

2019

# Morningstar Ratings and Their Predictive Ability of Mutual Fund Performance

Mason Murray  
*Pace University*

Follow this and additional works at: [https://digitalcommons.pace.edu/honorscollege\\_theses](https://digitalcommons.pace.edu/honorscollege_theses)



Part of the [Business Commons](#)

---

## Recommended Citation

Murray, Mason, "Morningstar Ratings and Their Predictive Ability of Mutual Fund Performance" (2019). *Honors College Theses*. 249.  
[https://digitalcommons.pace.edu/honorscollege\\_theses/249](https://digitalcommons.pace.edu/honorscollege_theses/249)

This Thesis is brought to you for free and open access by the Pforzheimer Honors College at DigitalCommons@Pace. It has been accepted for inclusion in Honors College Theses by an authorized administrator of DigitalCommons@Pace. For more information, please contact [nmcguire@pace.edu](mailto:nmcguire@pace.edu).

# **Morningstar Ratings and Their Predictive Ability of Mutual Fund Performance**

by

**Mason Murray**  
**Advised by Matthew Morey**  
**Lubin School of Business**  
**Pforzheimer Honors College**  
**Pace University**  
**One Pace Plaza**  
**New York, NY 10038**

### **Abstract**

This paper begins with a discussion of prior research in the field of mutual fund performance and its determinants, namely Morningstar ratings. Then an examination is conducted regarding Morningstar rating's predictive ability of the performance of U.S. Growth mutual funds from 2015 – 2018 using similar data and methodology as the study by Morey and Gottesman (2006). This study concludes that the ratings are in fact a strong predictor of performance for the selected funds over this time period and that the results are monotonic in that 5-star rated funds perform better than 4-star and so on to 1-star rated funds which are the lowest performers. The implications of these results support the use of Morningstar ratings but investors should remain wary about individual funds and also take note of the economic conditions during the time period that this study examines.

## **Table of Contents**

1. Introduction

2. Literature Review

    2a. Predictors of Outperformance

        2a(I). Expenses and Fees

        2a(II). Past Performance

    2b. Morningstar Ratings

3. Methodology

    3a. Performance

    3b. Predictive Performance

4. Conclusion

## 1. INTRODUCTION

A mutual fund is the go-to choice for investors to achieve diversification and market returns for low costs and transparent management. As of 2017, there was more money in mutual funds than the total net assets of all US commercial banks by over two trillions dollars. But how does the average investor go about selecting which mutual fund in which to invest their hard-earned money? Some might use fees, past returns, or management style, but the simplest criterion to examine would be that of a fund's Morningstar rating. Morningstar is a rating agency that assigns funds a star rating from one, which is the lowest, to five, which is the highest. In the age of Yelp reviews and Uber ratings, Americans are more familiar than ever with simple, straightforward rating systems and rely heavily on them to make decisions from who to hire to where to eat dinner. Because of this, many funds use their assigned Morningstar rating in their advertising rather than other traits such as their own return history.

On their website, Morningstar describes their star rating as “a purely quantitative, backward-looking measure of a fund's past performance.” The difference between comparing a fund's past performance to future performance and Morningstar rating to future performance is not only that the rating is relative to other funds of similar size and in similar categories but also that the ratings are monotonic whereas returns can be calculated to multiple decimals.

Given the wide use of Morningstar ratings, it is important to consider whether these ratings are worth being used to select mutual funds. This paper will discuss prior research in the field of mutual funds and Morningstar ratings and then perform an examination of whether these Morningstar ratings have any predictive ability over the future performance of the funds that they rate.

## 2. LITERATURE REVIEW

### Mutual Funds and Outperformance

In a recent paper by Gruber (2012, p. 3), mutual funds were described as “one of the fastest growing type of financial intermediary in the American economy.” Mutual funds are popular for a number of reasons, including ease of diversification and access to markets that an individual investor would not typically have on their own. These traits are derived from the structure of mutual funds in which investors pool their money for the purpose of investing in a wide range of assets such as stocks, bonds, commodities, derivatives, and other assets. These funds are run by portfolio managers who typically charge a small fee and the returns (or losses) are distributed back to the investors based on the amount invested. The performance of these funds is typically measured against a stated benchmark index outlined in the prospectus.

