Mutual Fund Age and Morningstar Ratings

Matthew R. Morey

The study reported here identified an age bias in the Morningstar mutual fund ratings. I found that the average overall star ratings of seasoned funds are consistently—and in many cases, significantly—higher than the average overall star ratings of younger funds. This bias is not the result of a survivorship bias but of the methodology Morningstar uses to calculate the ratings. If star ratings affect fund flows, then this age bias in the Morningstar ratings is of significance to the mutual fund industry and to investors.

With the tremendous growth of privately managed retirement accounts and the more than 10,000 mutual funds now available to investors, mutual fund ratings have never been so popular. As a consequence, many financial publications now devote a substantial amount of coverage to rating mutual funds. One of the most popular, if not the most popular, providers of these mutual fund ratings is Morningstar, Inc. Started in the mid-1980s, Morningstar has grown largely as a result of the success of its five-star rating system. Similar to the ratings of hotels, movies, or restaurants, Morningstar rates mutual funds on a scale of one to five stars, where one star is the worst rating and five stars is the best. Because of the rating system's simplicity and the way in which it mimics the ratings of so many other products consumers buy, the star rating system has truly become part of the accepted lexicon in mutual funds.

The Morningstar rating system has become so popular that a fund receiving a high star rating is often deemed by the public to have something like "a *Good Housekeeping* seal of approval" (Franecki 2000). In fact, some financial planners report that investors require them to invest only in funds with four- or five-star ratings (Franecki). Therefore, not surprisingly, many people believe investment flows in and out of mutual funds are closely related to the Morningstar ratings. For example, a recent study by Financial Research Corporation of Boston, which was reported in the *Wall Street Journal* (Franecki) found that in 1999, funds with four or five stars received inflows of $223.6 billion whereas funds with three or fewer stars had outflows of $132 billion.[1] Moreover, the heavy use of Morningstar ratings in mutual fund advertising (in some cases, the only mention of return performance in the mutual fund advertisement is the Morningstar rating) suggests that mutual fund firms believe that investors care about Morningstar ratings.

Given the inherent interest in the Morningstar ratings by investors, a key question is whether these ratings provide a clear and unbiased picture of mutual fund performance. The purpose of the study reported here was to investigate whether an age bias exists in the Morningstar ratings. In other words, can funds of different ages receive different Morningstar ratings in spite of similar performance? If such a bias could be documented, it would, of course, be of importance to investors and the mutual fund industry and it would also help explain the findings of Warshawsky, DiCarlantonio, and Mullan (2000) that older funds are more likely to show persistence in their Morningstar ratings than are younger funds.

The Age Bias

Morningstar has a number of rating systems, but the focus of this article is on what I call the "overall" Morningstar star rating. It is the most well known of the Morningstar rating systems and is the rating that Morningstar has popularized and marketed in its data since the company started providing ratings.

To calculate the overall rating, Morningstar follows several steps.[2] First, it classifies funds into one of four categories—domestic equity, foreign equity, municipal bond, or taxable bond. Second, using risk-adjusted returns and comparing those returns with those of other funds in the category, Morningstar calculates a fund's star rating for each of three time horizons—3 years, 5 years, and 10 years.[3] I will call these ratings the "time-specific" star ratings. Third, Morningstar weights the time-specific star ratings by the age of the fund. For

Matthew R. Morey is associate professor of finance at the Lubin School of Business, Pace University, New York.

funds with 10 years or more of returns (seasoned funds), Morningstar weights the 3-year star rating by 20 percent, the 5-year star rating by 30 percent, and the 10-year star rating by 50 percent. For funds with 5 to fewer than 10 years of return data (middle-aged funds), Morningstar weights the 3-year star rating by 40 percent and the 5-year star rating by 60 percent. For funds with fewer than 5 but at least 3 years of return data (young funds), Morningstar weights the 3-year star rating by 100 percent. Morningstar then takes this average number and rounds the number up if it has a decimal value of 0.5 or above and rounds it down if it has a decimal value below 0.5. For example, a seasoned fund that had a 4-star rating for the 3-year time horizon, a 4-star rating for the 5-year time horizon, and a 3-star rating for the 10-year time horizon would receive 3.5 stars (that is, 4 stars at 0.2 + 4 stars at 0.3 + 3 stars at 0.5). Morningstar would give this fund a four-star overall rating because the decimal value was at least 0.5. (If the fund had received a 3.4, the fund would have received a three-star overall rating.)

