Morningstar Quantitative Rating™ for funds

Methodology

Morningstar Quantitative Research
12 July 2018
Version 1.5

Content

Introduction

Morningstar has been conducting independent investment research since 1984. Traditionally, our approach has been to provide analyst-driven, forward-looking, long-term insights to help investors better understand investments. Morningstar has one of the largest independent manager research teams in the world, with more than 100 analysts globally covering more than 3,700 unique funds.

The Morningstar Analyst Rating™ for funds (the Analyst Rating) provides a forward-looking evaluation of how these funds might behave in a variety of market environments to help investors choose superior funds. It's based on an analyst's conviction in a fund's ability to outperform its peer group and/or relevant benchmark on a risk-adjusted basis through a full market cycle of at least five years.

The number of funds that receive an Analyst Rating is limited by the size of the Morningstar analyst team. To expand the number of funds we cover, we have developed a machine-learning model that uses the decision-making processes of our analysts, their past ratings decisions, and the data used to support those decisions. The machine-learning model is then applied to the "uncovered" fund universe and creates the Morningstar Quantitative Rating™ for funds (the Quantitative Rating), which is analogous to the rating a Morningstar analyst might assign to the fund if an analyst covered the fund. These quantitative ratings predictions make up what we call the Morningstar Quantitative Rating. With this new quantitative approach, we can rate nearly 6 times more funds in the global market.

Only open-end funds and exchange-traded funds that don't currently have an Analyst Rating and are in a category that Morningstar currently rates are eligible to receive a quantitative rating. With the introduction of the Quantitative Rating, we're extending a useful analytic tool to thousands of additional funds, providing investors with much greater breadth of coverage from the independent perspective they have come to know and trust from Morningstar.

Authors

Lee Davidson, CFA
Head of Quantitative Research
+1 312-244-7341
lee.davidson@morningstar.com

Xin Ling
Senior Quantitative Developer
+1 312-696-6191
xin.ling@morningstar.com

Madison Sargis
Associate Director of Quantitative Research
+1 312-244-7352
madison.sargis@morningstar.com

Timothy Strauts
Director of Quantitative Research, Funds
+1 312-384-3994
timothy.strauts@morningstar.com

Philosophy of Morningstar Quantitative Rating™ for Funds

Morningstar has been producing differentiated investment research since 1984. Although Morningstar research has expanded to equity, corporate credit, structured credit, and public policy, our roots are in the world of mutual funds. Traditionally, our approach has been to provide analyst-driven, forward-looking, long-term insights alongside quantitative metrics to further understanding of the investment landscape. Recently, we developed a way to combine our analyst-driven insights with our robust fund data offering to expand fund analysis beyond the capabilities of our manager research staff. With this new development, we will be able to cover 6 times more funds in the global market through empirical methods that are based on the proprietary ratings our analysts are already assigning to funds.

In general, there are two broad approaches that we could have chosen to expand our analyst-driven rating coverage in a quantitative way: either automate the analyst thought process without regard for output similarity; or, replicate the analyst output as faithfully as possible without regard for the analyst thought process. Attempting to mechanically automate a thought process introduces tremendous complexity, so we opted to build a model that replicates the output of an analyst as faithfully as possible.

Replicating the Analyst Rating was a desirable goal because Morningstar has demonstrated throughout its history that the recommendations of its analysts provide value to investors. Therefore, at the outset, it seemed plausible that if a statistical model could be created that replicated the decision-making process of analysts, then there stood a decent chance it would produce valuable results as well. Indeed, based on our 14-year back-test, this is exactly what we found.

But perhaps the most obvious benefit to investors of the quantitative set of ratings is the breadth of coverage and frequency of update. Our quantitative coverage universe is many times the size of our analyst-covered universe, and growing. It is limited only by our access to the necessary input data. Additionally, the Morningstar Quantitative Rating has the unique advantages of maintaining a monthly update cycle. Each fund's rating is refreshed on a frequency unsustainable by a fund analyst.

Of course, no rating system—quantitative or analyst—is valuable without empirical evidence of its predictive ability. We have rigorously tested the performance, accuracy, and stability of the Quantitative Rating. We have included in this document numerous studies performed on the ratings and will continue to enhance our methodologies over time to improve performance.

Morningstar Quantitative Rating Descriptions

The Quantitative Ratings are composed of the Morningstar Quantitative Rating™ for funds, Quantitative Parent Pillar, Quantitative People Pillar, Quantitative Performance Pillar, Quantitative Price Pillar, and Quantitative Process Pillar. A high level description of each rating is found below. The statistical model is described in the Overview Methodology section on page 4. The pillar rating methodology begins on page 5.

- Morningstar Quantitative Rating™ for funds: Comparable to Morningstar's Analyst Ratings for open-end funds and ETFs, which are the summary expression of Morningstar's forward-looking analysis of a fund. The Analyst Rating is based on the analyst's conviction in the fund's ability to outperform its peer group and/or relevant benchmark on a risk-adjusted basis over a full market cycle of at least five years. Ratings are assigned on a five-tier scale with three positive ratings of Gold, Silver, and Bronze; a Neutral rating; and a Negative rating. Morningstar calculates the Quantitative Rating using a statistical model derived from the Analyst Rating our fund analysts assign to open-end funds.

- Quantitative Parent Pillar: Comparable to Morningstar's Parent Pillar ratings, which provide Morningstar's analyst opinion on the stewardship quality of a firm. Morningstar calculates the Quantitative Parent Pillar using an algorithm designed to predict the Parent Pillar rating our fund analysts would assign to the fund. The Quantitative Rating is expressed as Positive, Neutral, or Negative.

- Quantitative People Pillar: Comparable to Morningstar's People Pillar ratings, which provide Morningstar's analyst opinion on the fund manager's talent, tenure, and resources. Morningstar calculates the Quantitative People Pillar using an algorithm designed to predict the People Pillar rating our fund analysts would assign to the fund. The Quantitative Rating is expressed as Positive, Neutral, or Negative.

- Quantitative Performance Pillar: Comparable to Morningstar's Performance Pillar ratings, which provide Morningstar's analyst opinion on the fund's performance pattern of risk-adjusted returns. Morningstar calculates the Quantitative Performance Pillar using an algorithm designed to predict the Performance Pillar rating our fund analysts would assign to the fund. The quantitative rating is expressed as Positive, Neutral, or Negative.

- Quantitative Price Pillar: Comparable to Morningstar's Price Pillar ratings, which provide Morningstar's analyst opinion on the fund's value proposition compared to similar funds sold through similar channels. Morningstar calculates the Quantitative Price Pillar using an algorithm designed to predict the Price Pillar rating our fund analysts would assign to the fund. The Quantitative Rating is expressed as Positive, Neutral, or Negative.

- Quantitative Process Pillar: Comparable to Morningstar's Process Pillar ratings, which provide Morningstar's analyst opinion on the fund's strategy and whether the management has a competitive advantage enabling it to execute the process and consistently over time. Morningstar calculates the

Quantitative Process Pillar using an algorithm designed to predict the Process Pillar rating our fund analysts would assign to the fund. The Quantitative Rating is expressed as Positive, Neutral, or Negative.

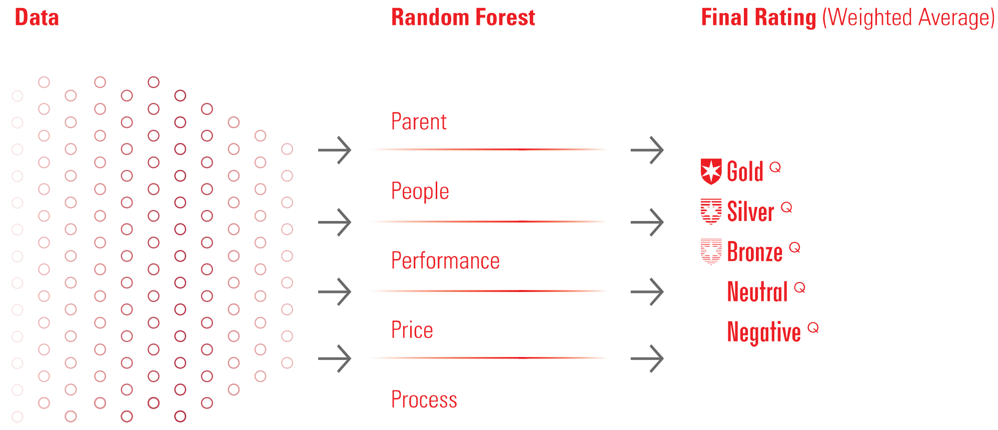
Overview of the Quantitative Rating Methodology

The Quantitative Rating consists of a series of 11 individual models working in unison that were designed to provide a best approximation for the Analyst Rating on the global universe of open-end funds and ETFs. Visually, you can think of the estimation as being a two-layered process. First we estimate the pillar ratings for each fund, and then we estimate the overall rating.

To estimate the pillar ratings, we chose a machine-learning algorithm known as a "random forest" to fit a relationship between the fund’s pillar ratings and its attributes. For each pillar, two random forest models were estimated that seek to determine the probability that fund will be rated Positive or Negative, respectively. Since there are five pillars, we estimated 10 individual random forest models to answer these questions and produce 10 probabilities (two per pillar). Then, at the pillar level, we aggregate these probabilities to produce one overall pillar rating.

After the pillar ratings are estimated, we needed to aggregate them into an overall fund rating. In order to do this, we used a multivariate linear regression. The final result is the Morningstar Quantitative Rating™ for funds.

Exhibit 1 Representation of a Morningstar Quantitative Rating Methodology



Source: Morningstar, Inc.