Outperformance is a commonly debated topic in the realm of mutual funds, as well as in the world of investing in general. For the purposes of mutual funds, this is

defined as the ability to deliver returns higher than that of the comparable benchmark of the fund. There is a substantial amount of research and resulting literature dedicated to answering the questions of whether mutual funds are capable of outperforming and if so, how. All of the studies result in different findings that land all across the board, up and down the spectrum. Lack of consensus is what keeps this topic continuously fascinating and out of it comes new and innovative ideas for measuring exactly how these funds function and what factors can lead to outperformance.

In Kosowski et al. (2006), the authors land on the far end of the spectrum that states that outperformance is not a function of luck and can be executed based on skill. By examining twenty-five years of mutual fund returns they negate the notion that outperformance can be explained by sampling variability and determine that some managers are in fact worth paying high-fees for and that manager selection is a vital step in the process of mutual fund investing. On the other hand, Fama and French (2011) represent the other extreme, concluding that fund outperformance is almost entirely attributed to luck and that no manager decisions show statistically significant in contributing to outperformance. Albeit, using slightly different methodology and time periods, it is still astounding that two studies examining the same topic can result in such varied conclusions.

## 2a. Predictors of Outperformance

## 2a(l). Expenses and Fees

Given the popularity of mutual funds as investment vehicles, selecting the correct mutual fund is still a very important decision. If a mutual fund is an individual's primary investment vehicle, even a couple of percentage points, over time and after compounding, can have major life implications. There are many different attributes of a mutual fund that one can use to select which fund they will end up investing in. One attribute is the fees that a fund charges, usually a small percentage of your investment or returns. Another is the expense ratio which is the amount of operating expenses, management fees, administrative fees, and other asset-based costs (excluding brokerage fees) incurred by the funds compared to the amount of overall fund assets. One can also select a fund based on the manager of the fund and the manager's style of investing or past performance.

Of these attributes, the predictor of outperformance most supported by relevant literature is that of expenses, specifically expense ratio. Studies by Carhart (1997), Dellva and Olson (1998), O'Neal (2004), Haslem, Baker and Smith (2008), and Gil-Bazo and Ruiz-Verdu (2009) all contribute supporting evidence in their conclusions that low expense ratio funds outperform funds with relatively higher expense ratios. In particular, the paper by Carhart (1997) which examines almost 2000 U.S. equity funds

from 1962 – 1993, demonstrates that a one percent increase in expense ratio results in a 1.54 percent decrease in fund performance. More recently, Gil-Bazo and Ruiz-Verdu (2009) offer support to this theory by finding that worse performing (load-adjusted) funds charge higher expense ratios than funds offering higher load-adjusted returns. A study from the Financial Research Corporation in 2002 (Predicting Mutual Fund Performance II: After the Bear (Financial Research Corporation)) as well as a Morningstar study by Kinneil (2010) found that a fund's expense ratio is not only the best predictor of outperformance but the only factor that produced statistically significant results.

The rationale behind the correlation between expense ratios and outperformance is that funds that are capable of outperforming consistently are in higher demand and therefore able to charge more to investors who wish to contribute to the fund. In this vein, Wermers (2000) and Kosowski et al. (2007) find that higher expenses are justified for some funds, as the managers are capable of outperformance via active management.

As with any other research around outperformance and mutual funds, there is a significant amount of findings that support that it is *not* expense ratio, rather a different factor or no factor at all. Cemers and Petajisto (2009) and Patejisto (2013) conclude that the degree to which the holdings in a mutual fund are different than those of the benchmark is the main driver of outperformance. A redemption fee is a fee applied to shares held for short time periods and studies by Dellva and Olson (1998), Finke,

Nanigian and Waller (2009), and Ismailescu and Morey (2012) find a positive relationship between the size and duration of the redemption fee and fund performance. There are also studies from Elton, Gruber and Blake (2003) and Cremers, Dreissen, Maenhout and Weinbaum (2009) that argue that funds with incentive fees outperform those without incentive fees. Incentive fees are paid to managers as a reward for strong performance and are similar to what is referred to “skin in the game” or high ownership stake by managers.

