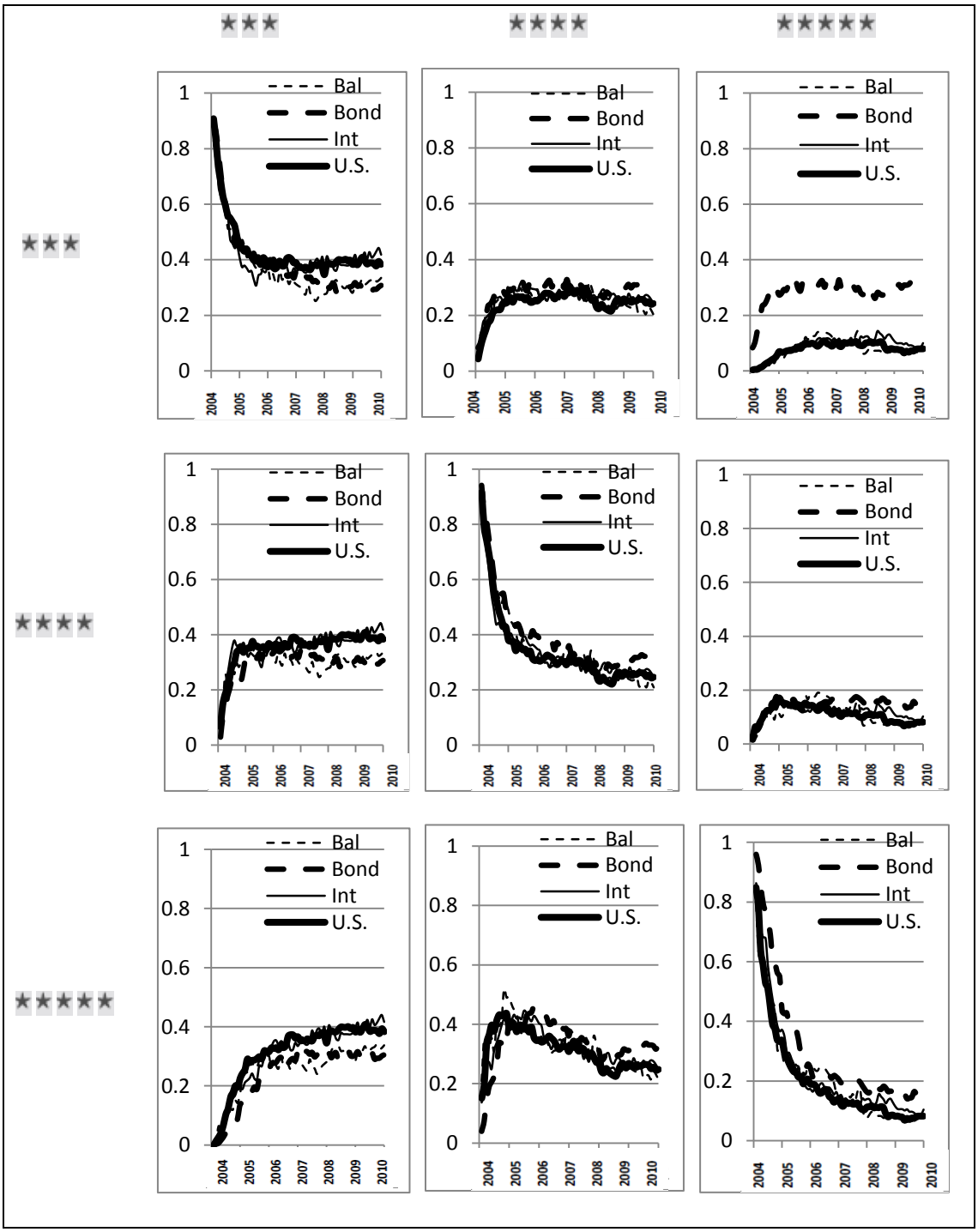


Figure 1. Graphs of Transition Probabilities over Time for Various Fund Categories**.** In this figure, the graph in the i-th row and j-th column shows the transition probabilities from state i to state j, where the state are indicated by the symbols (★★★, ★★★★ or ★★★★★) representing the star rating of the funds: , for each of the four groups of funds classified as 'U.S. Stock' , 'International Stock' , 'Bond' ('Taxable Bond' or 'Municipal Bond') and 'Balanced' funds categorised under Morningstar's 'US Broad Category' Classification). These funds have received monthly star ratings (Abbrev SR) over the period November 2004 (Month 1) – December 2010 (Month 73).

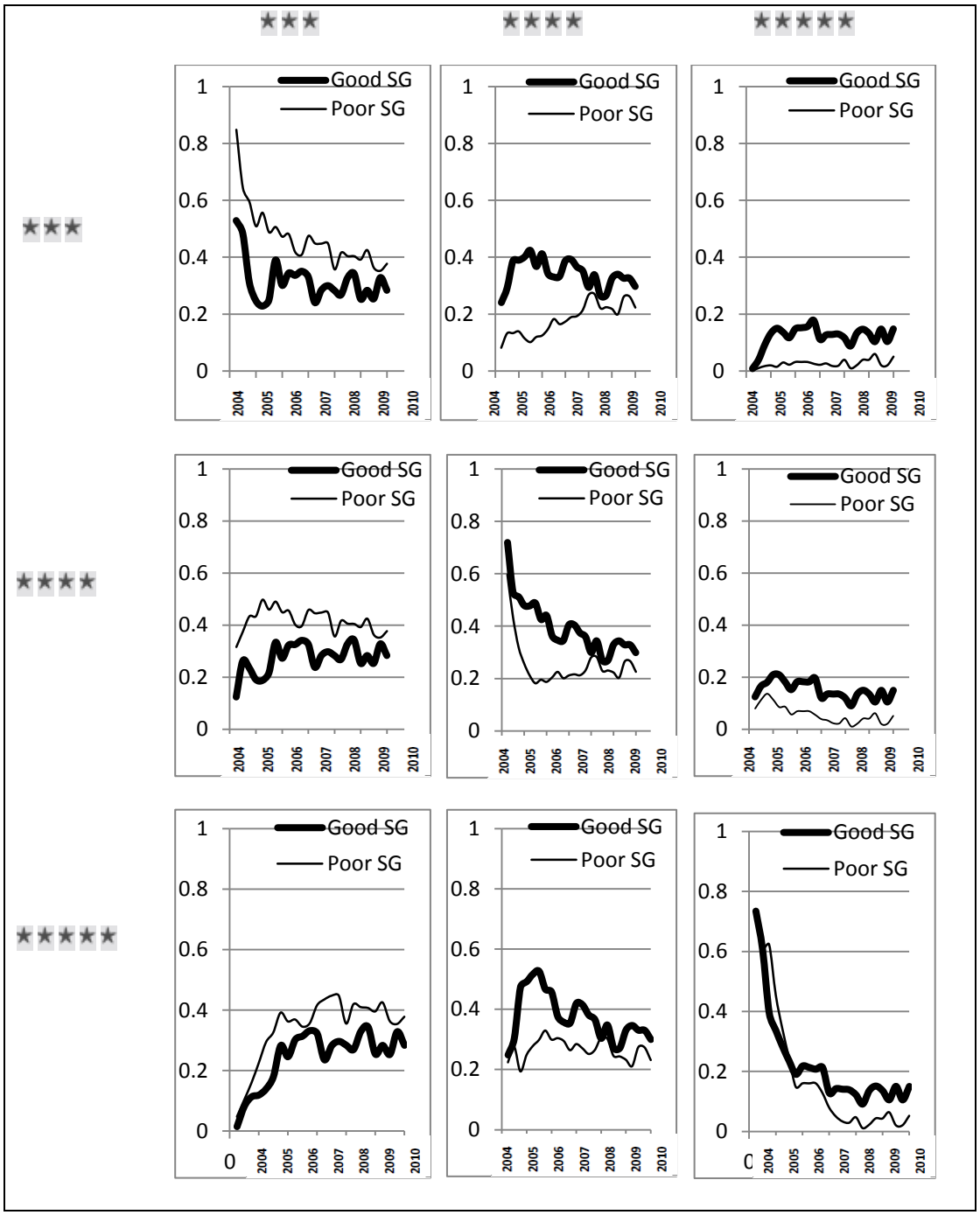


Figure 2. Graphs of Transition Probabilities over Time for Top and Bottom Stewardship Grade funds. In this figure, the graph in the i-th row and j-th column shows the transition probabilities from state i to state j, where the possible states are indicated by the symbols representing the star rating of the funds (★★★ , ★★★★ and ★★★★★) , for each of the two rating group – 'good stewardship grade' and 'poor stewardship grade'. Funds with SG = 'A' are placed in the 'Good SG' group while Funds with SG = 'D' or 'F' are placed in the 'Poor SG' group. These funds have received monthly star ratings (Abbrev SR) over the period November 2004 (Month 1) – December 2010 (Month 73) and have a stewardship grade (SG) reported for the month of November 2004 (month 1).

Table 3

Statistical Tests of Differences inTransition Probabilities

This table shows the Newey-West adjusted t-statistics and statistics (in parentheses) based on Wilcoxon's signed rank. *** indicates significance at the 1% level.