Morningstar Quantitative Rating—Pillar Rating Methodology

The five pillar ratings represent the foundation of the Analyst Rating. For the Quantitative Rating, the pillar ratings were estimated using a series of random forest models and rated on a scale of Positive, Neutral, and Negative.

In order to estimate the pillar ratings, data was collected for the funds that analysts have currently assigned pillar ratings. In total, 180-plus attributes and 10,000-plus rating updates were considered in order to train the random forest model. After numerous iterations, only the attributes most crucial to classifying each pillar rating were retained.

Each pillar rating is estimated using a combination of two random forest models. First, a model is estimated that seeks to distinguish funds based on whether that fund's pillar rating would be rated Positive. Second, a different model is estimated that seeks to distinguish funds based on whether that fund's pillar rating would be rated Negative. Each model puts out probability scores that the fund would be Positive and Negative. By combining these two probabilities via a weighted summation, a more robust estimator is achieved.

$$Estimated\ Pillar\ Rating = \frac{Prob(Positive)+[1-Prob(Negative)]}{2}$$

The output for these pillar ratings will, therefore, be on a scale of 0 to 1. The closer to 1 a fund's estimated pillar rating is, the more likely it is that the true pillar rating is Positive. Similarly, the closer to 0 a fund's estimated pillar rating is, the more likely that the true pillar rating is Negative. After the ratings were computed, thresholds were assigned that tended to correspond to natural distinctions between Negatives, Neutrals, and Positives for each pillar.

The intuition underlying this method is subtle, yet important. First, the weighted summation captures information about a fund along two dimensions—the likelihood that a fund's pillar is Positive and the likelihood that a fund is not Negative. In practice, this has the result of classifying many Neutral pillars as decidedly not Positive and not Negative.

Furthermore, by using two models to estimate a pillar rating, we are able to distinguish between data points that are important to each model individually. It makes intuitive sense that the data points that might indicate to an analyst to rate a fund Positive could be different from those that are used to rate a fund Negative. By adding in that flexibility, we dramatically improved our estimation. Empirically, several pillar models exhibited significant overlap in data points used to estimate each model, but that did not always hold.

Smoothing Algorithm

After raw pillar ratings have been computed, we implement a smoothing algorithm to reduce inter-month volatility. This algorithm takes the average of the current raw pillar rating and the two prior months' raw pillar ratings to create a three-month moving average. The three-month moving average was chosen to balance the desire to reduce unnecessary volatility of ratings month-to-month, but also allow the ratings to be adaptable to significant changes at the fund, such as a manager change.

People and Process Pillar Business Logic

After smoothing, we implement a business rule to ensure that People and Process Pillar ratings do not change depending on the share class. Technically, each fund share class will have their own People and Process Pillar ratings produced by the model, but we want to ensure that these are consistent for the same fund. To ensure this, we implement an asset-weighted average of raw People and Process Pillar ratings across share classes with the weights determined by share-class level net assets. In the case where net assets are not available, share class level ratings will be equally weighted. The final raw Pillar ratings, after smoothing and asset-weighting, are saved as the pillar rating estimate for the current month for each fund share class.

In the case where an analyst has rated a fund belonging to the same strategy, all other funds under that same strategy identifier will inherit the People and Process Pillar rating assignments as determined by the analyst. This ensures that the analyst view is leveraged whenever available to ensure consistency between the Analyst Rating and Quantitative Rating systems when it comes to the People and Process Pillars.

Parent Pillar Business Logic

In the same spirit, we implement one final business rule. In the case where there is an Analyst Rating for the Parent Pillar of a fund for a particular branding entity, we will suppress the Quantitative Parent Pillar for all funds from that particular branding entity and default to the analyst opinion. In this way, we ensure consistency of opinion between analyst and quant rating systems when it comes to the Parent Pillar.

Pillar Threshold

For those pillars where an analyst rating is not available, pillar labels (Positive, Neutral, or Negative) will be assigned according to a static threshold to the raw pillar ratings:

- If raw pillar rating < 0.25, then Negative
- If raw pillar rating <= 0.75 and >= 0.25, then Neutral
- If raw pillar rating > 0.75, then Positive

Calculating the Quantitative Rating

The final step in the Quantitative Rating involves predicting an overall rating on the scale of Negative, Neutral, Bronze, Silver, or Gold from our estimated pillar ratings. To accomplish this task, a multivariate linear regression has been employed. The model performs well when attempting to predict overall analyst ratings out-of-sample. Compared with other methods, the multivariate linear regression has obvious advantages in terms of transparency.

This model was built very simply. First, we take our sample of true Morningstar Analyst ratings and assign them numerical values: Negative = 1, Neutral = 2, Bronze = 3, Silver = 4, and Gold = 5. Then we assign the true pillar ratings numerical values: Negative = 0, Neutral = 0.5, and Positive =1. Then we run a multivariate linear regression to identify slope coefficients for each pillar. This model has the benefit of not only telling us how we might expect an overall rating to change given an incorrect pillar rating, but also how to construct overall ratings based on a set of pillar ratings. In short, it is an extremely simple, easy-to-interpret, and transparent model that works well in practice.

Exhibit 2 Sample Slope Coefficients for Each Pillar

Independent Variables	Coefficients
Parent	0.83
People	0.71
Performance	0.88
Price	0.64
Process	1.37

Source: Morningstar, Inc.

Based on the regression results, we see that each pillar weight is estimated at different values. For example, the Process Pillar appears to be the largest determinant of the overall ratings. The slope coefficients can be interpreted as follows: Given a change in the corresponding pillar rating (0 to 0.5 or 0.5 to 1) we can expect an X amount of change in the overall rating. For instance, say we increased a Parent Pillar rating to Positive from Negative (that is, to 1 from 0), then we would expect that the overall rating increases 0.83 units, where 1 unit is equal to 1 rating. The slope coefficients listed above are just examples. We re-estimate these slope coefficients each month when applying the model.

Rating Threshold

After pillar ratings have been assigned and regression weights estimated, we estimate the overall rating using the regression weights multiplied by the pillar ratings. Then, we use a chi-squared distribution algorithm to map these discrete overall ratings into a continuous distribution and use fixed percentile thresholds for final rating assignment. Exhibit 3 showcases these distribution breakpoints.

Exhibit 3 Rating Distribution Breakpoints

Morningstar Quantitative Rating™ for Funds	Breakpoints
Gold Q	> 95%
Silver Q	85%–95%
Bronze Q	70%–85%
Neutral Q	15%–70%
Negative Q	0%–15%

Source: Morningstar, Inc.

To increase the rating stability for funds near the breakpoints, we implement a buffering system. Between Negative—Neutral and Neutral—Bronze, the buffer is 2%. Between Bronze—Silver and

Silver—Gold, the buffer is 1%. A fund near the rating thresholds must move past the buffer before the rating changes. For example, a fund below the 15.0 percentile will need to move to the 17.0 percentile before the rating upgrades from Negative to Neutral. Similarly, a fund above the 15.0 percentile will need to move below the 13.0 percentile before being downgraded from Neutral to Negative.

Model Accuracy

The Morningstar Quantitative Rating model is constructed to mimic the rating assignment behavior of our manager research staff. While we believe that forecasting out-of-sample future performance is the most important aspect for investors, we have tested the accuracy of Quantitative Rating in its ability to match the Analyst Rating.

Much of the inconsistency we observe between the Quantitative Rating and the Analyst Rating is restricted to the 'Recommended' class of funds (Gold, Silver, and Bronze). Specifically, the model finds it difficult to distinguish the differences between the three different recommended ratings. However, the differences between Negative and Neutral, or Neutral and 'Recommended', appear to be quite readily captured by the model. Exhibit 4 shows the percentage agreement between the two rating systems. For example, if the Analyst Rating for a fund is assigned one of the three 'Recommended' ratings, the Quantitative Rating for that fund also comes up as 'Recommended' 77.8% of the time. There are very few instances of large disparities between Analyst Ratings and Quantitative Ratings. Funds rated Negative by analysts are only 'Recommended' 4.4% of the time. Conversely, funds 'Recommended' by analysts are rated Negative by the Quantitative Rating only 0.9% of the time. Overall, we are happy with the precision of the Quantitative Rating as we balance the desire to increase accuracy, avoid overfitting, and achieve strong future performance.

Exhibit 4 Percentage Agreement Between Quantitative Rating and the Analyst Ratings

	Quantitative Rating		
Analyst Rating	Negative Q	Neutral Q	**Recommended** Q
Negative	**55.07**	**40.53**	**4.41**
Neutral	**8.65**	**59.05**	**32.29**
Recommended	**0.96**	**21.20**	**77.84**

%
70
50
30
10

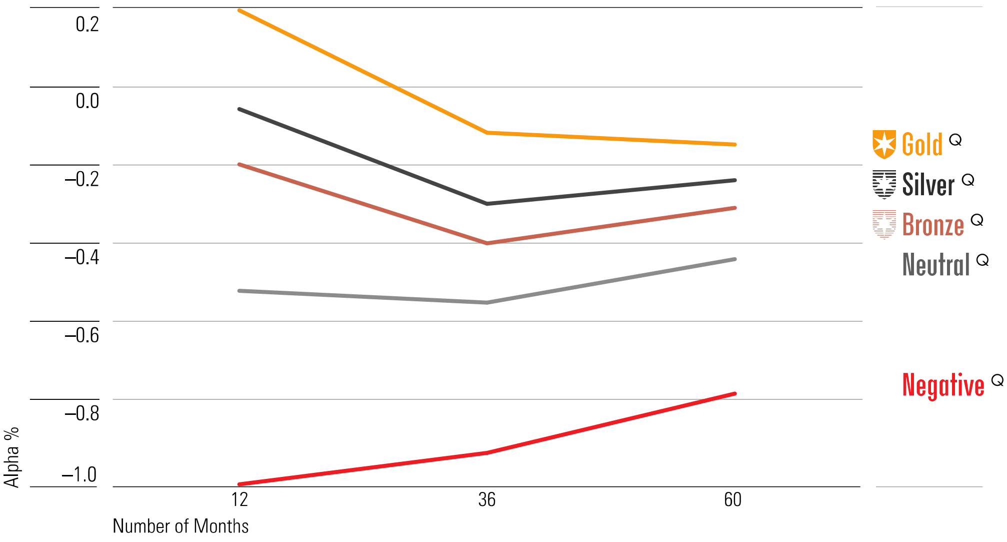
Source: Morningstar, Inc. Data as of May 2017.