So what is to be drawn from all of these studies that present varying and often contradictory results about expenses loads and other fees? Well, it’s important to note that *generally* researchers find a negative correlation between outperformance and a fund’s expense/loads, but this is not canon and some active managers can, in fact, deliver consistent returns that exceed those of the benchmark and are worthy of their fees. Incentive and manager fees seem to play some role in outperformance, but again not every study supports this. All in all, there is no one factor or attribute of a fund that can determine whether this fund or manager will be capable of delivering desirable returns.

## 2a(II). Past Performance

An obvious consideration when selecting a mutual fund is the past performance of a fund; that is, how has the fund done in the past regarding returns compared to the

returns of the benchmark. Like the factors discussed in the section prior, the findings from relevant literature on the subject vary significantly.

Studies that display that performance persists (outperformers will outperform again and vice versa) include those of Grinblatt and Titman (1992) and Hendricks, Patel and Zeckhauser (1994) and Elton, Gruber and Blake (1996). However, studies including Carhart (1997) find that this persistence is often due to momentum stocks and eventually conclude that the additional transaction costs of investing in these momentum stocks are so high that they eradicate gains from following this strategy. This means that persistence only ends up applying to poor performing funds. Thus the consensus is that outperforming will persist up to a few quarters while poor performance will typically persist longer and neither should necessarily be used entirely to select a fund.

## 2b. Morningstar Ratings

It is clear that of all the factors considered thus far, none are satisfactory in selecting a fund. Expenses, fees, and past performance were good predictors in some findings, bad in others, and inconclusive in most. There are many private providers of mutual fund “ratings.” The most famous and regarded of which is that from Morningstar. Morningstar uses an algorithm that involves a combination of risk-adjusted past performance and awards mutual funds a star rating from one to five. This is inherently

appealing to investors because of its simplicity and because it has become familiar and standard in the industry. These ratings have become so popular in fact that they have a significant impact on fund flows as evidenced in findings from Del Guericco and Tkac (2005). They conclude that a fund with a 5-star rating produces inflows of 53 percent above the normal flow while downgrades in star ratings can result in significant outflows, more so than would be justified solely from a change in performance.

Due to the commonality and usage of these ratings, a substantial amount of research has been done on whether ratings have any sort of predictive ability including studies from Blake and Morey (2000), Morey (2002a, 2002b, 2005) and Morey and Gottesman (2006). Once again we see contradictory results from different studies. Some studies find that funds that receive higher ratings relative to others do not result in outperformance, like Blake and Morey (2000). In their findings, the 5-star funds performed about the same as 3-star rated funds after ratings were established. Phillips and Kinniry (2010) support these findings.

In contrast, Morey and Gottesman (2006) again tested the predictive ability of Morningstar ratings after Morningstar updated their methodology in rating mutual funds. In this study, they found that the new rating system was, at least in some part, predictive of future fund performance as higher rated funds outperformed lower. In this study, they even found that funds with a rating one star lower than other funds

performed lower (so 5-star performed better than 4-star which performed better than 3-star and so on).

In conclusion, there is a wealth of dissonance, contradiction, and uncertainty in all the studies considered in this review as well as in the industry as a whole. No factor seems to be consistent in predicting performance, but on the other hand, every factor seems to play some sort of role. I have decided to reexamine the Morningstar rating's predictive ability because of the (perhaps exaggerated) common use of the ratings and the variability in past studies. The contrast in these findings, as well as the fact that they could be considered outdated, are justification for this study in which I recreate the Morey and Gottesman (2006) study with data from 2015-2018.

### 3. METHODOLOGY

I will take the Morningstar rating on January 1, 2015 for all US Equity Large Cap Growth mutual funds (single share class) and measure monthly mean returns, using the Sharpe ratio and the 30-day T-bill rate, through January 1, 2018. I will then use a regression analysis to compare the funds' ratings with the performance. A regression analysis allows a comparison of the performance of each group of funds based on their Morningstar rating. I will use the 3-star rated funds as the reference group because it is the center rating and will allow me to see how much better (or worse) 5-star and 4-star

funds perform and how much worse (or better) 2-star and 1-star funds perform, relatively.