	3-star	4-star	5-star
3-star	-8.27***	7.22***	12.46***
	(-1313***)	(1308***)	(1278***)
4-star	-9.19***	7.07***	16.14***
	(-1313***)	(1304***)	(1313***)
5-star	-9.07***	(5.49***)	(3.32***)
	(-1268***)	(1219***)	(920***)

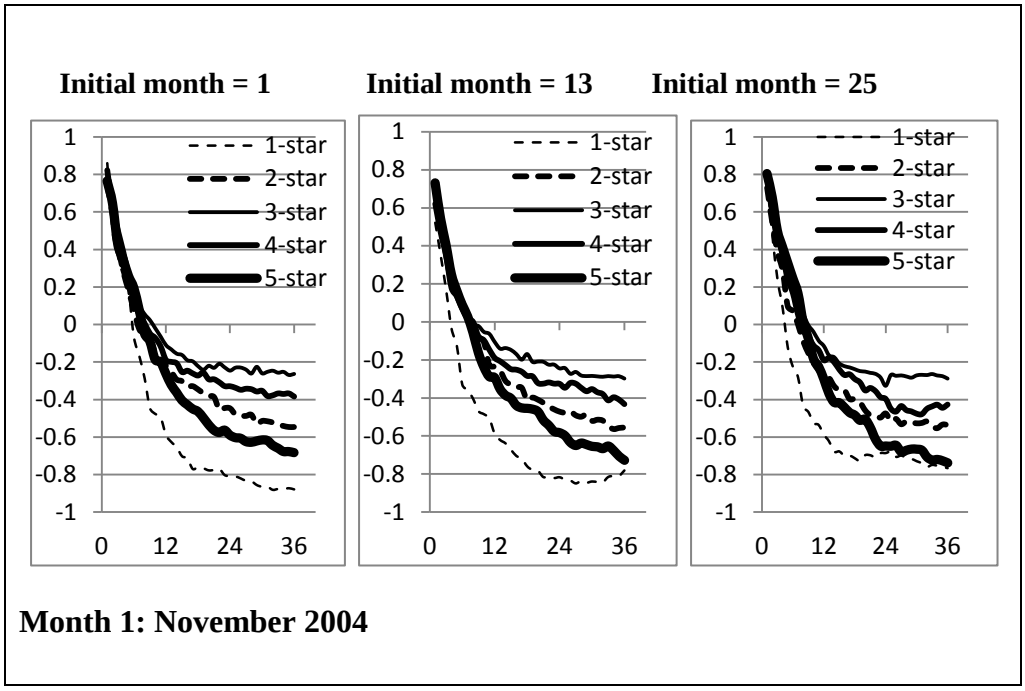
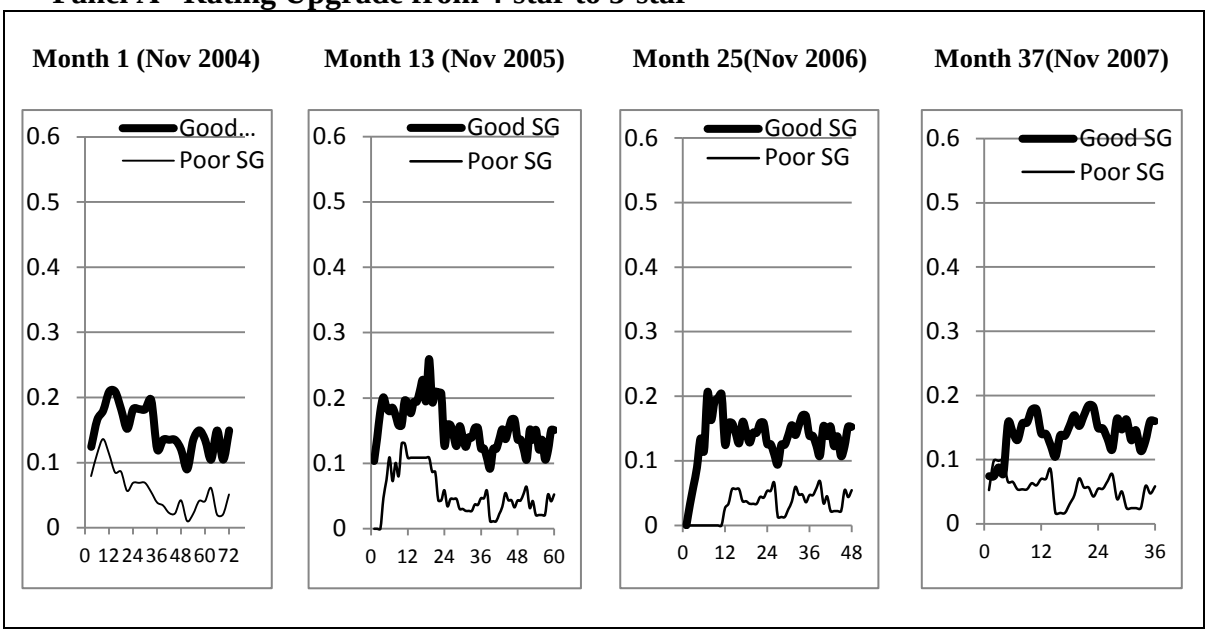


Figure 3. Graphs of Persistence Measures over Time. The figures above are the graph of d_i (s, t) versus time for initial month s = 1 (November 2004), 13 (November 2005) an 25 (November 2006) and for all five initial rating i (1-star to 5-star) and The sample consisting of funds classified as 'U.S. Stock' , 'International Stock' , 'Bond' ('Taxable Bond' or 'Municipal Bond') and 'Balanced' funds under Morningstar's 'US Broad Category' classification that receive monthly star ratings (Abbrev SR) over the period November 2004 (Month 1) – December 2010 (Month 73). The time t at which each graph first crosses the time-axis is a measure of persistence of rating i.

Panel A Rating Upgrade from 4-star to 5-star



Panel B Rating Downgrade from 4-star to 3-star

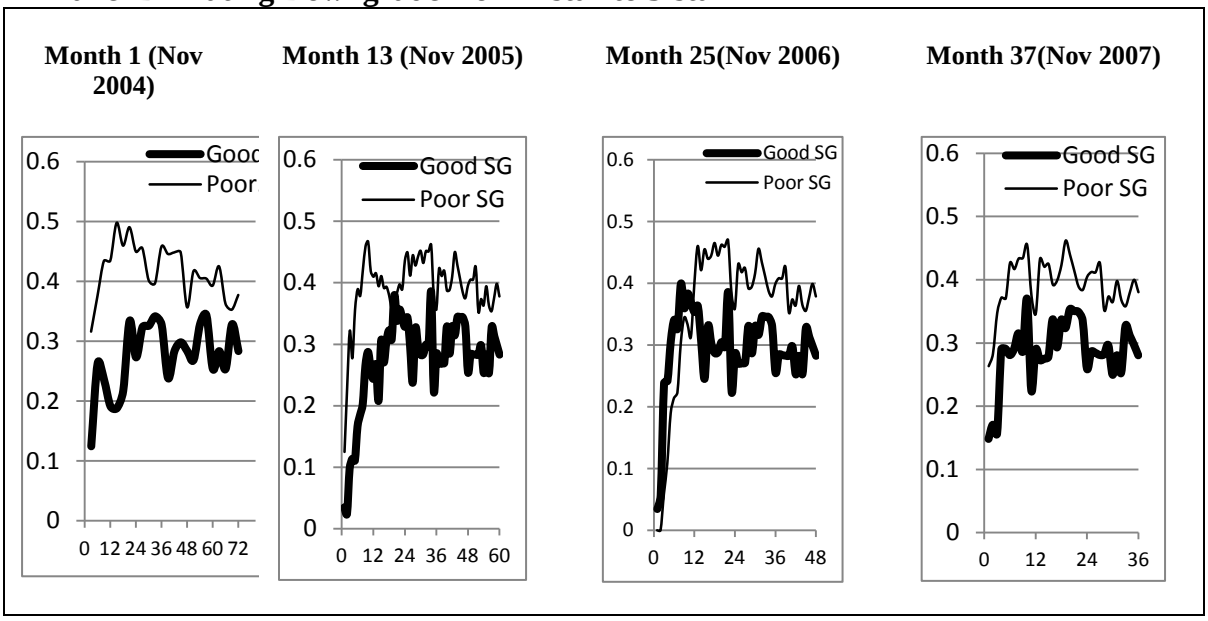


Figure 4. Graphs of Transition Probabilities for Different Starting Months. The figures below are the graphs of the time series of transition probabilities (probabilities of transition from 4-star to 5-star in Panel A and probabilities of transition from 4-strar to 3-star in Panel B) for different starting months (months 1(November 2004), month 13 (November 2005) an month 25(November 2006)) over a duration of 36 months. The sample consisting of funds classified as 'U.S. Stock' , 'International Stock' , 'Bond' ('Taxable Bond' or 'Municipal Bond') and 'Balanced' funds under Morningstar's 'US Broad Category' classification that receive monthly star ratings (Abbrev SR) over the period November 2004 (Month 1) – December 2010 (Month 73). .

Appendix

Two Classical Methods of Estimating Transition Matrices

Two classical methods of estimation transition probabilities are the cohort/multinomial method and the duration/continuous-time method. While the former is formulated in the discrete-time framework, the latter is used in the continuous-time setting.

Cohort -Method

Let us consider a sample of N mutual funds whose monthly rating, assumed to be an integer between 1 and 5 inclusive, are observed at the end of each month, over a period of $(T + 1)$ months from $t = 0$ to $t = T$ inclusive. For references hereafter and in other parts of this paper, we introduce the following notations:

(i) $n_i(t)$ = the total number of mutual funds with a rating of i at month t

(ii) $n_{ij}(t)$ = the number of rating transitions from i at time t - 1 to j at time t

(iii) $N_i(T) = \sum_{t=0}^{T-1} n_i(t)$, the total number of mutual funds with a rating of i – observed at time $t = 0, 1, \ldots, T - 1$.

(iv) $N_{ij}(T) = \sum_{t=0}^{T-1} n_{ij}(t)$, the total number of rating transitions from i at time t - 1 to j throughout the entire observation period

A simple method of calculating transition probabilities is the cohort method based on maximum likelihood principle. Specifically, the maximum-likelihood estimator for a one-month transition probability $p_{ij}(t)$ over the period [t - 1, t] is given by

$$\hat{p}_{ij}(t) = \frac{n_{ij}(t)}{n_i(t)}.$$

If we view the rating process as a time-homogeneous Markov chain, then the transition of ratings can be perceived as independent multinomial events, so that all the observations made over multiple periods can be aggregated as one large data set. Therefore, the transition probability p_{ij} over the time period [0, t] can be estimated by

$$\hat{p}_{ij} = \frac{N_{ij}(t)}{N_i(t)}.$$

One potential problem with the above maximum-likelihood method is that some estimated transition probabilities may turn out to be zero. To see why this might occur, consider a rating change from a 1-star grade to a 5-star grade. Based on our data, for a short period of one-year, there is no occurrence of such an event. This translates into $n_{15}(t) = 0$, and hence $\hat{p}_{15}(t) = 0$ for each t. However, theoretically, it is not impossible for such events to occur. Hence, the cohort method fails to capture rare events such as rating upgrades or downgrades by several notches. In their study of credit ratings, Lando and Skødeberg (2002) suggest the use of a continuous-time formulation when the full rating history is available. They argue that a continuous-time model only solves the afore-mentioned problem, but also facilitates the estimation of transition matrices over arbitrary time-horizons.