Performance

From a performance standpoint, the Quantitative Rating has historically provided highly significant, powerful predictions of future fund alpha and Sharpe ratios, as estimated from a variety of risk-models. The higher a fund is rated, the better its future performance over one-, three-, and five-year time horizons. For example, improving to a Gold rating from a Negative rating is associated with an increase in the average 36-month forward alpha by 0.81% annualized. Moreover, future performance of the funds rated under the Quantitative Rating is monotonic across the rating deciles, implying that there is even more valuable information to be gained from this system by getting more granular.

In Exhibit 5 below, we present this data showcasing that the Morningstar Quantitative Rating™ for funds methodology predicts out-of-sample CAPM alpha relative to the category benchmark net of fees with clear monotonicity over one-, three-, and five-year periods. The time period tested was January 2003 to December 2016. Returns are in U.S. dollars.

Exhibit 5 Morningstar Quantitative Rating™ for Funds Out-of-Sample Average CAPM Alpha



Source: Morningstar, Inc. Data as of Feb. 28, 2017.

For more performance testing results, please see Appendix D.

Stability

Finally, we see that the Quantitative Ratings are quite stable through time. When a fund receives a Negative rating, we would expect that it has only a 0.39% probability of receiving a Bronze rating a year later and a 0.09% probability of receiving a Gold or Silver rating. Similarly, when a fund receives a Gold rating, we would expect that the fund has only a 4.58% probability of receiving a Neutral rating a year later and a 0.11% probability of receiving a Negative rating. In other words, we tend to stick to our guns when rating funds using the Quantitative Rating.

Exhibit 6 Morningstar Quantitative Rating™ for funds Stability Transition Matrix: 1 Month

	Time + 1 Month				
Time	Negative	Neutral	Bronze	Silver	Gold
Negative	93.32	6.68	0.00	0.00	0.00
Neutral	2.02	95.79	2.18	0.02	0.00
Bronze	0.00	7.25	88.32	4.42	0.01
Silver	0.00	0.07	6.69	90.86	2.38
Gold	0.00	0.00	0.05	5.40	94.55

%
80
60
40
20

Source: Morningstar, Inc. Data as of Feb. 28, 2017.

Exhibit 7 Morningstar Quantitative Rating™ for funds Stability Transition Matrix: 12 Months

	Time + 12 Month				
Time	Negative	Neutral	Bronze	Silver	Gold
Negative	67.31	32.20	0.39	0.08	0.01
Neutral	8.08	78.06	11.07	2.48	0.32
Bronze	1.15	30.16	44.68	21.02	2.99
Silver	0.43	11.59	23.68	48.33	15.97
Gold	0.11	4.58	6.43	25.37	63.51

%
50
30
10
5

Source: Morningstar, Inc. Data as of Feb. 28, 2017.

Conclusion

The Morningstar Quantitative Rating™ for funds is intended to be predictive of future alpha, and extensive performance studies have affirmed that it is, in fact, performing as intended. For additional details, please refer to the Morningstar Quantitative Rating™ for funds FAQ document or feel free to contact us.

We expect that, over time, we will enhance the Quantitative Rating to improve performance. We will note methodological changes in this document as they are made.

References

Morningstar Analyst Rating for Funds Methodology Document. 2011.
http://hkbeta.morningstar.com/Productdata/Methodology/analyst_rating_methodology.pdf

Morningstar Quantitative Equity Ratings Methodology. 2012.
http://corporate.morningstar.com/US/documents/MethodologyDocuments/FactSheets/QuantitativeEquityCreditRatingsMeth.pdf

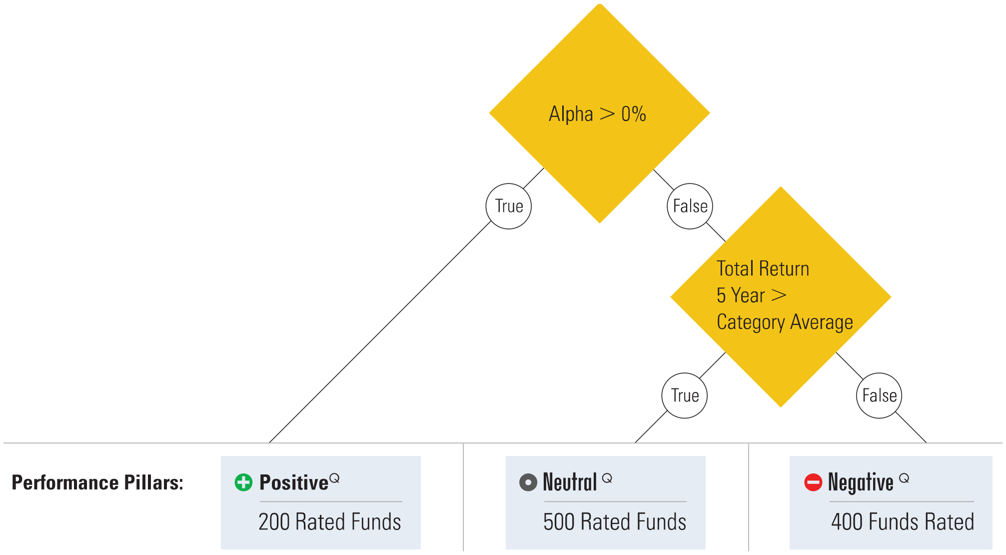
Appendix A: Random Forest

A random forest is an ensemble model, meaning its end prediction is formed based on the combination of the predictions of several submodels. In the case of a random forest, these submodels are typically regression or classification trees (hence the "forest" in "random forest"). To understand the random forest model, we must first understand how these trees are fit.

Regression Trees

A regression tree is a model based on the idea of splitting data into separate buckets based on your input variables. A visualization of a typical regression tree is shown in Exhibit 8. The tree is fit from the top down, splitting the data further into a more complex structure as you go. The end nodes contain groupings of records from your input data. Each grouping contains records that are similar to each other based on the splits that have been made in the tree.

Exhibit 8 Sample Regression Tree With Dummy Data



Source: Morningstar, Inc.

How are splits determined?

As you can see, the tree is composed of nodes that then are split until they reach terminal nodes that no longer split. Each split represents a division of our data based on a particular input variable, such as alpha, or total return five-year versus the category average (Exhibit 8). The algorithm determines where to make these splits by attempting to split our data using all possible split points for all of the input variables, and chooses the split variable and split point to maximize the difference between the variance of the unsplit data and the sum of the variances of the two groups of split data as shown in the following function.

$$VarDiff = \frac{\sum(y - \bar{y}_{presplit})^2}{N_{presplit}} - \left[\frac{\sum(y - \bar{y}_{left})^2}{N_{left}} + \frac{\sum(y - \bar{y}_{right})^2}{N_{right}}\right]$$

Intuitively, we want the split that maximizes the function because the maximizing split is the one which reduces the heterogeneity of our output variable the most. That is, the companies that are grouped on each side of the split are more similar to each other than the pre-split grouping.

A regression or classification tree will generally continue splitting until a set of user-defined conditions has been met. One of these conditions is the significance of the split. That is, if the split does not reduce heterogeneity beyond a user-defined threshold, then it will not be made. Another condition commonly used is to place a floor on the number of records in each end node. These conditions can be made more or less constrictive in order to tailor the model's bias-variance trade-off.

How are the end-node values assigned?

Each tree, once fully split, can be used to generate predictions on new data. If a new record is run through the tree, it will inevitably fall into one of the terminal nodes. The prediction for this record then becomes the arithmetic mean of the output variable for all of the training set records that fell into that terminal node.

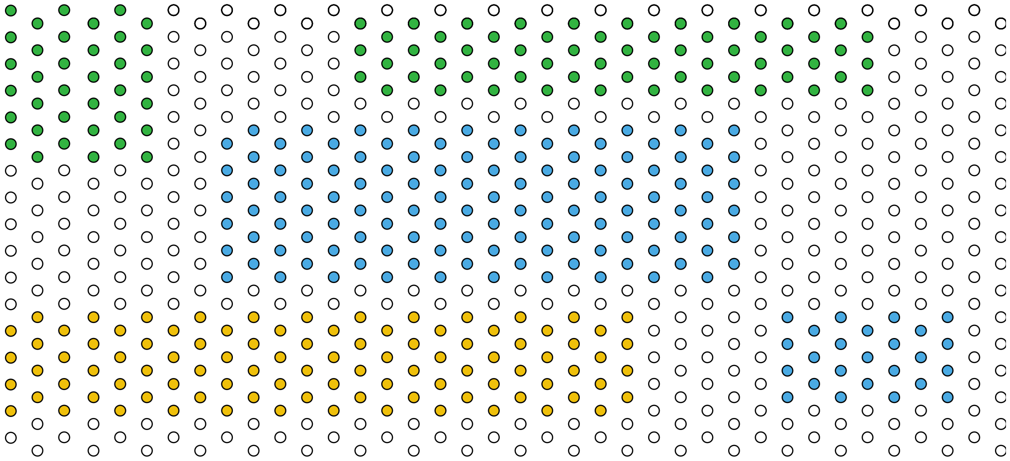
Aggregating the Trees

Now that we understand how trees are fit and how they can generate predictions, we can move further in our understanding of random forests. To arrive at an end prediction from a random forest, we first fit N trees (where N can be whatever number desired—in practice, 100 to 500 are common values) and we run our input variables through each of the N trees to arrive at N individual predictions. From there, we take the simple arithmetic mean of the N predictions to arrive at the random forest's prediction.