To select funds I use the January 2015 Morningstar Mutual Funds Data Disk. This disk provides data for funds as of January 1, 2015. From this disk, I then select all domestic equity funds that received an overall Morningstar mutual fund rating. This produced a sample of 4,700 funds. I use domestic equity funds as these styles are the most popular for domestic investors.

I then narrow the sample by eliminating replicate funds. Replicate funds are funds that are identical to another fund in our sample except that they are sold as different share classes. Since I do not want to over count the number of funds in the sample, I include only one of the fund's share classes in the sample. By eliminating these replicate funds, funds in other currencies, and those that did not have a ticker symbol or Morningstar rating on January 1, 2015, I attain a sample of 305 funds.

For each of the 305 funds, I then attempt to obtain the out-of-sample returns of these funds. For a large majority of the funds (approximately 80 percent) obtaining the out-of-sample returns is simply a matter of following the fund's future performance. However, for about 20 percent of the sample, the funds disappear prior to January 1, 2018 as a result of liquidations or mergers. This means that my study is affected by

survivorship bias, but the purpose of the results are for directional insight instead of an exact correlation calculation.

To calculate the ratings, Morningstar uses load-adjusted returns to reflect the extra expenses that the funds charge. However, the monthly fund returns that are available on Morningstar data are not adjusted for front or deferred loads. As a result, it is important for us to adjust the out-of-sample returns for loads when investigating whether the ratings are able to predict future performance. Moreover, for investors, load-adjusted performance is what they are typically concerned with.

### 3a. Performance

A risk-adjusted performance metric is used to measure out-of-sample performance: the Sharpe ratio. For each performance metric, only load-adjusted returns are examined.

Sharpe ratio:

where  $r_i$  are the monthly returns in excess of the 30-day T-bill rate,  $\sigma_i$  of the  $i$ th mutual fund during the out-of-sample period, and  $\sigma$  is the standard deviation of the  $r_i$ .

### 3b. Predictive Performance

A cross-sectional dummy variable regression analysis is used to examine the out-of-sample predictive performance:

where  $y_i$  is the out-of-sample performance metric for fund  $i$ .

$D5 = 1$  if the fund received 5-star overall Morningstar rating as of January 1, 2015, 0 if not,

$D4 = 1$  if the fund received a 4-star overall Morningstar rating as of January 1, 2015, 0 if not,

$D2 = 1$  if the fund received a 2-star overall Morningstar rating as of January 1, 2015, 0 if not,

$D1 = 1$  if the fund received a 1-star overall Morningstar rating as of January 1, 2015, 0 if not,

$i = 1$  through  $N$ , where  $N$  is the total number of funds in the sample.

In the above equation, the 3-star fund group for the Morningstar sample is the reference group for the dummy variable regressions. Hence, when using the Sharpe ratio as the out-of-sample performance measure, the coefficient,  $\gamma_0$  represents the expected Sharpe ratio when all the dummy variables are equal to 0, and the coefficients  $\gamma_1$  through  $\gamma_4$  represent the differences between the dummy variables and the reference group. For example, a negative  $\gamma_1$  implies the group of 4-star funds performs worse than the group of 5-star funds; a positive  $\gamma_1$  implies the group of 4-star funds outperforms the 5-star fund group. The t-statistics on the coefficients provide a test of

the significance of the difference between an individual dummy group and the reference group.

I use the 3-star funds as a reference group because it provides a center from which I can compare the performance of the higher-rated and lower-rated group funds. If the ratings accurately predict out-of-sample performance I should see increasingly positive (and significant) coefficients for five and four and increasingly negative (and significant) coefficients for two and one.

#### 4. RESULTS

The results of the dummy variable regression are presented in the table below.

Five has a positive coefficient that is greater than four which also has a positive coefficient and two has a negative coefficient which is greater than one which also has a negative coefficient. The p-values are all low enough for the results to be statistically significant at a 5% significance level, other than the two-star rated funds which are significant at a 15% level. Based on this data, Morningstar ratings are in fact predictive of mutual fund performance and funds with higher star ratings deliver higher load-adjusted returns than those with a rating lower relative rating.