Duration-Method

As discussed in the preceding section, for a continuous-time Markov chain, the transition matrix $\mathbf{P}(0,t)=\left(p_{i,j}(0,t)\right)$ over the period [0, t], for any $t > 0$, can be calculated from a generator matrix $\mathbf{\Lambda}=\left(\lambda_{ij}\right)$whose elements satisfy the conditions

$$\lambda_{ij} \geq 0 \text{ for } i \neq j \quad \text{and} \quad \lambda_{ii} = -\sum_{j \neq i} \lambda_{ij} .$$

The transition matrix $\mathbf{P}(0, t)$ is given by

$$\mathbf{P}(0,t) = \exp(t\mathbf{\Lambda}) ,$$

where $t\mathbf{\Lambda}$ is the matrix obtained by multiplying every entry of the matrix $\mathbf{\Lambda}$ by t and exp **M** of a square matrix **M** is the infinite power series

$$\exp(\mathbf{M}) = 1 + \mathbf{M} + \frac{1}{2!}\mathbf{M}^2 + \frac{1}{3!}\mathbf{M}^3 + ... + \frac{1}{n!}\mathbf{M}^n + ...$$

The computational costs involved in calculating the above series is huge due to the large number of matrix exponents involved. In practice, one could use the Jordan decomposition of **M** to significantly reduce computational complexity.

One way to estimate the matrix $\mathbf{\Lambda}$ is to first obtain an estimator, $\hat{\mathbf{P}} = \hat{\mathbf{P}}(0,1)$ for a one-step transition matrix **P** using the cohort method. Under certain conditions[6], the following infinite series of matrix sums

$$-\left(\mathrm{I}-\hat{\mathbf{P}}\right)-\frac{1}{2}\left(\mathrm{I}-\hat{\mathbf{P}}\right)^2-\frac{1}{3}\left(\mathrm{I}-\hat{\mathbf{P}}\right)^3-\frac{1}{4}\left(\mathrm{I}-\hat{\mathbf{P}}\right)^4-...-\frac{1}{n}\left(\mathrm{I}-\hat{\mathbf{P}}\right)^n-...$$

is known to converge to a matrix $\hat{\mathbf{Q}}$ (see Singer and Spilerman (1976) for a proof) such that $\hat{\mathbf{P}} = \exp(\hat{\mathbf{Q}})$. However, the matrix $\hat{\mathbf{Q}}$ is not guaranteed to have non-negative

[6] $\hat{\mathbf{P}}$ should satisfy some sufficient conditions. The mathematical details can be found in Israel et al. (2001)

off-diagonal entries, in which case it would fail to be a generator matrix for the Markov chain. It is also possible that a true generator exists even though the above series does not converge. In general, this is the well-known *embeddability problem*.

One can avoid the above problem by adopting the approach of Lando and Skødeberg (2002) in which the transition intensities for the generator matrix are estimated directly via existing estimation techniques. Specifically, we obtain the following maximum likelihood estimator for λ_{ij} (see Kuchler and Sorensen, 1998)

$$\hat{\lambda}_{ij} = \frac{\mathrm{N_{ij}(T)}}{\int_0^T \mathrm{N}_i(s)\, ds}, \; i \neq j .$$

In the expression on the right-side of the above formula, $\mathrm{N}_i(s)$ is the number of funds with rating *i* at time *s*. The term in the denominator is a count of the total fund-month spent at state *i* over the period [0, *T*]. It accounts for any period of time over which a fund carries a certain rating. Obviously, this estimator is inappropriate when no information about the exact transition time is available. For mutual fund ratings such as the star rating of Morningstar, the transition time can be taken to be end of every month[7].

[7] Morningstar calculates ratings at the end of each month and publishes them at the beginning of the following month. See Morningstar (2009).

Chapter 3 Grades Matter in Performance: Morningstar Stewardship Grades and Mutual Fund Performance

3.1 Introduction

Mutual fund investors have the daunting task of choosing which funds to invest in from thousands of available funds. We can analyze their dilemma as two seemingly mutually confounding problems. First, to fathom the future performance of the funds based on current available evidence, and second, to assess how well the mutual fund managers steward their investments under uncertain economic conditions.

It has been well established that the relationship between past and future short-term performance (or "hot hands") of mutual fund managers are tenuous at best (Jensen, 1969, Grinblatt and Titman, 1992 , Hendricks, Patel and Zeckhauser, 1993, Goetzmann and Ibbotson, 1994, Brown and Goetzmann, 1995, Elton, Gruber and Blake, 1996, Carhart, 1997). In this paper, we explore what role Morningstar Stewardship Grades (a corporate governance score for mutual funds) play in mutual fund performance.

The Morningstar Star Ratings have been widely used by retail and institutional investors alike as tools for selecting mutual funds. A comprehensive study on the influence of the Star Ratings found significantly large inflows in response to rating upgrades or initiation of top rating (Del Guercio and Tkac, 2008). The enormous popularity of the Star Ratings has prompted other researchers to study

their effectiveness as a performance measure. One such strand of research focuses on gauging the predictive ability of these ratings. It was documented that poor 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[1] found that best-rated funds outperform lower-rated funds over a three-year post-rating period (Morey and Gottesman, 2006).

Adam Smith's observation in *the Wealth of Nations*, "…*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*" (Smith, 1776, 700) almost prophetically lends its voice to the world of mutual funds governance in the beginning of the millennium.

Unsurprisingly, 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).

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

Corporate Finance literature has long established the nexus between corporate governance and mutual funds as large shareholders. Past academic interests have reviewed on how regulations, through restrictions on ownership concentration and control, have restrained institutional investors like mutual funds from playing an influential role in the governance of corporations whose assets they own (Roe, 1990, Shleiffer and Vishny, 1997, p. 38; for a more current outlook, see Bebchuk and Weisbach, 2010). However, interest in the corporate governance of the mutual funds themselves is of more recent vintage. Although academic research on corporate governance of large investors like mutual funds is still burgeoning, scholarly work on mutual fund corporate governance scores such as the Morningstar Stewardship Grades is relatively scarce (Li, Moshirian, Pham and Zein, 2006, Chou Ng and Wang, 2011 and Chen and Huang, 2011). Hitherto, to the best of our knowledge, the work by Wellman and Zhou (2007) is probably the most comprehensive study on the Stewardship Grades.

Using the first release of the Stewardship Grades since August 2004, Wellman and Zhou (2007) 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. Furthermore, their study on fund flows pattern reveals that upgrades and downgrades of Stewardship Grades lead to positive and negative fund flows respectively. In addition, they find that among the five stewardship components, only Fees and Board Quality exhibit significant explanatory power, thus demonstrating an indirect relation between corporate governance and fund performance. Their work corroborates

academic studies that report a positive association between firm valuation and corporate governance scores like the widely used G-index (Gompers, Ishii and Metrick, 2003).

Gil-Bazo and Ruiz-Verdu (2009) find some evidence that better governance is associated with fees that are more aligned with fund performance, thus offering a partial explanation for the anomaly that funds with worse before-fee performance charge higher fees. In a similar spirit, 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. Their findings suggest that funds with good corporate governance, as indicated by their Stewardship Grade, tend to act in the interest of their shareholders by voting responsibly for governance issues of their portfolio firms and investing only in well-governed firms.

Chen and Huang (2011) employ both OLS regression and quantile regressions to examine the contemporaneous relation between fund performance and corporate governance using both the 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. However, quantile regressions demonstrate a strong positive relation between Manager Incentive and fund performance at the right tail of the performance distribution. In addition, Board Quality is found to have a

significant ability to predict future performance when quantile regression is used. Along the same vein but performing dummy variable OLS regressions, Gottesman and Morey (2012) find no evidence that any of the stewardship component can consistently predict future performance. Hence among published work, there is at best mixed evidence of the effectiveness corporate governance components in performance prediction.

There has been growing interest among academics in using Stewardship Grades as a proxy for corporate governance quality to examine the role governance plays in various dimensions of fund management. Among the working papers that involve the use of Stewardship Grade are Casavecchia and Tooman (2012) and Lai, Tiwari and Zhang (2010). The former investigate how governance is associated with managerial herding behavior. Their results indicate that a higher manager incentive grade is associated with a lower intensity of managerial herding activities. The latter document three key results related to board quality of mutual funds. First, they report that for funds in the bottom quintile based on past performance, those with a good Board Quality grade suffer significantly lower outflows. Second, they find that for funds with bad boards, a negative past performance is strongly predictive of future negative performance. Finally, they document that following poor performance, funds with better boards are more likely to change their fund strategy compared to funds with bad boards. More recently, Kurniawan, How and Verhoeven (2012) explore whether governance matters to fund style drift. Their analysis provides evidence that style-drift is negatively related to individual stewardship components such as Board Quality, Fees Structure and Regulatory History. Such ongoing interest and work help

establish the unmistakable link between governance components and better operation of the fund, which consequently leads to performance.

Evidently, 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. Our findings complement existing studies on the relation between mutual fund performance and the performance-based Morningstar Star Ratings, thus providing some insights on the extent to which corporate governance of mutual fund should be considered by investors in searching for the best performing mutual funds. Mutual fund ratings have been widely publicized to, and often used by retail mutual fund distributors as a marketing tool for selling mutual funds. Individual investors also use ratings as a primary criterion for screening mutual funds. The results of our studies have important ramifications for retail investors' financial well-being.

We follow up with a brief preview of our contribution. 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)[2]. 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 *endogeneity bias* due to unobserved fund characteristics like managerial ability. Third, with the use of a dynamic panel data model, we demonstrate that even in the presence of lagged performance measures, a strong relationship between Stewardship Grades and performance holds. Finally, 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] The first principal component (FPC) is FPC = 0.46Board Quality+0.64Corporate Culture+0.30Fee Score+0.11Manager incentive+0.52Regulatory History. The details are presented in Appendix B.