A logical question at this point is: Why would the N trees we fit generate different predictions if we give them the same data? The answer is: They wouldn't. That's why we give each tree a different and random subset of our data for fitting purposes (this is the "random" part of "random forest"). Think of your data as represented in Exhibit 9.

Exhibit 9 Sample Random Forest Data Representation

● Random Data Subsets



Source: Morningstar, Inc.

A random forest will choose random chunks of your data, including random cross-sectional records as well as random input variables, as represented by the highlighted sections in Exhibit 9, each time it attempts to make a new split. While Exhibit 9 shows three random subsets, the actual random forest model would choose N random subsets of your data, which may overlap, and variables selected may not be adjacent. The purpose of this is to provide each of your trees with a differentiated data set, and thus a differentiated view of the world.

Ensemble models use a "wisdom of crowds" type of approach to prediction. The theory behind this approach is that many "weak learners," which are only slightly better than random at predicting your output variable, can be aggregated to form a "strong learner" so long as the weak learners are not perfectly correlated. Mathematically, combining differentiated, better-than-random, weak learners will always result in a strong learner or a better overall prediction than any of your weak learners individually. The archetypal example of this technique is when a group of individuals is asked to estimate the number of jelly beans in a large jar. Typically the average of a large group of guesses is more accurate than a large percentage of the individual guesses.

Random forests can also be used for classification tasks. They are largely the same as described in this appendix except for the following changes: Slightly different rules are used for the splitting of nodes in the individual tree models (Gini coefficient or information gain), and the predictor variable is a binary 0 or 1 rather than a continuous variable. This means that the end predictions of a random forest for classification purposes can be interpreted as a probability of being a member of the class designated as "1" in your data.

Appendix B: Pillar—Quantitative Models

Quantitative Parent Pillar Model

What are the Quantitative Parent Pillar threshold values?

In setting threshold values for the Parent Pillar, most Negative funds fell below 0.25 and most Positive funds fell above 0.75. Therefore, for the Parent Pillar, thresholds were assigned corresponding to Negative (0 to 0.25), Neutral [0.25 to 0.75), and Positive [0.75 to 1.0). A list of variables used in each of the random forest models (Positive and Negative) is below.

What variables are used in each of the random forest models (Positive and Negative)?

The variables are used in each model are below:

Exhibit 10 Input Variables for the Quantitative Parent Pillar Rating's Positive and Negative Random Forest Models

Parent	⊕ Positive^{Q}	⊖ Negative^{Q}
Asset Weighted Manager Tenure	Yes	Yes
Average Actual Management Fee Rank	Yes	Yes
Average Net Expense Ratio Rank	Yes	Yes
Average Max Management Fee Rank	Yes	Yes
Average Prospectus Operating Expense Ratio Rank	Yes	Yes
Average Manager Tenure	Yes	Yes
Average Morningstar Rating 3 Year	Yes	Yes
Average Morningstar Rating 5 Year	Yes	Yes
Average Morningstar Rating 10 Year	Yes	Yes
Average Morningstar Rating Overall	Yes	Yes
Retention 5 Year	Yes	Yes
Risk Adjusted Success Ratio 3 Year	Yes	Yes
Risk Adjusted Success Ratio 5 Year	Yes	Yes
Risk Adjusted Success Ratio 10 Year	Yes	Yes
Success Ratio 3 Year	Yes	Yes
Success Ratio 5 Year	Yes	Yes
Success Ratio 10 Year	Yes	Yes

Source: Morningstar, Inc.

How important is each of the variables in the model?

A summary of the most important variables for the Quantitative Parent Pillar model is below along with two estimates of their importance for classification. Darker shades of blue indicate that the variable is more important according to the two estimates of variable importance, whereas lighter shades of blue indicate the variable is less important.

Exhibit 11 Input Variable Importance for the Quantitative Parent Pillar Model

Input Variable	Decrease Accuracy	Decrease Gini
Asset Weighted Manager Tenure	63.2	154.5
Average Actual Management Fee Rank	61.5	238.7
Average Net Expense Ratio Rank	50.2	183.4
Average Max Management Fee Rank	48.9	203.0
Average Prospectus Operating Expense Ratio Rank	57.8	239.3
Average Manager Tenure	61.7	194.7
Average Morningstar Rating 3 Year	38.5	131.2
Average Morningstar Rating 5 Year	34.6	143.9
Average Morningstar Rating 10 Year	44.5	195.7
Average Morningstar Rating Overall	48.6	211.9
Retention 5 Year	74.5	245.5
Risk Adjusted Success Ratio 3 Year	33.6	103.8
Risk Adjusted Success Ratio 5 Year	40.9	256.4
Risk Adjusted Success Ratio 10 Year	44.0	171.1
Success Ratio 3 Year	36.6	86.8
Success Ratio 5 Year	39.9	170.7
Success Ratio 10 Year	50.1	202.7

60
50
45
40

215
200
180
120

Source: Morningstar, Inc.

Quantitative People Pillar

What are the Quantitative People Pillar threshold values?

In setting threshold values for the People Pillar, most Negative funds fell below 0.25 and most Positive funds fell above 0.75. Therefore, for the People pillar, thresholds were assigned corresponding to Negative (0 to 0.25), Neutral [0.25 to 0.75), and Positive [0.75 to 1.0).

What variables are used in each of the random forest models (Positive and Negative)?

The variables are used in each model are below:

Exhibit 12 Input Variables for the Quantitative People Pillar for Positive and Negative Random Forest Models

Parent	⊕ Positive^Q	⊖ Negative^Q
Actual Management Fee Rank	Yes	Yes
Alpha 5 Year—Category Average	Yes	Yes
Alpha 10 Year—Category Average	Yes	No
Alpha 5 Year—Category Benchmark	No	Yes
Asset Weighted Manager Tenure	Yes	No
Average Manager Tenure	Yes	No
Average Morningstar Rating 5 Year	No	Yes
Average Morningstar Rating 10 Year	Yes	No
Average Morningstar Rating Overall	No	Yes
Average Tenure	Yes	Yes
Information Ratio 5 Year—Category Average	Yes	Yes
Longest Tenure	Yes	Yes
Max Management Fee Rank	Yes	Yes
Net Expense Ratio Rank	Yes	Yes
Retention 5 Year	Yes	Yes
Sharpe Ratio 5 Year—Category Average	No	Yes
Success Ratio 5 Year	Yes	No

Source: Morningstar, Inc.

How important is each of the variables in the model?

A summary of the most important variables for the Quantitative People Pillar model is below along with two estimates of their importance for classification. Darker shades of blue indicate that the variable is more important according to the two estimates of variable importance, whereas lighter shades of blue indicate the variable is less important.

Exhibit 13 Input Variable Importance for the Quantitative People Pillar Model

Input Variable	Decrease Accuracy	Decrease Gini
Actual Management Fee Rank	58.1	167.2
Alpha 5 Year—Category Average	66.4	279.6
Alpha 10 Year—Category Average	46.9	96.8
Asset Weighted Manager Tenure	36.5	73.1
Average Manager Tenure	36.0	52.0
Average Morningstar Rating 10 Year	51.4	119.9
Average Tenure	77.6	282.7
Information Ratio 5 Year—Category Average	46.4	133.7
Longest Tenure	53.3	197.8
Max Management Fee Rank	60.8	157.2
Net Expense Ratio Rank	51.4	86.4
Retention 5 Year	87.5	405.7
Success Ratio 5 Year	33.5	74.5
Alpha 5 Year—Category Benchmark	18.0	77.4
Average Morningstar Rating 5 Year	22.6	180.9
Average Morningstar Rating Overall	19.7	130.7
Sharpe Ratio 5 Year—Category Average	9.5	102.8

60	215
50	150
40	100
20	75

Source: Morningstar, Inc.

Quantitative Performance Pillar Model

What are the Quantitative Performance Pillar threshold values?

In setting threshold values for the Performance Pillar, most Negative funds fell below 0.25 and most Positive funds fell above 0.75. Therefore, for the Performance Pillar, thresholds were assigned corresponding to Negative (0 to 0.25), Neutral [0.25 to 0.75), and Positive [0.75 to 1.0).

What variables are used in each of the random forest models (Positive and Negative)?

The variables are used in each model are below:

Exhibit 14 Input Variables for the Quantitative Performance Pillar for Positive and Negative Random Forest Models

Performance	⊕ Positive^{Q}	⊖ Negative^{Q}
Alpha 5 Year—Category Average	Yes	Yes
Alpha 10 Year—Category Average	Yes	Yes
Alpha 5 Year—Category Benchmark	Yes	Yes
Average Morningstar Rating 10 Year	No	Yes
Average Tenure	Yes	Yes
Beta 3 Year—Category Average	No	Yes
Down Capture 3 Year—Category Average	No	Yes
Information Ratio 5 Year—Category Average	Yes	No
Information Ratio 10 Year—Category Average	Yes	Yes
Information Ratio 5 Year—Category Benchmark	Yes	Yes
Longest Tenure	Yes	No
Morningstar Rating for 3 Year	No	Yes
Morningstar Rating for 5 Year	No	Yes
Overall Morningstar Rating	Yes	Yes
Trailing 5 Year Return Rank	Yes	Yes
Trailing 10 Yr Return Rank	Yes	No

Source: Morningstar, Inc.

How important is each of the variables in the model?