#### 5. CONCLUSION

In this study, I perform an examination of how the Morningstar rating system predicts future fund performance. I investigate the predictive power of the Morningstar rating system using all domestic Growth equity funds that were rated by Morningstar as

of January 1, 2015 and then examine the performance of all these funds over the next three years, out-of-sample, January 2015 – January 2018.

I find widespread support for the notion that the new Morningstar rating system can predict future performance, at least in the first three years out of sample.

Specifically, I find that higher rated funds, for the most part, significantly outperform lower rated funds. Moreover, the effect is relatively monotonic as even the next to lowest rated funds (2-star funds) significantly outperform the lowest rated funds (1-star funds). These results are quite different from those of Blake and Morey (2000) and Morey (2002b) which indicated that the previous Morningstar rating system did not predict future performance well.

One explanation for this disparity may be fund expenses. Research by Carhart (1997) has found that fund expenses are inversely related to future fund performance. That is, funds with low expenses seem to perform better in the future than those with high expenses. Thus, an explanation for our results could be that before the change in the ratings methodology, 5-star funds had similar expenses to the other funds and hence did not perform better in the future than other, lower-rated funds. Yet, after the change in the methodology, 5-star funds have significantly lower expenses than other funds and thus the high star rating can predict strong future performance.

Another, better explanation may be the “hot hands phenomenon” as first described by Hendricks, Patel and Zeckhauser (1988). They find that there is short-term persistence in fund performance. Given that my out-of-sample period is only three years long, I may be capturing this effect. Indeed, if the methodology of the rating system was flawed before June 2002, it is possible that the hot hands effect would not have shown up in performance prediction results from the old Morningstar ratings methodology. Now, with the improvements in the ratings methodology, the hot hands phenomenon may surface.

It’s important to take note of the time period and economic conditions of the time period in which these returns were recorded. 2015 – 2018 was a very positive economic environment in which the stock market consistently rose with less than historically average corrections or downturns. It also was a very low interest rate environment, coming off of the recession but after full recovery which allowed business to reinvest heavily. This was also a time period characterized by a notable amount of share buybacks from companies, inflating stock prices over time. Given all of this, mutual funds required little strategy other than tracking the market to deliver returns. The methodology of this study did not consider the benchmark to which these funds were matched but only relative performance to funds in the sample. This being said, funds that performed poorer over this time period could very well be positioned to outperform

in a bear market or a time period with more corrections and less stable economic forces.

The implications of these findings are far-reaching considering the wide use of Morningstar ratings by investors to determine which funds to invest in. The conclusion that the ratings are predictive of performance means that the heightened inflows into higher rated funds found by Del Guericco and Tkac (2005) are justified. It also means that Morningstar has successfully honed its rating methodology. Investors should remain wary, because although the sample reflects that the ratings are predictive of returns, it does not mean that individual funds necessarily will outperform or underperform based on rating.

## REFERENCES

- Blake, Christopher and Matthew Morey, 2000, "Morningstar Rating and Mutual Fund Performance", *Journal of Financial and Quantitative Analysis*, 35, 3, 451-83.
- Carhart, Mark, 1997, "On Persistence in Mutual Fund Performance", *Journal of Finance*, 52, 1, 57-82.
- Cremers, Martijn, Joost Driessen, Pascal Maenhout, and David Weinbaum. 2009. "Does Skin in the Game Matter? Director Incentives and Governance in the Mutual Fund Industry." *Journal of Financial and Quantitative Analysis*, 44 (6), 1345-1373.
- Cremers, Martijn. and Antti Petajisto. 2009. "How Active is Your Fund Manager? A New Measure That Predicts Performance." *Review of Financial Studies* 22: 3329-3365.
- Del Guercio, D. and P.A. Tkac, 2008. "Star Power: The Effect of Morningstar Ratings on Mutual Fund Flow." *Journal of Financial and Quantitative Analysis*, 43 (4), 907-936.
- Del Guercio, Diane and Paula A. Tkac, 2005, "The Effect of Morningstar Ratings on Mutual Fund Flows", Working Paper, Atlanta Federal Reserve Bank.
- Dellva, W.L., and Olson, G.T. 1998. "The relationship between mutual fund fees and expenses and their effects on performance." *The Financial Review*, 33, 85-104.