3.2 Data

Morningstar provides monthly Star Ratings, including the 3-year, 5-year and 10-year ratings (whichever available[3]) and Stewardship Grades. We further obtain all Stewardship Grade components (Corporate Culture (CC), Board Quality (BQ), Fees Score (FS), Manager Incentive (MI) and Regulatory History (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 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 ratio and the date of inception. We include only funds whose fund identifiers from Morningstar (identifier = 'Ticker') and CRSP (Fund Identifier = 'Nasdaq') databases match.

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[4]. We observe that only a small percentage of funds receive the best and worst mutual find ratings

[3] 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.

[4] 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 (Lutton et. al., 2011, p. 4-5)

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

Table 2 reports the descriptive statistics of the fund variables that we shall 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.

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 the 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 measures. 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 Rating. For the Stewardship Grade, the raw score is simply the arithmetic average of the cores 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 suggests there is a long term feedback relationship between the two variables. However, as Stewardship Grades are quite persistent (possibly non-stationary) while Star Ratings are not (i.e., stationary), 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 (Ghosh and Ng, 2013) and subsequent sections of this paper, we shall 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).

3.3 Methodology for Ranked Portfolio Tests and Regressions

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). Given the difficulty with which to extract the true components of

stewardship protection, we shall employ the principal component analysis to seek an objective or data-driven corporate governance score.

Finally, both the G-Index and the subset of drivers for managerial entrenchment (or the E-Index) are dependent on the questions, and are constructed previously with equal weightage on the constituent questions (Gompers et al., 2003, Bebchuk et al., 2008). However, as has been the case, the effectiveness of these questions to elicit corporate governance practices has been varying with time. To elucidate this problem, Bebchuk et al. (2008) observes:

"…institutional investors deciding which firms to include in their portfolios and which governance changes to press for would likely be better served if shareholder advisory firms were to use governance measures based on a small number of key provisions rather than attempt to count all the trees in the governance forest."

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):

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

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 period. Specifically, for each month over the period November 2004 (month 1) to May 2010 (month 67), we rank the sample funds by one or both of the Morningstar Stewardship Grades (abbrev. SG) and/or Morningstar Star Ratings (abbrev SR) and compute the difference in mean four-factor alpha for funds in any two groups. We then perform a t-test with Newey-West standard errors 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 funds 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 1 and 2, 2 and 3 and 1 and 3. We use the symbols 2_1, 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, We are thus 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 non-stationary process.[5]

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 is made mandatory.
3. Regulatory history score is changed from a scale of 0- 2 to -2 to 0.

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

[5] 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).

also observe a strong persistence of Stewardship Grades vis-à-vis the performance measures.

3.4 Empirical Results

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. However, 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 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. Nonetheless, 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 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 the top and bottom ranked portfolios (Table 4 Panel A and B, column SR*FPC). In summary, using Carhart (1997) one-year four factor alpha, we find little evidence (statistical or economic) that either the stewardship grade or the data-driven principal component score can predict short-term future performance without controlling for other factors.

The results are fundamentally different when we use a different more long-term measure of fund performance like the raw star rating score. We find that the top ranked portfolio outperforms the middle and bottom portfolios in a statistically and economically significant manner. For the period of November 2004 to May 2010, the top portfolio ranked by previous-month Stewardship grade outperformed the middle and bottom portfolio by 0.45 percent and 0.75 percent respectively. The corresponding figures when ranked by the objective data-driven FPC score are 0.3 percent and 0.58 percent respectively.

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 bottom portfolios ranked by their corresponding ratings. However, a stronger and a more consistent result is obtained when we consider a long term risk-adjusted return like the star rating.

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 observe 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 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. We first estimate the standard fixed effect model for the yearly data (collected in December, 2005-2011) assuming *strict exogeneity* of the regressors in Specification EX (Table 5, Panel A, Model Spec EX) (Wooldridge, 2010). In model (SR_EX), we find that the previous period's Star Rating has a positive and significant impact on the four-factor alpha. We further observe that previous period's size plays a significant negative role while age of the fund plays a significant but positive role in determining the risk adjusted returns. While the negative effect of size is consistent with the story that bigger funds can water down returns, the positive effect of age in this specification seems to be counter intuitive with extant literature (Berk and Green, 2004).

Furthermore, for the fixed effect (FE) regression models with just the Stewardship Grade (SG_EX) and the First Principal Component (FPC_EX), neither

the Stewardship Grade (SG) nor the First Principal Component (FPC) have a significant impact on short-term performance. In addition, in both models (SG_EX) and (FPC_EX) the effect of previous period's size becomes statistically insignificant in determining short-term performance as opposed to the specification (SR_EX). Age, on the other hand, shows a negative but statistically insignificant association with the short term risk adjusted performance. Such results might have led to a popular perception that corporate governance scores do not play a strong enough role in determining fund performance. We however think otherwise. The insignificance of the SG and FPC scores in the models (SG_EX) and (FPC_EX) might be attributed to the violation of the strict exogeneity assumption in these specifications that make the coefficient estimates inconsistent.

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 include 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 U.S. 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 analysed.

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. Both under strict exogeneity (Model EX) and incorporating endoeneity (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.

Predictive Panel Regressions of Long term Performance and Stewardship

The main objective of pursuing good governance is to ensure a long-term and 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, Del Guercio and Tkac, 2008).

Hence, 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. Table 6 Panel A, specification EX uses the assumption of strict exogeneity in the Fixed Effect panel data model. We observe that in Model (SGA_EX) all the components of the Stewardship score except lagged values of Corporate Culture (CC) are statistically significant, and so are the

controls with lagged values of average manager tenure and longest manager tenure (measures of stability), size and age (measures of maturity) and absolute flow. We do, however, find that while size has a positive impact on Star Rating, age shows a significant negative impact. This dichotomy is persistent for models (SG_EX) and (FPC_EX) where we use the Stewardship Grade and the FPC score respectively consistent with findings in Bebchuk et. al. (2009).

In Table 6 Panel B models of specification EX for data on or after 2007, we also find a positive effect of average manager tenure being higher, and a slight positive effect (in Model SGA_EX) of the longest manager tenure. This last result could be real (say, previously discussed managerial entrenchment, as documented in Brown et al., 1996, Ding and Wermers, 2009) or an artifact of the possible endogeneity in the model.

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 ratio 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 that past Star Rating plays 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 EX and 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.

In the standard time fixed effects model with strict exogeneity reported in Table 6 Panel C Specification EX, we find that all the Stewardship components are significant with Fee Score having a negative coefficient. The other variables that exhibit significant positive effect include lagged values of turnover, size, manager tenure average and absolute flow. We also observe that the raw Stewardship and FPC scores have a significant positive impact on Star Rating when controlled for fund characteristics, of which only expense ratio shows a significantly negative impact.

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. These results corroborate findings on effect on the E-index (Bebchuk et al., 2009).

One exciting part about the monthly data is that the number of observations on time domain is sufficiently large to facilitate 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 turn 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 Models FPC_EX and SG_EX).

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 ratio 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, 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.

3.5 Conclusions

According to the 2012 Investment Company Fact Book, 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 leads us 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.

Stewardship Grades have been given by Morningstar since November 2004 to provide investor with an indication of how well a mutual fund performs its fiduciary duties. The evidence of the link between good corporate governance and performance have been ephemeral at best. In this paper, we evaluate and ascertain this linkage for mutual funds, and hence, give a simple alternative to the recombination of individual

Stewardship scores to an objective measure that could have implications on long-term performance.

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.

Finally, we explore the implications for investors, both institutional and retail, in evaluating mutual funds with other factors like size, age, manager tenure, flow, expense ratio and turnover.

Summarizing, we find some evidence that a new fiduciary grade based on principal component analysis of the component grades possesses a stronger predictive

power than the Stewardship Grade itself. This suggests that a more informative and reliable corporate governance rating can be obtained by putting weightings on the individual stewardship components based on Principal Components Analysis. In our study on the relation between the two ratings, we find that strong Granger causality exists between the two ratings. This relation holds even when we control for all potential determinants. To avoid spurious regression, we perform unit root tests on the time series of both ratings to ascertain that not both series are non-stationary. It turns out that the Star Rating is a stationary process while the Stewardship Grade cannot be proven to be stationary. However, the difference series do not Granger cause each other.

With an increasing awareness of the importance of corporate governance, investors are likely to include governance quality as one of their criteria for screening mutual funds. Given the popularity of the Morningstar Star Ratings, the Stewardship score (or better still, the proposed FPC score) has the potential of becoming a standard 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. Moreover, our findings complement existing work on the predictive ability of mutual fund ratings and persistence of mutual fund performance. We acknowledge that our analyses are subject to some limitations such as the use of raw scores as continuous variables and the short duration of our data set, although this is the longest one analyzed in the literature. Nonetheless, we hope our application of appropriate econometric techniques can help to instigate further research on the efficacy of

mutual fund ratings by using more robust methods that can better handle panel data involving ordinal variables such as the Stewardship scores.

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, 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)

Rating	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 t-test with Newey-West adjusted standard errors 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	0.017	0.0211	0.4509***	0.3119***
	(0.0131)	(0.0157)	(0.0496)	(0.0179)	(0.0289)	(0.0357)	(0.0254)
3_1	0.0946***	-0.0094	-0.0445	-0.0077	0.0509	0.7490***	0.5833***
	(0.0239)	(0.0305)	(0.0473)	(0.0154)	(0.0366)	(0.0366)	(0.0385)

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	0.036	0.0587*	0.4103***	0.2917***
	(0.0162)	(0.0202)	(0.0458)	(0.0256)	(0.0316)	(0.0452)	(0.0340)
3_1	0.1313***	0.0089	0.0216	-0.0053	0.0677	0.7140***	0.4875***
	(0.0241)	(0.0455)	(0.0435)	(0.0246)	(0.0460)	(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 EX) A static panel fixed time effect model.