This weighting and rounding system produces an asymmetry in the ratings that is illustrated in **Table 1**. The table considers two realistic situations. In the first situation, the three-year time-specific star rating falls from four stars to three stars (or rises from four stars to five stars) while the other time-specific star ratings remain at four stars. This situation might take place because the very recent returns of the fund are somewhat worse (better) than those of the other funds in the category. The 5-year and 10-year time-specific ratings will not change because they are based on longer-term returns.

In the second situation, the 3-year *and* the 5-year time-specific star ratings fall from four stars to three stars (or rise to five stars) while the 10-year time-specific rating remains at four stars. This situation might take place because recent returns of the fund are worse (better) than the other funds in the category. Again, the 10-year time-specific rating does not change because it is based on longer-term returns.

Table 1. Example Calculation of Overall Star Rating from the Time-Specific Ratings

Age of Fund	3-Year Star Rating	5-Year Star Rating	10-Year Star Rating	Overall Star Rating
Situation 1. Three-year star rating falls from four to three stars (rises to five stars); other time-specific ratings remain at four stars				
Seasoned fund	4 stars	4 stars	4 stars	4 stars
	3	4	4	4
	5	4	4	4
Middle-aged fund	4	4	—	4
	3	4	—	4
	5	4	—	4
Young fund	4	—	—	4
	3	—	—	3
	5	—	—	5
Situation 2. Three-year and five-year star rating falls from four stars to three stars (rises to five stars); other time-specific ratings remain at four stars				
Seasoned fund	4 stars	4 stars	4 stars	4 stars
	3	3	4	4
	5	5	4	5
Middle-aged fund	4	4	—	4
	3	3	—	3
	5	5	—	5
Young fund	4	—	—	4
	3	—	—	3
	5	—	—	5

Note: Weights of seasoned funds (10 or more years of historical returns): 3-year rating, 20 percent; 5-year rating, 30 percent; 10-year rating, 50 percent. Middle-aged funds (5 but fewer than 10 years of historical returns): 3-year rating, 40 percent; 5-year rating, 60 percent. Young funds (3 but fewer than 5 years of historical returns): 100 percent three-year rating.

Table 1 shows that in Situation 1, the decrease or increase in the three-year time-specific ratings has no effect on the overall ratings of seasoned and middle-aged funds. The overall ratings of young funds, however, do decrease or increase. In Situation 2, the decrease or increase in the three-year and five-year ratings only has the effect of increasing or keeping constant the overall ratings of the seasoned funds. For middle-aged funds, however, the increase or decrease in the three-year and five-year ratings can cause the overall ratings to increase or decrease. The point is that a one-star increase in the three-year and five-year ratings can drive up the overall ratings of seasoned funds but a one-star decrease will not drive down the overall ratings. Conversely, the overall ratings of middle-aged and young funds can change up or down more easily.

The asymmetry in the ratings of seasoned funds produces a bias in the overall ratings of seasoned funds that is not present for younger funds. Seasoned funds systematically receive higher overall star ratings—not because they receive higher time-specific ratings but, rather, because of the Morningstar weighting and rounding system. The bias is illustrated in **Table 2**, which shows the ratings for domestic equity funds in June 1999 but is representative also of other categories of funds and other time periods. The table presents the overall and time-specific Morningstar ratings data for seasoned, middle-aged, and young funds and demonstrates that for seasoned funds, the average 3-year, 5-year, and 10-year time-specific ratings are all lower than the average overall star rating. The middle-aged and young funds have overall ratings that are similar to their time-specific ratings. Indeed, in this example, the average of the overall ratings for seasoned funds is higher than that of the middle-aged funds, despite seasoned funds having lower time-specific average star ratings.