A summary of the most important variables for the Quantitative Performance Pillar model is below along with two estimates of their importance for classification. Darker shades of blue indicate that the variable is more important according to the two estimates of variable importance, whereas lighter shades of blue indicate the variable is less important.

Exhibit 15 Input Variable Importance for the Quantitative Performance Pillar Model

Input Variable	Decrease Accuracy	Decrease Gini
Alpha 5 Year—Category Average	55.7	228.4
Alpha 10 Year—Category Average	45.7	78.1
Alpha 5 Year—Category Benchmark	53.0	142.8
Average Tenure	82.5	247.7
Information Ratio 5 Year—Category Average	35.5	77.9
Information Ratio 10 Year—Category Average	60.6	160.3
Information Ratio 5 Year—Category Benchmark	46.8	179.3
Longest Tenure	49.1	89.4
Overall Morningstar Rating	87.9	471.7
Trailing 5 Year Return Rank	47.0	248.4
Trailing 10 Year Return Rank	48.9	97.2
Average Morningstar Rating 10 Yr	57.1	205.5
Beta 3 Yr - Category Average	11.5	48.3
Down Capture 3 Year—Category Average	11.7	48.4
Morningstar Rating for 3 Year	19.3	61.1
Morningstar Rating for 5 Year	12.6	77.4

60	215
50	150
40	100
20	75

Source: Morningstar, Inc.

Quantitative Price Pillar Model

What are the Quantitative Price Pillar threshold values?

In setting threshold values for the Price Pillar, most Negative funds fell below 0.25 and most Positive funds fell above 0.75. Therefore, for the Price Pillar, thresholds were assigned corresponding to Negative (0 to 0.25), Neutral [0.25 to 0.75), and Positive [0.75 to 1.0).

What variables are used in each of the random forest models (Positive and Negative)?

The variables are used in each model are below:

Exhibit 16 Input Variables for the Quantitative Price Pillar for Positive and Negative Random Forest Models

Performance	⊕ Positive^{Q}	⊖ Negative^{Q}
Acquired Expense Ratio Rank	Yes	Yes
Actual Management Fee Rank	Yes	Yes
Max Management Fee Rank	Yes	Yes
Net Expense Ratio Rank	Yes	Yes
Prospectus Operating Expense Ratio Rank	Yes	Yes
Prospectus Net Expense Ratio Rank	Yes	Yes

Source: Morningstar, Inc.

How important is each of the variables in the model?

A summary of the most important variables for the Quantitative Price Pillar model is below along with two estimates of their importance for classification. Darker shades of blue indicate that the variable is more important according to the two estimates of variable importance, whereas lighter shades of blue indicate the variable is less important.

Exhibit 17 Input Variable Importance for the Quantitative Price Pillar Model

Input Variable	Decrease Accuracy	Decrease Gini
Acquired Expense Ratio Rank	92.5	128.0
Actual Management Fee Rank	86.1	193.9
Max Management Fee Rank	104.3	156.2
Net Expense Ratio Rank	266.7	1,329.9
Prospectus Net Expense Ratio Rank	97.3	334.5
Prospectus Operating Expense Ratio Rank	87.2	661.8

150
100
95
90

700
500
200
150

Source: Morningstar, Inc.

Quantitative Process Pillar Model

What are the Quantitative Process Pillar threshold values?

In setting threshold values for the Process Pillar, most Negative funds fell below 0.25 and most Positive funds fell above 0.75. Therefore, for the Process Pillar, thresholds were assigned corresponding to Negative (0 to 0.25), Neutral [0.25 to 0.75), and Positive [0.75 to 1.0).

What variables are used in each of the random forest models (Positive and Negative)?

The variables are used in each model are below:

Exhibit 18 Input Variables for the Quantitative Process Pillar for Positive and Negative Random Forest Models

Process	⊕ Positive^{Q}	⊖ Negative^{Q}
Alpha 5 Year—Category Average	Yes	Yes
Alpha 10 Year—Category Average	Yes	No
Alpha 5 Year—Category Benchmark	Yes	Yes
% Assets in Top 10 Holdings	No	Yes
Asset Weighted Manager Tenure	Yes	No
Average Morningstar Rating 10 Year	Yes	No
Average Morningstar Rating Overall	No	Yes
Average Net Expense Ratio Rank	No	Yes
Average Actual Management Fee Rank	No	Yes
Average Max Management Fee Rank	No	Yes
Average Prospectus Operating Expense Ratio Rank	No	Yes
Average Tenure	Yes	Yes
Information Ratio 5 Year—Category Average	No	Yes
Information Ratio 5 Year—Category Benchmark	Yes	No
Longest Tenure	Yes	Yes
Max Management Fee Rank	No	Yes
Net Expense Ratio Rank	No	Yes
Number of Holdings	Yes	Yes
Retention 5 Year	Yes	Yes
Risk Adjusted Success Ratio 5 Year	Yes	No
Trailing 5 Year Return Rank	Yes	No

Source: Morningstar, Inc.

How important is each of the variables in the model?

A summary of the most important variables for the Quantitative Process Pillar model is below along with two estimates of their importance for classification. Darker shades of blue indicate that the variable is more important according to the two estimates of variable importance, whereas lighter shades of blue indicate the variable is less important.

Exhibit 19 Input Variable Importance for the Quantitative Process Pillar Model

Input Variable	Decrease Accuracy	Decrease Gini
Alpha 5 Year—Category Average	50.9	211.9
Alpha 10 Year—Category Average	46.1	95.2
Alpha 5 Year—Category Benchmark	59.6	180.0
Asset Weighted Manager Tenure	43.6	69.5
Average Morningstar Rating 10 Yr	56.1	123.1
Average Tenure	61.7	161.3
Information Ratio 5 Year—Category Benchmark	30.3	61.5
Longest Tenure	66.5	155.5
Number of Holdings	81.7	
ank	12.9	81.4
Average Prospectus Operating Expense Ratio Rank	17.6	132.4
Information Ratio 5 Year—Category Average	12.1	61.2
Max Management Fee Rank	26.0	102.2
Net Expense Ratio Rank	15.2	43.7

Decrease Accuracy: 70, 50, 30, 15

Decrease Gini: 150, 125, 100, 75

Source: Morningstar, Inc.

Appendix C: Input Data FAQ

How do we normalize the input data?

After all data is calculated and collected, we cross-sectionally normalize the data by region to be mean zero and standard deviation 1. This puts everything into the same units (in terms of standard deviation), which makes the data a bit easier to interpret.

How do we assign regions?

In order to normalize by region, we need to know what funds belong to what regions. Countries are assigned to regions based on the Morningstar Region classification system. We assign funds to regions based on the fund's domicile, unless the fund's domicile is not contained within the set of Available for Sale countries. In that case, we choose an Available for Sale country depending on which of those countries belongs to the domicile with the most industrywide assets (for example, U.S. > emerging-markets Asia).

How do we handle missing data?

In the case of missing data, we cross-sectionally impute the median value of the region to which the fund is assigned. We use region-level imputation, as opposed to category-level, because we want to have a relatively broad sample of funds on which to draw. Sometimes imputing the median value in the place of missing data can be harmful, especially when you need to calculate an average (for example, a regression), but in this case we believe that we are justified since the random forest algorithm splits data based on percentiles in the distribution and does not require us to reliably estimate moments. Imputed values will be treated as "average" and hence likely to pull the final ratings decisions toward Neutral. We think that, in the absence of any data, the average fund should probably be Neutral and would be the stance that an analyst would take a priori before any data about the fund was presented to them.

How do we handle category changes?

Input data reflects information available at a given time. Therefore, historical data incorporates the fund's historical category. For performance-related metrics where we require a time series of a fund's category average performance or category index return, we use the monthly track record reflecting the fund's category for that specific month.

What data points are category-relative?

First, most data points will be calculated relative to the category (for example, category average alpha, success ratios, return ranks, beta, fee ranks, star ratings, and so on), but some will not (for example, tenure, retention ratio, or number of holdings). We prefer to use category-relative data points where possible, but tended to refrain when the data point was more operational in nature.

What currency do we use for calculating fund performance statistics?

To estimate fund performance, we convert all fund and index returns to U.S. dollars prior to running our regressions. This eliminates any effects due to the difference in currency return.

What does "average" stand for?

Average stands for an equally weighted average of all share classes given a branding ID.

When are the input data and ratings updated?

The input data and ratings are updated on the 21st day of each month.

Why are fees taken account for people?

Here fees are directly related to how much a fund charges by managing money for clients, for two reasons. One, our model testing shows that fees do help explain the variance in the People Pillar rating. Two, fees empirically affect all pillars directly or indirectly.

Why do we use the input variables Percentage of Assets in Top 10 Holdings, and Number of Holdings, for the Process Pillar? What is the effective relationship between these variables and the pillar rating?

Percentage of Assets in Top 10 Holdings is a good indicator to measure how concentrated a fund's portfolio is. The higher the top 10 asset percentage, the more concentrated the portfolio. Such portfolios are implicitly taking on higher risk. Number of Holdings reflects both directions of the concentration of a portfolio—is it very concentrated, or over-diversified? Both variables reflect a fund’s investment philosophy and actual investment process.

Appendix D: Performance of the Morningstar Quantitative Rating™ for funds

Have we performed any testing on the Morningstar Quantitative Rating?

Yes, many tests. The Morningstar Quantitative Rating methodology went through a four-year vetting period. The resultant methodology has proven to be one of the most successful ratings systems Morningstar has developed on a variety of tests. The following questions describe the tests performed.

How accurate are the Morningstar Quantitative Ratings compared with the Analyst Ratings?

Approximately 95% of funds lie within 1 rating of true rating. See Exhibits 4 for the distribution of differences.