(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)

Dependent variable: One-year four-factor risk-adjusted alpha (Dec 2005, Dec 2006, 2007, Dec 2008, Dec 2009 and Dec 2010)

	(Model Spec EX)			(Model Spec EN)			(Model Spec DYN)			
Explanatory Variables	(SR_EX)	(SG_EX)	(FPC_EX)	(SR_EN)	(SG_EN)	(FPC_EN)	(SGA_DY)	(SR_DY)	(SG_DY)	(FPC_DY)
intercept	-0.2326***	-0.0041	0.0250	-0.1418**	-0.5565**	-0.5306**				
	(0.0851)	(0.0835)	(0.0774)	(0.0983)	(0.2220)	(0.2237)				
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	0.0415***			0.0617				0.0511*		
	(0.0127)			(0.0402)				(0.0880)		
lagged SG raw score		0.0056			0.1134**				0.4091	
		(0.0122)			(0.0489)				(0.3401)	
lagged FPC			0.0016			0.2546**				0.2202
			(0.0205)			(0.1149)				(1.2901)
lagged FF alpha							-0.3509***	-0.7581*	-0.8813**	-0.7588**
							(0.0957)	(0.2418)	(0.3573)	(0.3504)
lagged turnover ratio	-0.0011	-0.0020	-0.0020	-0.0011	-0.0011	-0.0017	0.0176	-0.0004	-0.0036	-0.0021
	(0.0010)	(0.0015)	(0.0015)	(0.0009)	(0.0007)	(0.0012)	(0.0139)	(0.0007)	(0.0023)	(0.0043)
lagged size	-0.0165***	-0.0036	-0.0018	-0.0147**	-0.0388**	-0.0360**	-0.5873***	-0.3291***	-0.2007	-0.2710
	(0.0047)	(0.0098)	(0.0091)	(0.0068)	(0.0193)	(0.0187)	(0.1547)	(0.0651)	(0.2049)	(0.2170)
lagged age	0.0322**	-0.0073	-0.0086							
	(0.0126)	(0.0166)	(0.0167)							
Fixed Time Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Fund Effect	No	No	No	No	No	No	No	No	No	No
Fixed Fund Effect (Difference)	No	No	No	No	No	No	Yes	Yes	Yes	Yes
N	626	532	532	626	531	531	500	500	500	500

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

Dependent variable: One-year four-factor risk-adjusted alpha (Dec 2007, Dec 2008, Dec 2009 and Dec 2010)						
	(Specification EX)			(Specification EN)		
	(SR_EX)	(SG_EX)	(FPC_EX)	(SR_EN)	(SG_EN)	(FPC_EN)
Intercept	-0.3388***	-0.0752	-0.0434	-0.1675	-0.6342*	-0.6421*
	(0.0792)	(0.0980)	(0.0921)	(0.1351)	(0.3237)	(0.3388)
lagged BQ						
lagged CC						
lagged FS						
lagged MI						
lagged RH						
Lagged SR raw score	0.0517***			0.0628		
	(0.0142)			(0.0589)		
lagged SG raw score		0.0127			0.1355*	
		(0.0168)			(0.0746)	
lagged FPC			0.0149			0.3207*
			(0.0282)			(0.1149)
lagged turnover ratio	-0.0009	-0.0013	-0.0014	-0.0008	-0.0010	-0.0016
	(0.0010)	(0.0016)	(0.0016)	(0.0009)	(0.0011)	(0.0015)
lagged size	-0.0160**	-0.0064	-0.0039	-0.0117	-0.0497	-0.0469
	(0.0063)	(0.00139)	(0.0122)	(0.0098)	(0.0304)	(0.0302)
lagged age	0.0447***	5.54E-5	-0.0006			
	(0.0122)	(0.0232)	(0.0234)			
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	No	No	No
N	626	526	526	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 EX)	A static panel fixed time effect model.
(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 EX) A static panel fixed time effect model.

(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. 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

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

Dependent variable: star rating Raw Scores (SRRaw) : Dec 2005, Dec 2006, 2007, Dec 2008, Dec 2009 and Dec 2010

	(Specification EX)			(Specification EN)			(Specification DY)		
	(SGA_EX)	(SG_EX)	(FPC_EX)	(SGA_EN)	(SG_EN)	(FPC_EN)	(SGA_DY)	(SG_DY)	(FPC_DY)
intercept	2.6930***	3.1035***	3.0787***	-2.3316	-1.4026***	-1.3375***			
	(0.1600)	(0.1185)	(0.1845)	(9.4294)	(0.1599)	(0.1761)			
lagged BQ	0.294***			-6.1152			1.5233		
	(0.0050)			(8.5596)			(1.5346)		
lagged CC	0.0382			-3.1755			-0.5023		
	(0.0556)			(7.9687)			(0.9719)		
lagged FS	-0.0925***			0.2201			1.6888		
	(0.0354)			(2.8163)			(1.3333)		
lagged MI	0.0729***			1.2526			-0.1882		
	(0.0192)			(2.0506)			(0.8381)		
lagged RH	0.1851***			8,9256			0.8795		
	(0.0371)			(13.6769)			(0.6528)		
lagged SG raw score		0.0705***			0.7440***			0.7020**	
		(0.0196)			(0.0489)			(0.3014)	
lagged FPC			0.1783***			1.7089***			1.7719**
			(0.0312)			(0.1111)			(0.5547)
lagged SR raw score							0.1980***	0.4947***	0.4328***
							(0.0715)	(0.0544)	(0.0524)
lagged Turnover ratio	0.0039	0.0049	0.0041	0.0109	0.0087**	0.0039	-0.0055	-0.0024	-0.0052*
	(0.0028)	(0.0030)	(0.0029)	(0.0354)	(0.00394)	(0.0043)	(0.00522)	(0.0025)	(0.0028)
lagged size	0.1497***	0.1405***	0.1339***	0.2388	-0.0913***	-0.0786***	-0.1940**	-0.2969***	-0.2525***
	(0.0178)	(0.0193)	(0.0139)	(0.1843)	(0.0332)	(0.0257)	(0.0859)	(0.0654)	(0.0614)
lagged avg manager tenure	0.0112***	0.0141***	0.0101						
	(0.0019)	(0.0025)	(0.0066)						
lagged longest manager tenure	0.0143***	0.0153***	0.0183***						
	(0.0028)	(0.0037)	(0.0046)						
lagged absolute flows	0.0005**	0.0003**	0.0005						
	(0.0002)	(0.0001)	(9.50E-05)						
lagged age	-0.2834***	-0.3423***	-0.3096***						
	(0.04005)	(0.0419)	(0.0327)						
Fixed Time Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Fund Effect	No	No	No	No	No	No	No	No	No
Fixed Fund Effect (Difference)	No	No	No	No	No	No	Yes	Yes	Yes
N	536	536	536	536	536	536	508	508	508

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

Dependent variable: star rating Raw Scores (SRRaw). : Dec 2007, Dec 2008, Dec 2009 and Dec 2010									
	(Specification EX)			(Specification EN)			(Specification DY)		
	(SGA_EX)	(SG_EX)	(FPC_EX)	(SGA_EN)	(SG_EN)	(FPC_EN)	(SGA_DY)	(SG_DY)	(FPC_DY)
intercept	2.4896***	3.0297***	2.9591***	1.1733	-1.4776***	-1.4497***			
	(0.2765)	(0.2145)	(0.2437)	(3.8057)	(0.2432)	(0.1120)			
lagged BQ	0.2464***			-2.8588			3.5513*		
	(0.0466)			(4.6754)			(1.9048)		
lagged CC	0.0038			0.1352			-2.1204		
	(0.0590)			(2.4047)			(1.7410)		
lagged FS	-0.1488***			-0.5677			2.3248		
	(0.0217)			(0.6533)			(2.5091)		
lagged MI	0.0721**			0.6922			2.2919		
	(0.0386)			(0.5085)			(1.1419)		
lagged RH	0.1891***			2.6490			1.6128		
	(0.0641)			(5.1095)			(1.7169)		
lagged SG raw score		0.0322***			0.7894***			1.2171***	
		(0.0040)			(0.0717)			(0.4386)	
lagged FPC			0.1024**			1.8177***			2.9350***
			(0.0431)			(0.1045)			(0.8201)
lagged SR raw score							0.1574	0.4355***	0.3196***
							(0.1062)	(0.0763)	(0.0680)
lagged turnover ratio	0.0030	0.0041	0.0034	0.0076	0.0032	-0.0017	-0.0082	-0.0057	-0.0105**
	(0.0376)	(0.0040)	(0.0034)	(0.0225)	(0.0028)	(0.0044)	(0.0071)	(0.0036)	(0.0046)
lagged size	0.1445***	0.1340***	0.1268***	0.1655	-0.1388***	-0.1120***	-0.2375	-0.3617***	-0.2913***
	(0.0236)	(0.0234)	(0.0175)	(0.1517)	(0.0382)	(0.0250)	(0.1544)	(0.0919)	(0.0824)
lagged avg manager tenure	0.0137***	0.0114**	0.0097						
	(0.0038)	(0.0041)	(0.0088)						
lagged longest manager tenure	0.0142**	0.0185***	0.0212***						
	(0.0062)	(0.0065)	(0.0061)						
Absolute flows	0.0004***	0.0004**	0.0004***						
	(0.0001)	(0.0001)	(0.0001)						
Age	-0.2192***	-0.2553***	-0.2477***						
	(0.0577)	(0.0561)	(0.0436)						
Fixed Time Effect	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Fixed Fund Effect	No	No	No	No	Yes	Yes	No	No	No
Fixed Fund Effect (Difference)	No	No	No	No	No	No	Yes	Yes	Yes
N	527	527	527	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_EX)	(FPC_EX)	(SG_EX)	(SGA_EN)	(FPC_EN)	(SG_EN)	(SGA_DY)	(FPC_DY)	(SG_DY)
intercept	1.2204***	1.4737***	1.5023***	1.2182***	-1.8331***	-1.6111***			
	(0.0233)	(0.0313)	(0.0278)	(0.0174)	(0.0848)	(0.0480)			
lagged BQ	0.2801***			0.3079***			-0.0129***		
	(0.0149)			(0.0157)			(0.0001)		
lagged CC	0.1382***			0.1799***			-0.0287***		
	(0.0165)			(0.0160)			(0.0008)		
lagged FS	-0.1228***			-0.1336***			-0.0593***		
	(0.0081)			(0.0080)			(0.0035)		
lagged MI	0.1238***			0.1687***			0.0189***		
	(0.0120)			(0.0120)			(0.0003)		
lagged RH	0.1648***			0.1815***			0.0911***		
	(0.0107)			(0.0115)			(0.0008)		
lagged SG raw score			0.0895***		1.8445***	0.7531***			0.0013***
			(0.0058)		(0.0422)	(0.0137)			(1.32E-06)
lagged FPC		0.2197***						0.0016***	
		(0.0135)						(0.0002)	
lagged SR raw							0.6930***	0.6929***	0.6883***
							(0.0001)	(3.49E-05)	(8.69E-06)
Lagged turn ratio	0.0054***	0.0053***	0.0057***	0.0041***	0.0043***	0.0088***	0.0005***	0.0007***	0.0005***
	(0.0005)	(0.0006)	(0.0005)	(0.0006)	(0.0009)	(0.0009)	(4.41E-05)	(4.77E-06)	(1.45E-06)
lagged size	0.1390***	0.1223***	0.1417***	0.1405***	-0.0680***	-0.0729***	0.0552**	0.0296***	0.0587***
	(0.0045)	(0.0044)	(0.0024)	(0.0038)	(0.0075)	(0.0077)	(0.0003)	(0.0001)	(1.52E-06)
lagged exp.ratio	-0.0034***	-0.0039***	-0.0038***						
	(0.0008)	(0.0010)	(0.0009)						
lagged avg manager tenure	0.0232***	0.0276***	0.0275***						
	(0.0008)	(0.0009)	(0.0009)						
lagged absolute flows	0.0006***	0.0006***	0.0006***						
	(0.0001)	(0.0001)	(0.0001)						
Lagged age							-0.5264***	-0.4115***	-0.4293***
							(0.0017)	(0.0008)	(1.18E-05)
D_crisis								-0.0520***	
								(0.0004)	