To follow this bias further, I examined the ratings of domestic equity funds for every quarter after fall 1991.[4] The results, in **Table 3**, illustrate

Table 2. Distribution of Stars Organized by Age: Example from the June 1999 Morningstar Data Disk

Category	3-Year Time-Specific Rating	5-Year Time-Specific Rating	10-Year Time-Specific Rating	Overall Rating
Seasoned funds				
Number of funds	748	748	748	748
Number of 5-star funds	67	70	74	78
Number of 4-star funds	184	160	169	195
Number of 3-star funds	262	263	261	270
Number of 2-star funds	168	176	169	142
Number of 1-star funds	67	79	75	63
Average rating	3.02	2.95	3.00	3.11
Standard deviation of ratings	1.09	1.12	1.12	1.09
Middle-aged funds				
Number of funds	1,130	1,130	—	1,130
Number of 5-star funds	130	117	—	112
Number of 4-star funds	266	263	—	267
Number of 3-star funds	397	394	—	394
Number of 2-star funds	214	247	—	249
Number of 1-star funds	123	109	—	108
Average rating	3.06	3.03	—	3.02
Standard deviation of ratings	1.15	1.12	—	1.11
Young funds				
Number of funds	1,165	—	—	1,165
Number of 5-star funds	107	—	—	107
Number of 4-star funds	234	—	—	234
Number of 3-star funds	407	—	—	407
Number of 2-star funds	302	—	—	302
Number of 1-star funds	115	—	—	115
Average rating	2.93	—	—	2.93
Standard deviation of ratings	1.10	—	—	1.10

that the average overall star rating of seasoned funds is almost always higher than that for young funds. Only in 1 case out of 37, December 1994, is the average overall star rating of the seasoned funds less than that of the young funds. Moreover, in 26 of 37 cases, the seasoned funds average rating is significantly higher. The average overall star rating of seasoned funds is also generally higher than that for middle-aged funds. In 34 of the 37 cases, the seasoned funds have a higher average rating, although the difference is generally not as great as that uncovered in the comparison of seasoned and young funds.

Survivorship Bias

My results indicate that methodological bias exists toward giving seasoned funds higher overall star

Table 3. Average Overall Ratings Organized by Age of Fund: Domestic Equity Funds

				t-Test of Difference in Overall Ratings		
Date of Disk	Seasoned Funds	Middle-Aged Funds	Young Funds	Seasoned Funds minus Young Funds	Seasoned Funds minus Middle-Aged Funds	Middle-Aged Funds minus Young Funds
September 1991	3.187	2.878	3.077	1.112	3.562**	–1.900*
December 1991	3.189	2.922	2.913	2.788**	3.121**	0.087
March 1992	3.131	3.026	2.880	2.533**	1.265	1.398
June 1992	3.115	3.025	2.935	1.673*	1.088	0.819
September 1992	3.165	3.018	2.892	2.500**	1.802*	1.136
December 1992	3.221	3.058	2.885	3.200**	2.129**	1.597
March 1993	3.173	2.995	2.810	3.274**	2.260**	1.650*
June 1993	3.122	3.019	2.859	2.391**	1.350	1.438
September 1993	3.125	3.014	2.829	2.965**	1.447	1.816*
December 1993	3.117	3.009	2.930	1.820*	1.461	0.752
March 1994	3.044	3.051	2.957	0.915	–0.102	0.981
June 1994	3.069	3.044	3.017	0.593	0.332	0.297
September 1994	3.045	3.082	3.024	0.252	–0.507	0.658
December 1994	3.048	3.031	3.097	–0.611	0.251	–0.797
March 1995	3.066	2.991	3.037	0.358	1.065	–0.548
June 1995	3.155	2.971	2.924	3.012**	2.545**	0.576
September 1995	3.155	2.950	2.983	2.371**	2.845**	–0.422
December 1995	3.142	2.941	2.926	3.161**	2.914**	0.213
March 1996	3.137	2.960	2.954	2.691**	2.606**	0.090
June 1996	3.121	2.937	2.969	2.288**	2.773**	–0.472
September 1996	3.092	2.955	2.958	2.174**	2.090**	–0.052
December 1996	3.055	3.067	3.004	0.881	–0.191	1.005
March 1997	3.074	3.074	2.932	2.523**	0.008	2.231**
June 1997	3.076	3.025	2.935	2.510**	0.819	1.489
September 1997	3.086	3.004	2.963	2.287**	1.370	0.694
December 1997	3.095	3.050	2.913	3.451**	0.745	2.468**
March 1998	3.082	3.012	2.950	2.573**	1.194	1.155
June 1998	3.105	3.011	2.929	3.408**	1.653*	1.585
September 1998	3.087	3.032	2.907	3.508**	0.989	2.550**
December 1998	3.101	3.032	2.914	3.601**	1.282	2.456**
March 1999	3.108	3.024	2.901	4.019**	1.577	2.596**
June 1999	3.111	3.023	2.928	3.560**	1.695*	2.055**
September 1999	3.100	3.031	2.958	2.742**	1.363	1.591
December 1999	3.084	3.018	3.043	0.825	1.348	–0.551
March 2000	3.067	3.041	3.057	0.195	0.541	–0.365
June 2000	3.068	3.043	3.060	0.150	0.526	–0.407
September 2000	3.065	3.041	3.038	0.558	0.521	0.069