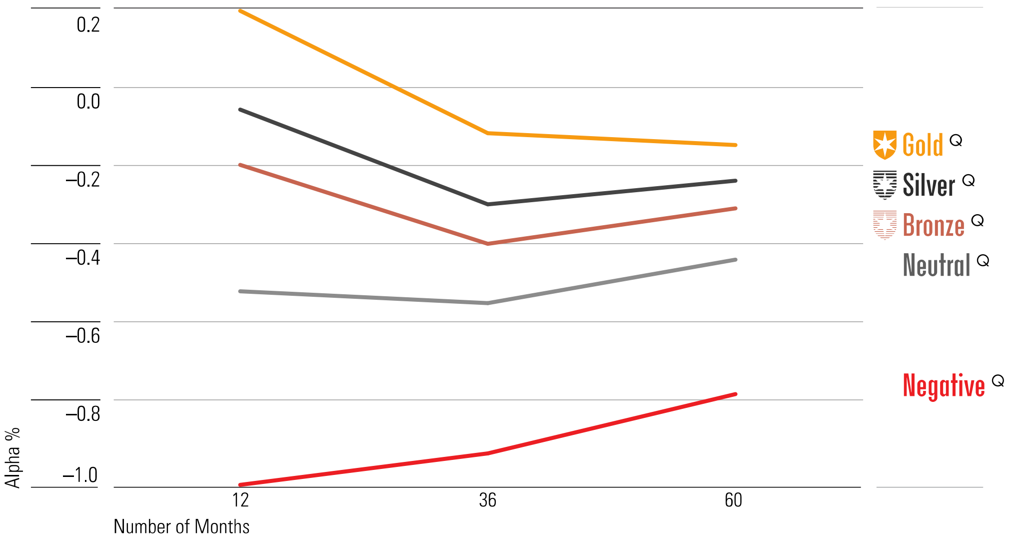
Is the Morningstar Quantitative Rating stable?

Yes, the ratings are highly stable over time. Gold-rated funds maintain their Morningstar Medalist rating one year later with 98.9% frequency. Overall, Morningstar Medalists remain as Medalists with 80.9% frequency one year later. Exhibit 6 shows the average movements within one month. In Exhibit 7, we show the average movements over a one-year time horizon.

How well does the Morningstar Quantitative Rating predict out-of-sample CAPM alpha relative to the category average?

The Morningstar Quantitative Rating methodology predicts out-of-sample CAPM alpha relative to the category benchmark net of fees with clear monotonicity over one-, three-, and five-year periods. See Exhibit 20 for results. The time period tested was January 2003 to February 2017. Returns are in U.S. dollars.

Exhibit 20 Morningstar Quantitative Rating Out-of-Sample Average CAPM Alpha

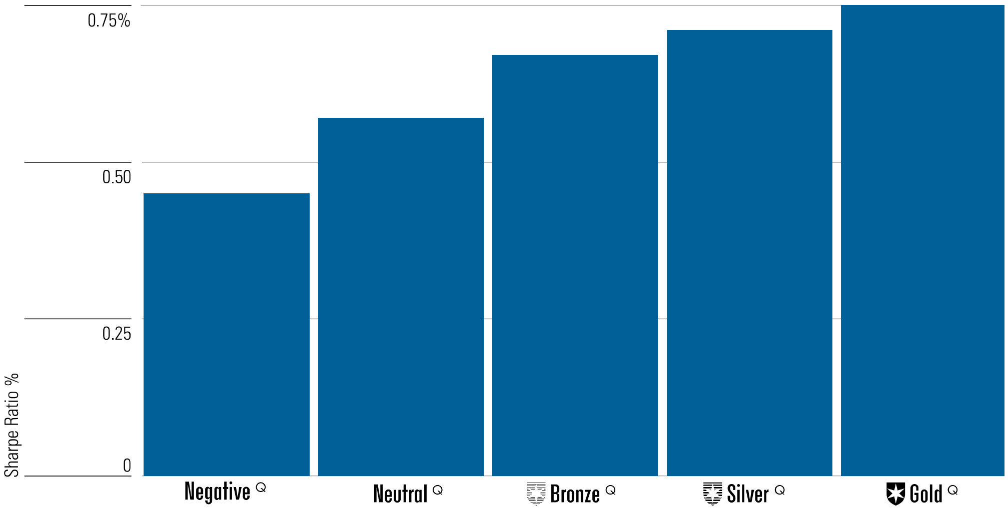


Source: Morningstar, Inc. Data as of Feb. 28, 2017

How well does the Morningstar Quantitative Rating predict out-of-sample Sharpe ratio relative to the category average?

The Morningstar Quantitative Rating methodology predicts out-of-sample Sharpe ratio relative to the category average with clear monotonicity. See Exhibit 21 for results. The time period tested was January 2003 to February 2017. Returns are in U.S. dollars.

Exhibit 21 Morningstar Quantitative Rating Out-of-Sample Sharpe Ratio

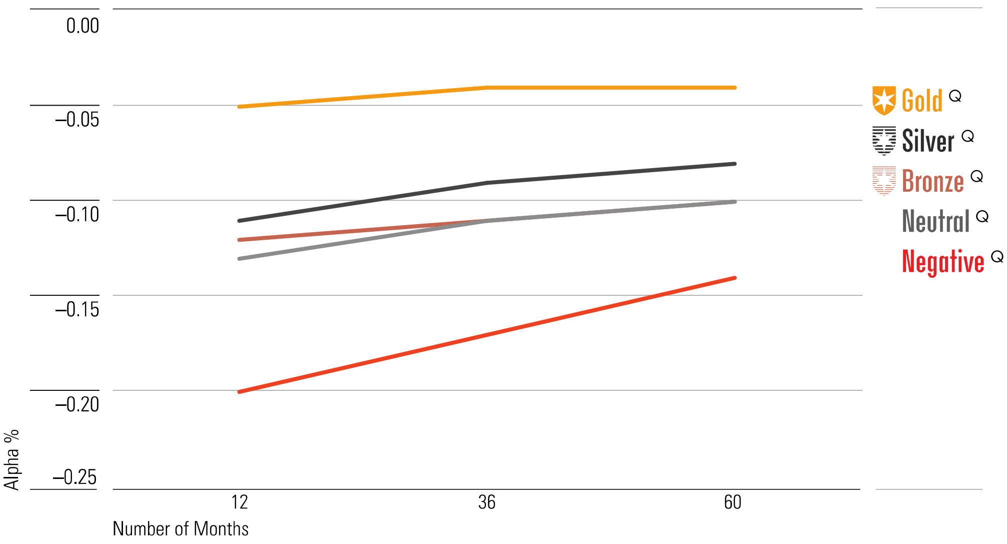


Source: Morningstar, Inc. Data as of Feb. 28, 2017

How well does the Morningstar Quantitative Rating predict out-of-sample Carhart alpha for equity funds?

The Morningstar Quantitative Rating methodology predicts out-of-sample Carhart alpha for equity funds. Alpha is estimated net of fees. Higher rated funds tend to have higher alphas. See Exhibit 22 for results. The time period tested was January 2003 to February 2017. Returns are in U.S. dollars.

Exhibit 22 Morningstar Quantitative Rating Out -of-Sample Average Carhart Alpha

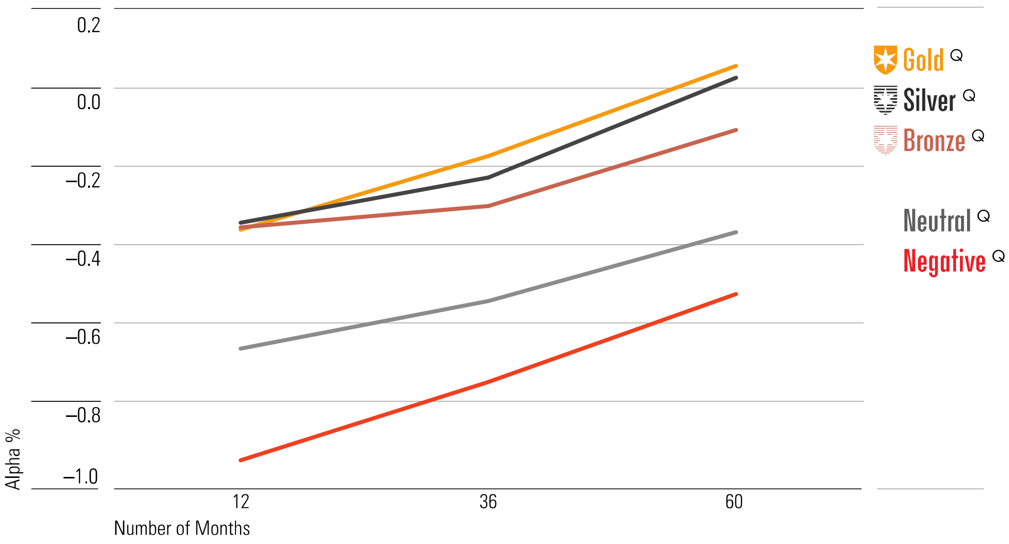


Source: Morningstar, Inc. Data as of Feb. 28, 2017

How well does the Morningstar Quantitative Rating predict out-of-sample Fama-French 5-factor alpha for fixed-income funds?

The Morningstar Quantitative Rating methodology predicts out-of-sample and Fama-French 5-factor alpha for fixed-income funds. Alpha is estimated net of fees. Higher rated funds tend to have higher alphas. See Exhibit 23 for results. The time period tested was January 2003 to February 2017. Returns are in U.S. dollars.

Exhibit 23 Morningstar Quantitative Rating Out -of-Sample Average Fama-French 5-Factor Alpha



Source: Morningstar, Inc. Data as of Feb. 28, 2017

How well does the Morningstar Quantitative Rating predict out-of-sample CAPM alpha after controlling for fees?