	(SGA_EX)	(FPC_EX)	(SG_EX)	(SGA_EN)	(FPC_EN)	(SG_EN)	(SGA_DY)	(FPC_DY)	(SG_DY)
Fixed Time Effect	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Fixed Fund Effect	No	No	No	No	No	No	No	No	No
Fixed Fund Effect (Difference)	No	No	No	No	No	No	Yes	No	No
N	623	623	623	625	625	625	588	588	588

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

Dependent variable: star rating Raw Scores (SRRaw) Sample : January 2007 to May 2011

	Specification EX			Specification EN		
	(SGA_EX)	(FPC_EX)	(SG_EX)	(SGA_EN)	(FPC_EN)	(SG_EN)
intercept	1.2567***	1.5825***	1.57911***	1.2202***	-2.0030****	-1.5869***
	(0.0437)	(0.0450)	(0.0349)	(0.0330)	(0.0910)	(0.0697)
lagged BQ	0.2203***			0.2408***		
	(0.0205)			(0.0203)		
lagged CC	0.0888***			0.1391***		
	(0.0178)			(0.0181)		
lagged FS	-0.1642***			-0.1750***		
	(0.0049)			(0.0049)		
lagged MI	0.1643***			0.2125***		
	(0.0152)			(0.0146)		
lagged RH	0. 1566***			0.1793***		
	(0.0184)			(0.0188)		
lagged SG raw score			0.05092***			0.7654***
			(0.0030)			(0.0179)
lagged FPC		0.1201***			1.9533****	
		(0.0113)			(0.0458)	
lagged turnover ratio	0.0056***	0.0058***	0.0060***	0.0038***	0.0012	0.0060***
	(0.0006)	(0.0007)	(0.0007)	(0.0007)	(0.0008)	(0.0008)
lagged size	0.1503***	0.1383***	0.1366***	0.1550***	-0.0940****	-0.1004***
	(0.0029)	(0.0039)	(0.0043)	(0.0029)	(0.0089)	(0.0090)
lagged expense ratio	-0.0037***	-0.0046***	-0.0045***			
	(0.0008)	(0.0009)	(0.0009)			
lagged avg manager tenure	0.0254***	0.0311***	0.0309***			
	(0.0012)	(0.0013)	(0.0012)			
lagged absolute flows	0.0005***	0.0005***	0.0005***			
	(0.0001)	(0.0001)	(0.0001)			
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	No	No	No
N	606	606	606	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					0.0265****		0.0702***
					(6.92E-05)		(1.84E-05)
lagged SR raw score	0.7433***	0.7459***	0.7460***	0.7457***	0.7441****	0.7463***	0.6920***
	(7.60E-06)	(0.0001)	(0.0042)	(2.03E-05)	(1.02E-05)	(6.90E-05)	(5.51E-06)
lagged SR raw score (lag 2)							0.1430***
							(4.58E-06)
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			-0.0757****				
			(0.0187)				
Lagged First Difference of SG raw score				-0.0271***			
				(2.46E-06)			
lagged turnover ratio	0.0006***	0.0003***	-0.0002	0.0003***	0.0008****	0.0006***	0.0005***
	(1.29E-07)	(6.50E-06)	(0.0012)	(2.47E-07)	(1.60E-06)	(3.61E-07)	(5.13E-07)
lagged size	0.0384***	0.0402***	-0.0121	0.0406***	0.0087****	0.0430***	-0.0576**
	(0.0001)	(0.0002)	(0.0193)	(5.47E-06)	(5.17E-05)	(1.04E-05)	(1.01E-05)
lagged age	-0.5558***	-0.3897***	-0.4238***	-0.3911***	-0.4566****	-0.4412***	-0.2976***
	(0.0015)	(0.0006)	(0.0407)	(8.51E-05)	(0.0002)	(2.40E-05)	(0.0001)
D_crisis(-1)			-0.0570*		-0.0500****		-0.0486***
			(0.0314)		(6.03E-05)		(1.02E-05)

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 A
Morningstar Stewardship Grades Methodologies

Five Stewardship Components

The Morningstar Stewardship Grade is calculated based on five components:

(i) Board Quality (BQ),

(ii) Corporate Culture (CC),

(iii) Fees (Fees),

(iv) Manager Incentives (MI)

(v) Regulatory History (RH).

We first describe in detail the above five Stewardship components. What we present here are based on methodologies used prior to a revamp in 2007. We articulate the details of the methodology changes at the end of this Appendix.

Board Quality

The board quality score is determined from the following four factors, each worth up to 0.5 point:

(i) Does the board act consistently to protect the interests of shareholders.

Examples of positive action taken by the board include dismissing or replacing underperforming fund managers and disapproving attempt by fund management to merge poor-performing funds with more successful funds. Ding and Wermers (2005) document evidence supporting the hypothesis that the replacement of fund managers is beneficial to shareholders. It is found that on the average, fund managers who are replaced by board directors underperform their peers, and that incoming managers outperform those replaced by one percentage point per year.

(ii) Do independent directors have significant investments in the funds? (Maximum score = 0.5 point)

The highest score of 0.5 point can be earned if at least 75% of a board's directors invest in the funds they oversee with an amount exceeding his/her aggregate annual compensation for serving on a board.

Chen, Goldstein and Jiang (2006) and Cremers et al. (2005) have independently examined director ownership in funds. The former report that the optimal contracting hypothesis holds: in the absence of other control mechanisms, directors tend to own shares in the funds that offer high expected benefits. Indeed, directors are found to hold more shares in actively managed funds such as small-cap equity funds than in, for example, bond funds. The latter find that there exists a strong positive association between fund performance and directors' stakes in the funds.

(iii) Is a board overseeing too many funds to the extent that its ability to protect shareholders' interest will be compromised. [6]

Ferris and Yan (2007) show that directors who oversee many funds have a higher chance of being implicated in a fund scandal. Their results support Morningstar's view that board's effectiveness would be adversely affected by "over-burdened" directors.

(iv) Does the fund meet the requirement of the Security and Exchange Commission (SEC) that at least 75% of the board's directors are independent?

Morningstar does not consider current and former employees as well as family members of both fund company and fund services providers as independent. The belief that board independence is positive for fund performance is affirmed by

[6] This criterion was dropped after the methodology change in 2007.

Khorana (1996) who demonstrates that the degree of independence of a fund's board has a positive association with the quality of fund governance. It is found that underperforming managers are more likely to be replaced when the board has a higher proportion of independent directors.

Corporate Culture

For this component, Morningstar considers a wide spectrum of factors.

(i) Has the fund management company launched "trendy" funds just to chalk up assets, regardless of whether the timing to launch such funds is appropriate. For example, many funds that are narrowly-focused on technology stocks were launched during the time when the technology sector was at the verge of collapse. Indeed, the bubble burst in 2001, causing many investors to suffer heavy losses.

(ii) Has the fund management company closed funds at an appropriate asset size or has it allowed the size to reach an unacceptable level. This question is crucial because it is difficult to manage a fund with a huge asset base effectively and profitably. Fund managers might be forced, due to liquidity and other consideration, to take large positions in stocks which might not offer the best potential returns.

As part of fund managers' compensation is derived from management fees which are in turn a fraction of the asset size of the fund, unscrupulous fund managers might want to continue growing the size of their fund's fund asset base to reap higher monetary gains.

Chen, Hong, Huang and Kubik (2004) find strong evidence that fund size erodes fund performance. Furthermore, the adverse effect that fund size has on fund returns is most pronounced for funds that invest in illiquid assets such as stocks having small capitalization, thus suggesting that liquidity concerns could in part explain this effect.

(iii) Does the fund implement measures such as high back-end loads to discourage frequent redemption of funds.

(iv) Does the firm communicate effectively with shareholders. For example, the management are expected produce comprehensive publications such as updated fund fact sheets and portfolio managers' reports for all shareholders on a regular basis.

(v) Has the firm used soft dollars which are payments made to the fund service providers. Soft dollars which are incorporated into brokerage fees will neither be reported nor included in the calculation of fund's expense ratio. Control on the use of soft dollars benefits shareholders. Funds paying high soft dollar commissions will be penalised.