*Significant at the 10 percent level.
**Significant at the 5 percent level.

ratings, but another reason may be at work for seasoned funds receiving higher overall ratings—survivorship bias. What can happen is that the seasoned funds receive relatively higher three- and five-year time-specific ratings because the poorly performing seasoned funds progressively drop out of the sample, leaving only the best performers within the seasoned fund group. As a result, the average overall star ratings of seasoned funds will be higher because the seasoned funds' three-year and five-year star ratings are higher.

To investigate whether survivorship bias is causing the higher overall ratings for the seasoned funds, I used a simple dummy variable regression analysis. I examined the following equation:

$$Rating_i = \alpha_0 + \gamma_1(Age_i) + u, \quad (1)$$

where $Rating_i$ is the overall Morningstar star rating and Age_i is a 0,1 dummy variable for age of the fund. The interpretation of the regression is straightforward. The coefficient γ_1 represents the overall star rating differences between the dummy variable and the reference groups. Hence, if 0 represents young funds and 1 represents seasoned funds, a negative (positive) γ_1 implies the group of seasoned funds has a lower (higher) overall star rating than the young funds group. The *t*-statistics on the coefficients provide a test of the significance of the difference between the dummy group and the reference group. The constant term represents the mean overall star rating of the reference group.

Using Equation 1, I examined three cases:

- *Case 1.* For each of the 30 samples (June 1993 to September 2000), I examined only young and seasoned funds that had received three stars for the three-year time-specific star rating. In this way, in essence, I controlled for the three-year rating. Young funds were the reference group, receiving 0 for the dummy variable, and seasoned funds received a 1. Hence, a positive and significant γ_1 implies that the seasoned funds had a higher overall star rating than the young funds, despite all the funds in the sample receiving three stars for the three-year rating period.
- *Case 2.* For each of the 30 samples, I examined only middle-aged and seasoned funds that received three stars for both the three-year and the five-year time-specific star rating. Hence, I controlled for three-year and five-year ratings. Middle-aged funds were the reference group, receiving a 0 for the dummy variable, and the seasoned funds received a 1. A positive and significant γ_1 implies that the seasoned funds had a higher overall star rating than the middle-aged funds, despite all the funds in the sample receiving three stars for the three-year and five-year rating periods.
- *Case 3.* For each of the 30 samples, I examined only young and middle-aged funds that received three stars for the three-year time-specific star rating. Young funds were the reference group, receiving a 0 for the dummy variable, and middle-aged funds received a 1. A positive and significant γ_1 implies that the middle-aged funds had a higher overall star rating than the young funds, despite all the funds in the sample receiving three stars for the three-year rating period.

Table 4, which presents the results of the regressions for each of the three cases, shows that it is the bias in the methodology and not survivorship bias that is causing the higher average ratings for seasoned funds.[5] For Case 1, every single γ_1 is positive and 28 of the 30 are significant at the 5 percent level (29 of the 30 at the 10 percent level). For Case 2, all 30 γ_1 coefficients are positive and strongly significant. Only for Case 3 is a general absence of positive and significant coefficients found. These results suggest that, although no pattern is discernible in the difference of the overall ratings of middle-aged and young funds, evidence does exist that, on average, seasoned funds consistently receive higher overall star ratings than young and middle-aged funds, even after short-term (time-specific) star rating performance has been controlled for.[6]

Conclusions

Investigating the relationship of mutual fund age and Morningstar ratings, I found that the average overall star ratings of seasoned funds are consistently—and in many cases, significantly—higher than the average overall star ratings of middle-aged and young funds. This bias is not a result of survivorship bias but, rather, of the methodology Morningstar uses to calculate the ratings. Specifically, I found that the weighting and rounding systems used by Morningstar make a decline in overall ratings relatively more difficult for seasoned funds.