After controlling for fees, we find a strong relationship between the Morningstar Quantitative Rating and subsequent CAPM alpha. We performed a double sort on funds: first by the Quantitative Rating and then by fee quintile. For each subgroup, we averaged the forward 12-, 36-, and 60-month CAPM alpha. As expected, as fees decrease, average CAPM alpha increases. We also find that higher ratings are correlated with higher CAPM alphas. See Exhibits 24, 25, and 26 for results. The time period tested was January 2003 to February 2017. Returns are in U.S. dollars.

Exhibit 24 12-Mo CAPM Alpha % — Double Sort on Fees and Morningstar Quantitative Rating

	Fee Quintile Worst				Best	
MQR	▤	▤	▤	▤	▤	Average
Negative Q	–1.82	–1.15	–0.80	–0.67	–0.50	–0.99
Neutral Q	–0.93	–0.62	–0.51	–0.33	–0.11	–0.50
Bronze Q	–0.36	–0.36	–0.14	–0.07	–0.02	–0.19
Silver Q	–0.26	–0.07	0.02	0.10	–0.06	–0.06
Gold Q	0.26	0.25	0.57	–0.09	–0.03	0.19
Average	–0.62	–0.39	–0.17	–0.21	–0.14	—

%
0.25
0.00
–0.25
–0.75

Source: Morningstar, Inc.

Exhibit 25 36-Mo CAPM Alpha % — Double Sort on Fees and Morningstar Quantitative Rating

	Fee Quintile Worst				Best	
MQR	▤	▤	▤	▤	▤	Average
Negative Q	–1.74	–1.16	–0.69	–0.54	–0.46	–0.92
Neutral Q	–0.85	–0.64	–0.54	–0.48	–0.20	–0.54
Bronze Q	–0.53	–0.49	–0.41	–0.31	–0.18	–0.38
Silver Q	–0.44	–0.32	–0.26	–0.17	–0.25	–0.29
Gold Q	0.05	–0.08	0.18	–0.46	–0.23	–0.11
Average	–0.70	–0.54	–0.34	–0.39	–0.26	—

%
0.00
–0.25
–0.50
–1.00

Source: Morningstar, Inc.

Exhibit 26 60-Mo CAPM Alpha %—Double Sort on Fees and Morningstar Quantitative Rating

	Fee Quintile					
	Worst				Best	
MQR						Average
Negative Q	–1.22	–1.07	–0.77	–0.46	–0.33	–0.77
Neutral Q	–0.58	–0.54	–0.43	–0.40	–0.21	–0.43
Bronze Q	–0.29	–0.34	–0.34	–0.31	–0.23	–0.30
Silver Q	–0.22	–0.14	–0.22	–0.32	–0.25	–0.23
Gold Q	0.06	0.01	0.01	–0.52	–0.28	–0.14
Average	–0.45	–0.42	–0.35	–0.40	–0.26	—

%
0.00
–0.25
–0.50
–1.00

Source: Morningstar, Inc.

How well does the Morningstar Quantitative Rating predict out-of-sample Carhart alpha after controlling for fees?

After controlling for fees, we find a strong relationship between the Morningstar Quantitative Rating and subsequent Carhart alpha. We performed a double sort on funds: first by fee decile and second by the Quantitative Rating decile. For each subgroup, we averaged the trailing 12-, 36-, and 60-month Carhart alpha. As expected, as fees decrease, average Carhart alpha increases. See Exhibits 27, 28, and 29 for results. The time period tested was January 2003 to February 2017. Returns are in U.S. dollars. Only equity funds are included in the sample.

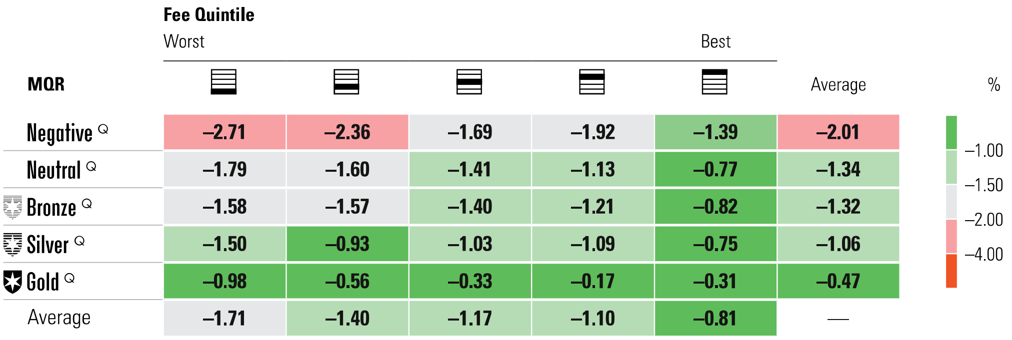
Exhibit 27 12-Mo Carhart Alpha %—Double Sort on Fees and Morningstar Quantitative Rating

	Fee Quintile					
	Worst				Best	
MQR						Average
Negative Q	–3.25	–2.58	–2.10	–2.17	–1.76	–2.37
Neutral Q	–2.21	–1.95	–1.65	–1.22	–0.86	–1.58
Bronze Q	–1.90	–1.77	–1.36	–1.30	–0.82	–1.43
Silver Q	–1.94	–1.19	–1.29	–1.33	–0.87	–1.32
Gold Q	–1.52	–0.62	–0.39	–0.34	–0.26	–0.63
Average	–2.16	–1.62	–1.36	–1.27	–0.91	—

%
–0.50
–1.00
–1.50
–2.00

Source: Morningstar, Inc.

Exhibit 28 36-Mo Carhart Alpha % — Double Sort on Fees and Morningstar Quantitative Rating

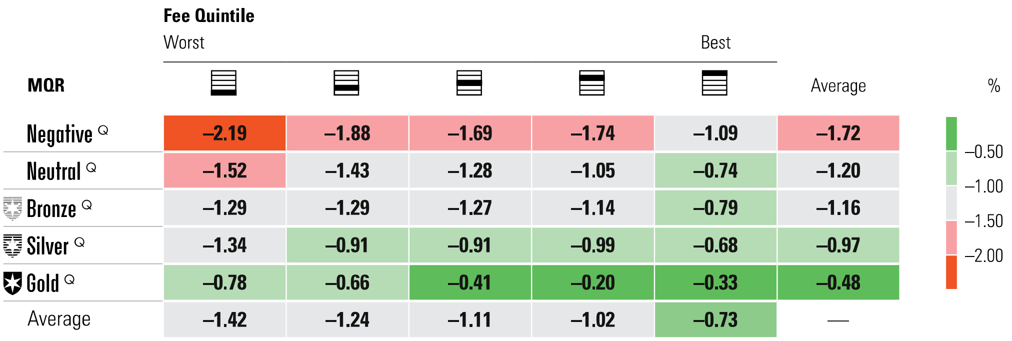


Fee Quintile

MQR	Worst				Best	Average
Negative Q	**–2.71**	**–2.36**	**–1.69**	**–1.92**	**–1.39**	**–2.01**
Neutral Q	**–1.79**	**–1.60**	**–1.41**	**–1.13**	**–0.77**	**–1.34**
Bronze Q	**–1.58**	**–1.57**	**–1.40**	**–1.21**	**–0.82**	**–1.32**
Silver Q	**–1.50**	**–0.93**	**–1.03**	**–1.09**	**–0.75**	**–1.06**
Gold Q	**–0.98**	**–0.56**	**–0.33**	**–0.17**	**–0.31**	**–0.47**
Average	**–1.71**	**–1.40**	**–1.17**	**–1.10**	**–0.81**	—

Source: Morningstar, Inc.

Exhibit 29 60-Mo Carhart Alpha % — Double Sort on Fees and Morningstar Quantitative Rating



Fee Quintile

MQR	Worst				Best	Average
Negative Q	**–2.19**	**–1.88**	**–1.69**	**–1.74**	**–1.09**	**–1.72**
Neutral Q	**–1.52**	**–1.43**	**–1.28**	**–1.05**	**–0.74**	**–1.20**
Bronze Q	**–1.29**	**–1.29**	**–1.27**	**–1.14**	**–0.79**	**–1.16**
Silver Q	**–1.34**	**–0.91**	**–0.91**	**–0.99**	**–0.68**	**–0.97**
Gold Q	**–0.78**	**–0.66**	**–0.41**	**–0.20**	**–0.33**	**–0.48**
Average	**–1.42**	**–1.24**	**–1.11**	**–1.02**	**–0.73**	—

Source: Morningstar, Inc.

How well does the Morningstar Quantitative Rating predict out-of-sample Fama-French 5-factor alpha after controlling for fees?

After controlling for fees, we find a strong relationship between the Morningstar Quantitative Rating and subsequent Fama-French 5-factor alpha. We performed a double sort on funds: first by fee decile and second by the Quantitative Rating decile. For each subgroup, we averaged the trailing 12-, 36-, and 60-month Fama-French 5-factor alpha. As expected, as fees decrease, average Fama-French 5-factor alpha increases. See Exhibits 30, 31, and 32 for results. The time period tested was January 2003 to February 2017. Returns are in U.S. dollars. Only fixed-income funds are included in the sample.

Exhibit 30 12-Mo Fama-French 5-Factor Alpha %—Double Sort on Fees and Morningstar Quantitative Rating

	Fee Quintile					
	Worst				Best	
MQR						Average
Negative Q	–3.08	–1.79	–2.06	–1.76	–0.78	–1.89
Neutral Q	–2.31	–1.89	–1.77	–0.70	–0.42	–1.42
Bronze Q	–1.99	–1.30	–0.70	–0.30	–0.26	–0.91
Silver Q	–2.12	–1.44	–0.59	–0.08	–0.20	–0.89
Gold Q	–2.05	–1.31	–0.56	–0.35	–0.34	–0.92
Average	–2.31	–1.55	–1.14	–0.64	–0.40	—

%
–0.50
–1.00
–1.50
–2.00

Source: Morningstar, Inc.