(vi) Does the firm have a high manager retention rate? It is believed a high retention rate is associated with a conducive working culture where employees are happy and perform well. Investors who buy funds from a firm with a high manager retention rate will be more confident that the manager does not change over a long term. Changing of fund manager can be disruptive to investors' investment plan as they have to re-assess the new manager and his/her investment strategy.

Fees

Mutual fund investors pay various levels of fees. Fees related to distribution and redemption, commonly known as front-end loads, or sales charges, are paid at the time of transaction. Back-end loads or redemption fees are paid when investors sell the fund. Management fees are paid on a regular (usually annual) basis via direct deduction from the funds' assets. Funds can come in various share classes. Although all classes hold the same securities and are managed by the same portfolio manager, they have different fees structure. In addition, as mentioned in the preceding

paragraph, fund sponsors make soft dollar arrangement in which fund managers pay higher brokerage commissions to research and brokerage services, incurring another layer of fees for shareholders. Morningstar evaluates funds based on two aspects of the fees structure, each worth 1 point.

(i) One aspect is Fees Comparison. A fund receives 0.5 points if its expense ratio is lower than the median expense ratio of all funds within the same category group and having the same share class[3]. An additional 0.5 points is awarded to funds having an expense ratio within the lowest 25th percentile.

(ii) The other aspect is Fees Trends. A fund receives 1 point if its expense ratio decreases as its assets grows, or if there is evidence that it will lower the expense ratio when its size increases. Funds that charge additional fees such as performance fees, typically a fraction of excess returns over a certain benchmark, will be viewed less favorably by Morningstar and hence tend to receive lower score.

Manager Incentives

Two aspects, each worth 1 point, will be evaluated:

(i) The first aspect is Fund Ownership. Does a portfolio manager invest a significant amount of money in the fund he oversees?

Managers with more than US 1 million or more than a third of their liquid net worth invested in their funds will be given 1 point. For investment of $500,000 - $1 million, managers receive 0.5 points. In cases where the fund size is small, fund managers can invest in other funds of the same firm to earn partial credit. As of 2005, fund managers of US mutual funds are required to disclose the amount of their wealth invested in the fund they manage, in the following seven ranges:

(i) $0,

(ii) $1-10,000,

(iii) $10,001-50,000,

(iv) $50,001-100,000,

(v) $100,001-500,000,

(vi) $500,000-1,000,000

(vii) above $1,000,000.

The above disclosure requirement is one of the series of new regulations enacted by the Security Exchange and Commissions in 2004 in response to fund scandals discovered then.

Fund ownership, according to a recent work by Khorana et al. (2007), is positively correlated to the risk-adjusted returns of funds, with fund performance improving by as much as three basis points for each basis point increase in managerial ownership. The results of their work support the notion that managerial ownership gives managers more incentives to generate higher returns for fund's shareholders, and is an important determinant of fund performance. Hence, disclosure on the level of managerial ownership offers investors valuable information which they can use when making their investment decision.

(ii) Does the compensation scheme reward portfolio manager based on long-term performance or short-term asset appreciation. Funds with incentives geared towards short-term growth will be viewed less favorably, and hence given lower score. Conversely, funds whose managers are compensated based on long-term fund performance instead of asset growth will generally receive higher rating.

Morningstar instructs fund companies to complete a survey which details the compensation structure of their fund managers as well the level of their investment in the funds they manage. Morningstar believes that fund managers' incentives have a strong influence on the quality of management. A fund manager whose compensation is tied to short-term out-performance of its benchmark (e.g. performance fees) will

have a tendency to take excessive risk, as documented in Brown, Harlow and Starks (1996**)**

Regulatory History

At the point of assessment, Morningstar examines regulatory History at the fund management level over its past three years of history. Funds found with severe breaches of certain regulations might get the lowest score of -2. Funds free from regulatory violations or potential fund indictments receive the highest score of 2. Funds found to have breached certain rules but have remedial actions in place will get a score between -2 and 2, depending on their level of commitment to reform.

Derivation of Stewardship Grades

To determine the final stewardship grade, Morningstar computes a score for each of these criteria. Prior to 2007, each criterion carried a maximum score of 2 points. For Regulatory History, the lowest possible score was -2. For each of the other four criteria, the minimum score was 0. Based on these scores, a qualitative grade published in Morningstar Fund Reports would be assigned according to Table 1 below:

Table A1. Qualitative Grade For Stewardship components (Prior to July 2007)

Score	Qualitative Grade
2.0	Excellent
1.5	Good
1	Fair
0.5	Poor
<= 0	Very Poor

The sum of the scores assigned to these five criteria was used to determine the overall stewardship grade as outlined in Table 2.

Table A2. Qualitative Grade For Overall Stewardship Grade (Before 2007)

Score	Stewardship Score
9 – 10	A
7 - 8.5	B
5 - 6.5	C
3 - 4.5	D
<= 2.5	F

Table A3. Qualitative Grade For Stewardship components (In and after July 2007)

Score	Letter Grade
Full credit	Excellent
¾ credit	Good
½ credit	Fair
¼ credit	Poor
No credit	Very Poor

Most of the data that the Morningstar analysts use are obtained from funds' *Statement of Additional Information*, a regulatory filing made annually with the Securities and Exchange Commission. Although Morningstar aims to update funds' Stewardship Grades annually, there could be delays due to reasons such as a change in control at a fund company, a change in coverage at Morningstar or scheduling difficulties.

Methodology Changes

Since the first release of grades in August 2004, Morningstar has revamped the Stewardship Grade methodology twice, the first in 2006 and the second in 2007. Prior to 2006, Morningstar compared the fees of a fund with other funds in the same category. Since 2006, comparison had been made among funds that invest in similar types of securities or adopt similar investment strategies. The main changes made in 2007 concerned the weightage of the five stewardship grade components. Specifically, in 2007, Morningstar increased the weighting of corporate culture from 2 of 10 points to 4 of 10 points. In addition, there were also some changes made to the way fund board quality, fee, and regulatory history are assessed. Prior to 2007, one factor that was considered in assessing board quality is the number of funds directors oversee. Under the new methodology, this workload-related criterion was replaced with an emphasis on the performance of the directors regardless of the number of funds they oversee. The three criteria used in assessing the fund boards are (1) Independence of the boards (25% of board quality score) (2) Director ownership in the shares of funds they oversee (25% of board quality score) (3) Director's stewardship of the funds they oversee[7] (50% of board quality score)

Finally, changes were also made to the way Morningstar analysts work. Prior to 2007, funds from the same family are covered by a team of analysts. Since 2007, Stewardship Grades to funds in a family are given by one lead analyst for that family. The lead analyst typically makes in-person due-diligence visit to the fund company to gather the data necessary to evaluate each fund, and proposes a set of scores to a committee of analysts who oversee the methodology.

[7] Board obtains high scores if (i) the fees charged by the funds that the board oversees are lower than the peer-group norm (2) board is prepared to close funds that could face constraints in investment strategies (e.g. funds with a huge fund size) (3) board goes far enough to fire poor-performing managers

Appendix B

Principal Component of Stewardship Component Scores

The FPC (First Principal Component) score that we use in this paper is calculated from a yearly sample of funds (December 2004 – December 2010. Specially, we calculate the first principal component (FPC) based on each of the following seven yearly data sets: Dec 2004, Dec 2004 – Dec 2005, …., Dec 2004 – Dec 2010, where only the December data are used. We then calculate the average FPC. The results are shown in Panel A. To ensure that our results are not affected by the choice of samples from which FPC is derived, we compute the first principal component based on monthly samples. The results are given in Panel B.

Panel A. Yearly Samples (December, U.S. Stock Funds)

SG Component	Dec 04	Dec 04 - 05	Dec 04 - 06	Dec 04 - 07	Dec 04 - 08	Dec 04 - 09	Dec 04 -10	**New PCA(Average)**
BQ	0.496	0.499	0.505	0.470	0.436	0.405	0.414	**0.46**
CC	0.623	0.636	0.632	0.645	0.654	0.663	0.660	**0.64**
FS	0.298	0.268	0.302	0.296	0.300	0.316	0.327	**0.30**
MI	0.163	0.119	0.093	0.083	0.085	0.099	0.145	**0.11**
RH	0.500	0.510	0.496	0.518	0.534	0.536	0.515	**0.52**

Panel B. 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	**New PCA(Average)**
BQ	0.4984	0.4996	0.4956	0.4591	0.4271	0.3962	**0.46**
CC	0.6321	0.6372	0.6365	0.6487	0.6571	0.6641	**0.65**
FS	0.2944	0.2830	0.3005	0.2938	0.3092	0.3245	**0.30**
MI	0.0984	0.0898	0.0824	0.0773	0.0907	0.1148	**0.09**
RH	0.5057	0.5062	0.5021	0.5255	0.5310	0.5324	**0.52**

References

Aalen O. O., Johansen S., 1978. An empirical transition matrix for non-homogeneous Markov chain based on censored observations. Scandinavian Journal of Statistics 5, 141 – 150.

Adkisson, J. A., Fraser, Don R., 2003. Reading the stars: age bias in Morningstar ratings. Financial Analysts Journal 59, 24 – 27.

Anderson, T. W., Goodman L. A., 1957. Statistical inference about Markov chains. Ann. Math. Statistics 28, 89 – 110.

Anderson, K., Hansen L. S., Keiding N., 1991. Non- and semi-parametric estimation of transition probabilities from censored observations of a non-homogeneous Markov process. Scandinavian Journal of Statistics 18, 153 – 167.

Baker, M., Wurgler, J., 2000. Investor sentiment and the cross-section of stock returns. Journal of Finance 61, 1645–1680.

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

Bangia, A., Diebold, F. X., Kronimus, A., Schagen, C., Schuermann, T., 2002. Ratings migration and the business cycle with application to credit portfolio stress testing. Journal of Banking and Finance 26, 445 – 474.

Bebchuk, L. A., 2002. The case against board veto in corporate takeovers. University of Chicago Law Review 69, 973–1035.

Bebchuk, L. A., Cohen A., 2005. The costs of entrenched boards. Journal of Financial Economics 78, 409–433.