These results have two implications. First, if investors care about the ratings—and a great amount of anecdotal evidence says that they do—then these results imply that investors should be careful in interpreting the overall ratings as signals of past performance, much less future performance. My findings, together with those of Blake and Morey (2000), which documented that overall star rating does not have much ability to predict winning funds, provide investors with ample evidence

Table 4. Results of Dummy Variable Regressions with Control for Short-Term Ratings: Domestic Equity Funds

Date of Disk	Case 1	Case 2	Case 3
June 1993	0.166*	0.287**	0.021
September 1993	0.211**	0.304**	0.015
December 1993	0.277**	0.333**	0.007
March 1994	0.189**	0.300**	–0.021
June 1994	0.325**	0.370**	–0.007
September 1994	0.225**	0.360**	0.031
December 1994	0.153**	0.285**	–0.014
March 1995	0.185**	0.253**	0.001
June 1995	0.129**	0.200**	0.095
September 1995	0.196**	0.252**	–0.022
December 1995	0.112**	0.177**	–0.047
March 1996	0.073	0.196**	0.054
June 1996	0.149**	0.179**	–0.025
September 1996	0.106**	0.127**	–0.040
December 1996	0.093**	0.109**	0.034
March 1997	0.140**	0.201**	0.035
June 1997	0.183**	0.162**	–0.022
September 1997	0.162**	0.218**	–0.080**
December 1997	0.139**	0.198**	–0.019
March 1998	0.119**	0.160**	–0.019
June 1998	0.145**	0.185**	–0.063**
September 1998	0.177**	0.213**	–0.014
December 1998	0.144**	0.183**	–0.039
March 1999	0.119**	0.153**	–0.016
June 1999	0.149**	0.168**	0.003
September 1999	0.121**	0.196**	0.046*
December 1999	0.143**	0.155**	–0.001
March 2000	0.108**	0.157**	0.076**
June 2000	0.128**	0.166**	0.106**
September 2000	0.069**	0.136**	0.061**

Note: Data did not exist to conduct the test for the September 1991–March 1993 period (see Appendix A).

*Significant at the 10 percent level.
**Significant at the 5 percent level.

that they need to look beyond the ratings when deciding which fund(s) to invest in.

Second, mutual fund rating services may want to use a single consistent time horizon to evaluate funds. Systems that weight time horizons by the age of the fund, such as those used in the Morningstar overall star ratings, can lead to biases that render the ratings more subjective than objective. Morningstar, to its credit, has tried to address this issue in several ways. It has emphasized the time-specific ratings, which compare all funds with the same time horizon. Indeed, these ratings are now available on Morningstar data products, whereas before the fall of 1996, the time-specific ratings were not normally made available. In addition, Morningstar has developed a "category rating" that uses a slightly different methodology from that used for the overall rating and, most importantly, bases all ratings on three-year historical returns.

This research was supported by the TIAA-CREF Institute Research Grant Program. I would like to thank Doug Fore, Mark Hulbert, and Mark Warshawsky.

Appendix A. Data

For my analysis, I used all of the quarterly *Morningstar On-Disk* or *Principia* programs from their inception in September 1991 (September means that the data were current up to September 30) to September 2000 (37 total disks). For each of the 37 disks, I picked all of the rated funds that were in the domestic equity category. Hence, I had 37 separate samples of funds.

I used the domestic equity category rather than others (foreign equity, municipal bond, taxable bond) for several reasons. First, this category usually contains more funds than any of the other categories. Second, the domestic equity category was in existence throughout the September 1991 to September 2000 time period. Third, using domestic equity funds is in the tradition of Blume (1998) and of Blake and Morey, who also examined this category.

The creation of each of the 37 samples was not straightforward. The domestic equity category includes many types of funds. Moreover, the types of funds included in the category have changed over time. From September 1991 to September 1996, the domestic equity category included aggressive growth, equity-income, growth, growth-income, and small-company funds, as well as foreign equity funds and so-called specialty funds. In addition, hybrid funds—such as convertible bond funds, balanced funds, and asset allocation funds—were included in the domestic equity category by Morningstar, but their star ratings were not calculated by the same methods as used for the other funds in the category. As a result of this calculation difference, I excluded the hybrid funds from the samples from September 1991 to September 1996. From December 1996 onward, the domestic equity category still contained the aggressive growth, equity-income, growth, growth-income, small-company, and specialty funds but contained two changes from the earlier periods. First, hybrid funds were included and their ratings were calculated in the same fashion as all the other funds. Hence, for the samples after December 1996, I included the hybrid funds in the samples. Second, the foreign equity funds were excluded from the domestic equity category, so the samples after December 1996 do not include any foreign equity funds.