Elton, E.J., Gruber, M.J., and Blake, C.R. 2003. "Incentive fees and mutual funds." *The Journal of Finance*, 58(2), 779-804.

Elton, E.J., Gruber, M.J., and Blake, C.R. 2012. "Does Mutual Fund Size Matter? The Relationship Between Size and Performance." *Review of Asset Pricing Studies*, 2 (1), 31-55.

Gruber, M. J. (1996), Another Puzzle: The Growth in Actively Managed Mutual Funds. *The Journal of Finance*, 51: 783-810. doi:10.1111/j.1540-6261.1996.tb02707.x

Fama, Eugene, and Kenneth French. 2011. "Luck versus skill in the cross-section of mutual fund returns." *Journal of Finance* 65, 1915–1947.

Financial Research Corporation. 2002. *Predicting Mutual Fund Performance II: After the Bear*. Boston: Financial Research Corporation Publications.

Finke, Michael, David Nanigian and William Waller. 2009. "Redemption Fees: Reward for Punishment." Working Paper, Texas Tech University.

Gottesman, Aron and Matthew Morey. 2006. "Manager Education and Mutual Fund Performance" *Journal of Empirical Finance*, 2006, 13, 2, 145-182.

Gottesman, Aron and Matthew Morey. 2012. "Does a Mutual Fund's Corporate Culture Predict Fund Performance." *Review of Financial Economics*, 2012, 21, 69-81.

Gil-Bazo, J., and P. Ruiz-Verdu. 2009. "The Relation between Price and Performance in the Mutual Fund Industry." *Journal of Finance*, 64 (5), 2153-2183.

Grinblatt, Mark and Sheridan Titman. 1992. "The Persistence of Mutual Fund Performance." *Journal of Finance*, 47, 1977-1984.

Grinblatt, Mark and Sheridan Titman. 1994. "A Study of Monthly Mutual Fund Returns and Performance Evaluation Techniques." *Journal of Financial and Quantitative Analysis*, 29 (3), 419-444.

Haslem, J.A., H.K. Baker, and D.M. Smith. 2008. "Performance and characteristics of actively managed retail equity mutual funds with diverse expense ratios." *Financial Services Review*, 17, 49-68.

Hendricks, D., J. Patel, R. Zeckhauser, 1988, "Hot Hands in Mutual Funds: Short-Run Relative Persistence of Relative Performance 1974-1988", *Journal of Finance*, 48, 93- 130.

Ismailescu, Iuliana and Matthew Morey. 2012. "Redemption Fees and the Risk-Adjusted Performance of International Equity Mutual Funds." *Journal of Investment Management*, 2012, 10, 3, 21-33.

Kinnel, Russell. 2010. "How Expense Ratios and Star Ratings Predict Success." *Morningstar Fund Investor*, 18(12).

Kosowski, Robert, Allan Timmermann, Hal White, and Russ Wermers. 2006, "Can Mutual Fund "Stars" Really Pick Stocks? New Evidence from a Bootstrap Analysis." *Journal of Finance*, 61, 2551-2595.

Morey, Matthew, R., 2002a, "Mutual Fund Age and Morningstar Ratings", *Financial Analysts Journal*, March/April, 56-63.

Morey, Matthew, 2002b, "Rating the Raters: An Investigation into Mutual Fund Rating Services", *The Journal of Investment Consulting*, November/December, 5, 2, 30-50.

Morey, Matthew, 2005, "The Kiss of Death? A 5-Star Morningstar Rating", *Journal of Investment Management*, 2005, 3, 2, 41-52.

O'Neal, Edward.S. 2004. "Purchase and redemption patterns of U.S. equity mutual funds." *Financial Management*, 63-90.

Petajisto, Antti. 2013. "Active Share and Mutual Fund Performance." *Financial Analysts Journal*, 2013, 69(4):73-93

Phillips, Christopher and Francis M. Kinniry. 2010. "Mutual Fund Ratings and Future Performance." Working Paper, Vanguard Investments.

Wermers, Russ. 2000. "Mutual fund performance: An empirical decomposition into stock-picking talent, style, transaction costs, and expenses." *Journal of Finance* 55, 1655-1703.