Bebchuk L., Cohen, A., Ferrell A., 2009. What matters in corporate governance? The Review of Financial Studies 22, 783–827.

Bebchuk, L. A., Weisbach, M. S., 2010. The state of corporate governance research. The Review of Financial Studies 23, 939–961.

Berk, J. B., Green, R. C., 2004. Mutual fund flows and performance in rational markets. Journal of Political Economy 112, 1269 – 1295.

Blake, C., More, M. R., 2000. Morningstar ratings and mutual fund performance. Journal of Financial and Quantitative Analysis 35, 451-483.

Blume M., 1998. An anatomy of Morningstar ratings. Financial Analysts Journal 54, 19 – 27.

Bollen, Nicholas P., Jeffrey A., Busse J. A., 2005. Short-term persistence in mutual fund performance. Review of Financial Studies 18, 569 – 597.

Brown K. C., Harlow W. V., Starks L. T., 1996. Of tournaments and temptations: an analysis of managerial incentives in the mutual fund industry. Journal of Finance 51, 85-110.

Brown, S. J., Goetzmann W., 1995, Performance persistence, Journal of Finance, 50(2), 679 – 698.

Carhart, M. M., 1997. On persistence in mutual fund performance. Journal of Finance 51, 57-82.

Casavecchia L., Tooman A., 2012. Managerial herding, investor's sensitivity and the role of mutual fund internal governance. Working Paper. University of Technology, Syney and ANZ Banking Group.

Chen Carl R., Huang Y., 2011. Mutual fund governance and performance: a quantile regression analysis of Morningstar's stewardship grade. Corporate Governance: An International Review, 19, 311 – 333.

Chen, Q., Goldstein, I., Jiang W., 2008. Directors' ownership in the U.S. mutual fund industry. Journal of Finance 63, 2629 - 2702.

Chen J., Hong H., Huang M. and Kubik J. D., 2004, Does fund size erode mutual fund performance? Role of liquidity and organization. The American Economic Review 94, 5, 1276 – 1302.

Chou J., Ng L., Wang Q., 2011. Are better governed funds better monitors? Journal of Corporate Finance 17, 1254 – 1271.

Cremers, M., Driessen J. , Maenhout P., Weinbaum D., 2009. Does skin in the game matter? Director incentives and governance in the mutual fund industry. Journal of Financial and Quantitative Analysis 44, 1345 –1373

Del Guercio, D., Dann, L., Partch, M., 2003. Governance and boards of directors in closed-end investment companies. Journal of Financial Economics 69, 111 – 115.

Del Guercio D., Tkac P. A., 2008. Star power: the effect of Morningstar ratings on mutual fund flows. Journal of Financial and Quantitative Analysis 43, 907 - 936.

Ding, B., Wermers, R., 2005. Mutual fund performance and governance structure: the role of portfolio manager and boards of directors. Working Paper. University of Maryland.

Duret, A.S., Hereil, P., Mitaine, P., Moussavi, N., Roncalli. T., 2008. Fund rating systems and performance predictability. Working Paper, Université d'Évry.

Elton E.J., Gruber M.J., Das S., Hlavka M., 1993. Efficiency with costly information: a reinterpretation of evidence from managed portfolios. Review of Financial Studies 6, 1–22.

Elton, E. J., Gruber M. J., Blake C. R., 1996. The persistence of risk-adjusted mutual fund performance. Journal of Business, 69, 133 – 157.

Fama, E., French, K. R., 1993, The common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3-56.

Fama, Eugene F., James D. MacBeth, 1973. Risk, return and equilibrium: empirical tests. Journal of Political Economy 81, 607 – 636.

Ferris S. P., Yan. X., 2007. Do independent directors and chairmen matter? The role of boards of directors in mutual fund governance. Journal of Corporate Finance 13, 392-420.

Garnier O., Pujol Thierry, 2007. Can today's stars be used to read the stars of the future? Case study of equity funds distributed in France between 2000 and 2005..Les Cahiers Scientifiques de l'AMF

Ghosh A., Ng W. S., 2013. Persistence of Mutual Fund Ratings. Working Paper. Singapore Management University.

Gil-Bazo J., Ruiz-Verdu P., 2009. The relation between price and performance in mutual fund industry. Journal of Finance, 64, 2153 – 2183.

Gill, R., Johansen S., 1990. A survey of product-integration with a view towards applications in survival analysis. The Annals of Statistics 184, 1501 – 1555.

Goetzmann, W. N., Ibbotson, R., 1994, Do winners repeat? Patterns in mutual funds selection. Journal of Portfolio Management 20, 9 –18.

Goh J., Ghosh A., Ng W. S., 2013. Grades Matter in Performance: Morningstar Stewardship Grades and Mutual Fund Performance. Working Paper. Singapore Management University.

Gompers, P. A., Ishii, J., Metrick., A., 2003. Corporate governance and equity prices. Quarterly Journal of Economics 118, 107–155.

Gottesman A., Morey M., 2012. Mutual fund corporate governance and performance. Review of Financial Economics, 21, 69 – 81.

Grinbatt, M., Titman, S., 1992. The Persistence of mutual fund performance. Journal of Finance 47, 1977–1984.

Hendricks, D., Patel, J., Zeckhauser, R., 1993. Hot hands in mutual funds: short-run persistence of relative performance, 1974-1988. Journal of Finance 48, 93-130.

Hereil P., Mitaine P., Moussavi N., Roncalli T., 2010. Mutual fund ratings and performance persistence. White Paper. Quant Research by Lyxor.

Huij J., Verbeek M., 2007. Cross-sectional learning and short-run persistence in mutual fund performance. Journal of Banking and Finance 31, 973–997.

Investment Company Institute, 2012. Mutual Fund Fact Book, Washington DC.

Israel, R. B., 2001, Finding generators for Markov chains via empirical transition matrices with applications to credit ratings. Mathematical Finance, 2, 245-265.

Jafry, Y., T. Schauermann, 2004. Measurement, estimation and comparison of credit migration matrices. Journal of Banking and Finance, 28, 2603 – 2639.

Jain P. C., Wu J. S., 2000. Truth in mutual fund advertising: evidence on future performance and fund flows. Journal of Finance 55, 937 – 958.

Jarrow, R. A., F. Lando, S. M. Turnbull, 1997. A Markov model for the term structure of credit risk spreads. Review of Financial Studies. 10, 481–523.

Jegadeesh N., Titman S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48, 65–91.

Jensen, M. C., 1968. The performance of mutual funds in the period 1945-1964. The Journal of Finance 23, 389–416.

Khorana A., 1996. Top management turnover. An empirical investigation of mutual fund managers. Journal of Financial Economics 40, 403-427.

Khorana, A, Nelling E., 1998. The determinants and predictive ability of mutual fund ratings. The Journal of Investing 7, 61–66

Khorana, A., Servaes, H., Wedge, L., 2007. Portfolio manager ownership and fund performance. Journal of Financial Economics 85, 179 – 204.

Kong S. X., Tang D. Y., 2008. Unitary boards and mutual fund governance. Journal of Financial Research 31, 193 – 224.

Kuchler. U., Sorensen, M., 1997. Exponential Families of Stochastic Processes, New York, Springer-Verlag.

Kurniawan M., How J., Verhoeven P., 2012. Monitoring style drift: evidence from equity funds. Working Paper. Queensland University of Technology.

Lai S., Tiwari A., Zhang J., 2009. Mutual fund flows, performance persistence board quality. Working Paper, Singapore Management University.

Lando D., Skødeberg T. M. 2002. Analyzing rating transitions and rating drift with continuous observations. Journal of Banking and Finance, 26, 423-444.

Li, D., Moshirian, F., Pham, P.K., Zein, J., 2006. When financial institutions are large shareholders: the role of macro corporate governance environments. Journal of Finance 59, 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., 1995. Returns from investing in equity mutual funds 1971 – 1991. Journal of Finance 50, 549 – 572.

Morey R. M., 2005, The kiss of death: a 5-star Morningstar fund rating? The Journal of Investment Management 3, 41–52.

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

Morey, R. M., 2002. Mutual fund age and Morningstar ratings. Financial Analysts Journal 58, 2, 56 – 63.

Morningstar Inc., 2007. Fact sheet: the Morningstar Stewardship Grade for Funds URL: http://morningstaradvisor.com/uploaded/pdf/Stewardship.pdf

Morningstar Inc., 2009. The Morningstar rating methodology. URL: http://corporate.morningstar.com/no/documents/MethodologyDocuments/MethodologyPapers/MorningstarFundRating_Methodology.pdf

Navone M., 2011. Investor's distraction and strategic re-pricing decisions. Journal of Banking and Finance, 26, 1291 – 1303.

Nickell P., Perraudin. W., Varatto, S., 2000. Stability of rating transitions. Journal of Banking and Finance 24, 203 – 227.

Rao, D. N., 2010. Portfolio turnover and its effect on performance of equity-oriented mutual fund schemes: an empirical study in the Indian context. Working Paper. Kings University.

Roe, M. J., 1990. Political and legal restraints on ownership and control of public corporations. Journal of Financial Economics 27, 7 – 42.

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

Singer B., Spilerman S., 1976. The representation of social processes by Markov models. American Journal of Sociology 82, 1 – 54 .

Sirri, E. R., Tufano, P., 1998. Costly search and mutual fund flows. Journal of Finance 53, 1589 – 1622.

Truck S., Rachev S. T., 2011, Changes in migration matrices and credit VaR – a new class of difference indices. Working Paper.

Wahal, S., Wang, A. 2011. Competition among mutual funds. Journal of Financial Economics 99, 40–59.

Warshawsky M., DiCarlantonio M., Mullan L., 2000. The persistence of Morningstar ratings. Journal of Financial Planning 13, 110-128.

Wellman, J., Zhou, J., 2007. Corporate governance and mutual fund performance: a first look at the Morningstar stewardship grades. Working Paper, Binghamton University.

Wermers, R., 2000. Mutual fund performance: An empirical decomposition into stock-picking Talent, Style, Transaction Costs And Expenses. Journal of Finance 55, 1655–1695.

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