For the 3-, 5-, and 10-year time-specific star ratings, the data ratings are not uniform throughout all the Morningstar data disks. The disks from September 1991 to March 1993 do not provide any information about the 3-year, 5-year, and 10-year star ratings. Hence, because the overall star ratings are the weighted average of the time-specific ratings, I had no way to assess how the overall ratings were being calculated for this period. The disks from June 1993 to June 1996 also do not provide these time-specific star ratings, but they do provide the 3-, 5-, and 10-year Morningstar return and Morningstar risk numbers from which the time-specific ratings were calculated. With these numbers, I could construct the time-specific star ratings.

Notes

1. A recent working paper by Del Guerico and Tkac (2001) also documents that there is a strong relationship between Morningstar ratings and fund flows.
2. For a thorough description of the Morningstar methodology, see Blume (1998) or Sharpe (1998).
3. Note that Morningstar does not give an overall star rating to funds with fewer than three years of historical returns.
4. Appendix A contains an explanation of these data. All the information about the Morningstar data came from the *Morningstar On-Disk*, *Principia*, *Principia-Plus*, and *Principia-Pro Manuals* (various editions from 1991 to 2000) and the *Morningstar Mutual Fund Sourcebooks* for 1993 and 1994.
5. An explanation of these data is given in Appendix A. Note also that the constant values are not presented here because they all equaled 3.00 (three stars); by taking only funds that received three stars in the three-year and/or five-year star ratings, I restricted the overall ratings of the younger funds in the regression (regardless of whether they were young funds in Cases 1 and 3 or middle-aged funds in Case 2) to all be equal to three stars. Also note that the sample sizes of these regressions were always quite large; they ranged from 132 observations for Case 2, March 1994, to a high of 1,071 observations for Case 3, September 2000. Moreover, no instance of fewer than 40 observations of any particular age of fund occurred.
6. A possible explanation for these results is that I examined only three-star funds for the three-year and/or five-year ratings, but the seasoned funds are in the top half of this three-star class. As a result, the seasoned funds are really doing better than the younger funds, yet we cannot discern this performance because the range is so large for three-star funds. Indeed, if the seasoned funds are in the top half of the three-star group, the result may be that the seasoned funds do relatively better in the 10-year rating period, where the seasoned funds are being compared only with themselves. The end of this process may be that the seasoned funds receive higher overall star ratings. The evidence shows only weak support, however, for this hypothesis. This evidence is available on request.

References

Blake, Christopher, and Matthew R. Morey. 2000. "Morningstar Ratings and Mutual Fund Performance." *Journal of Financial and Quantitative Analysis*, vol. 35, no. 3 (September):451–481.

Blume, Marshall. 1998. "An Anatomy of Morningstar Ratings." *Financial Analysts Journal*, vol. 54, no. 2 (March/April):19–27.

Del Guercio, Diane, and Paula A. Tkac. 2001. "Star Power: The Effect of Morningstar Ratings on Mutual Fund Flows." Working Paper 2001-15, Federal Reserve Bank of Atlanta (August).

Franecki, David. 2000. "Fund Ratings and Recent Results Diverge." *Wall Street Journal* (May 3):C27.

Morningstar On-Disk, Principia, Principia-Plus, and Principia-Pro Manuals. Various editions for 1991–2000. Chicago, IL: Morningstar Publications.

Morningstar Mutual Fund Sourcebook. 1993 and 1994. Chicago, IL: Morningstar Publications.

Sharpe, William. 1998. "Morningstar Performance Measures." *Financial Analysts Journal*, vol. 54, no. 4 (July/August):21–33.

Warshawsky, Mark, Mary DiCarlantonio, and Lisa Mullan. 2000. "The Persistence of Morningstar Ratings." *Journal of Financial Planning*, vol. 13, no. 9 (September):110–121.