Exhibit 31 36-Mo Fama-French 5-Factor Alpha %—Double Sort on Fees and Morningstar Quantitative Rating

	Fee Quintile					
	Worst				Best	
MQR						Average
Negative Q	–2.52	–1.72	–1.62	–1.41	–0.54	–1.56
Neutral Q	–2.04	–1.70	–1.39	–0.66	–0.29	–1.21
Bronze Q	–1.75	–1.19	–0.64	–0.41	–0.10	–0.82
Silver Q	–1.64	–1.09	–0.45	–0.16	–0.16	–0.70
Gold Q	–1.49	–0.92	–0.50	–0.12	–0.04	–0.61
Average	–1.89	–1.32	–0.92	–0.55	–0.22	—

%
–0.50
–1.00
–1.50
–2.00

Source: Morningstar, Inc.

Exhibit 32 60-Mo Fama-French 5-Factor Alpha %—Double Sort on Fees and Morningstar Quantitative Rating

	Fee Quintile Worst				Best	
MQR	▤	▤	▤	▤	▤	Average
Negative Q	–2.29	–1.46	–1.06	–0.98	–0.15	–1.19
Neutral Q	–1.82	–1.36	–1.09	–0.33	–0.01	–0.92
Bronze Q	–1.50	–0.93	–0.36	0.04	0.25	–0.50
Silver Q	–1.15	–0.50	–0.10	0.19	0.18	–0.28
Gold Q	–1.17	–0.30	–0.16	0.31	0.18	–0.23
Average	–1.59	–0.91	–0.56	–0.16	0.09	—

%
🟩
–0.50
🟩
–1.00
⬜
–1.50
🟥
–2.00
🟧

Source: Morningstar, Inc.

How well does the Morningstar Quantitative Rating predict out-of-sample CAPM alpha after controlling for the Morningstar Analyst Rating?

After controlling for Analyst Rating, we find a strong relationship between the Morningstar Quantitative Rating and subsequent CAPM Alpha. See Exhibits 33, 34, and 35 for results. The time period tested was January 2003 to February 2017. Returns are in U.S. dollars. All asset classes are included in the study.

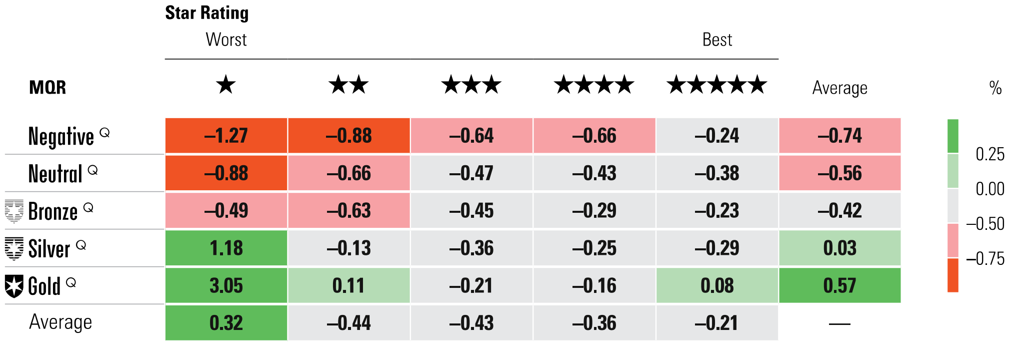
Exhibit 33 12-Mo CAPM Alpha %—Double Sort on Morningstar Quantitative Rating and Analyst Rating

	Star Rating Worst				Best	
MQR	★	★★	★★★	★★★★	★★★★★	Average
Negative Q	–1.20	–1.01	–0.75	–0.98	–1.01	–0.99
Neutral Q	–1.48	–0.66	–0.39	–0.19	–0.05	–0.55
Bronze Q	–0.98	–0.54	–0.33	0.01	0.07	–0.36
Silver Q	0.22	0.61	–0.19	–0.03	0.04	0.13
Gold Q	0.87	–0.01	0.15	0.11	0.42	0.31
Average	–0.51	–0.32	–0.30	–0.21	–0.11	—

%
🟩
0.25
🟩
0.00
⬜
–0.50
🟥
–1.00
🟧

Source: Morningstar, Inc.

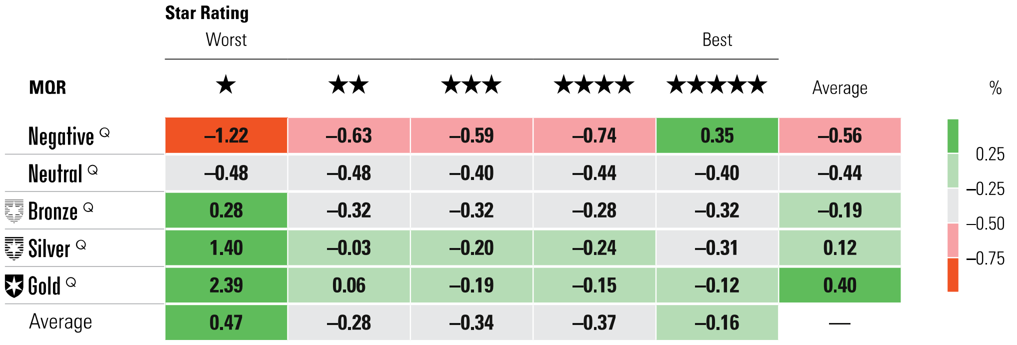
Exhibit 34 36-Mo CAPM Alpha %—Double Sort on Morningstar Quantitative Rating and Analyst Rating



	Star Rating Worst				Best	
MQR	★	★★	★★★	★★★★	★★★★★	Average
Negative Q	**–1.27**	**–0.88**	**–0.64**	**–0.66**	**–0.24**	**–0.74**
Neutral Q	**–0.88**	**–0.66**	**–0.47**	**–0.43**	**–0.38**	**–0.56**
Bronze Q	**–0.49**	**–0.63**	**–0.45**	**–0.29**	**–0.23**	**–0.42**
Silver Q	**1.18**	**–0.13**	**–0.36**	**–0.25**	**–0.29**	**0.03**
Gold Q	**3.05**	**0.11**	**–0.21**	**–0.16**	**0.08**	**0.57**
Average	**0.32**	**–0.44**	**–0.43**	**–0.36**	**–0.21**	—

Source: Morningstar, Inc.

Exhibit 35 60-Mo CAPM Alpha %—Double Sort on Morningstar Quantitative Rating



	Star Rating Worst				Best	
MQR	★	★★	★★★	★★★★	★★★★★	Average
Negative Q	**–1.22**	**–0.63**	**–0.59**	**–0.74**	**0.35**	**–0.56**
Neutral Q	**–0.48**	**–0.48**	**–0.40**	**–0.44**	**–0.40**	**–0.44**
Bronze Q	**0.28**	**–0.32**	**–0.32**	**–0.28**	**–0.32**	**–0.19**
Silver Q	**1.40**	**–0.03**	**–0.20**	**–0.24**	**–0.31**	**0.12**
Gold Q	**2.39**	**0.06**	**–0.19**	**–0.15**	**–0.12**	**0.40**
Average	**0.47**	**–0.28**	**–0.34**	**–0.37**	**–0.16**	—

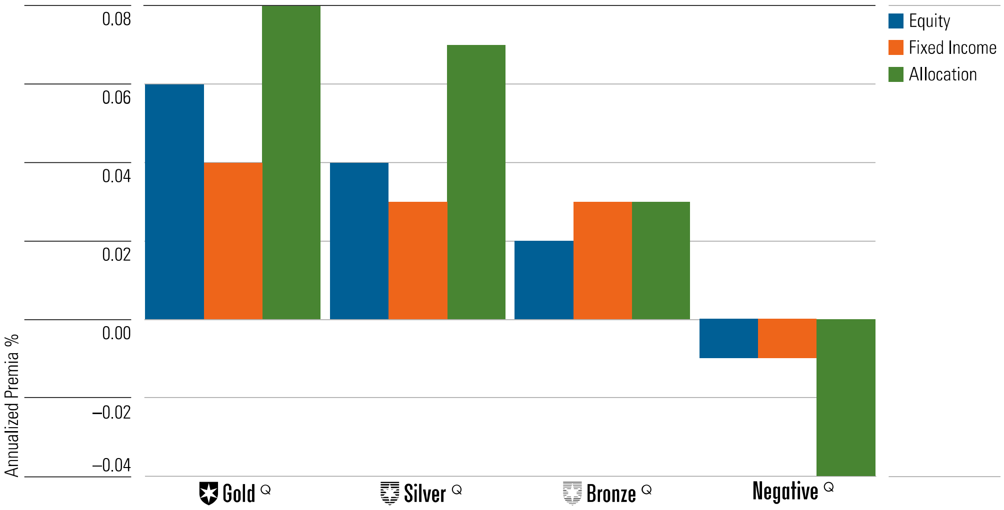
Source: Morningstar, Inc.

Is there a premium for investing in the Morningstar Quantitative Rating?

Yes. There is a positive premium for investing in the Gold-, Silver-, and Bronze-rated funds by the Morningstar Quantitative Rating. There is a negative premium for investing in the Negative-rated funds by the Quantitative Rating. These premiums exist across all three asset classes rated—equity, fixed-income, and allocation. Furthermore, for each asset class, the largest premium exists in Gold-rated funds and monotonically decreases with the ratings.

To determine the premium associated with each rating level and asset class, we ran separate Fama-MacBeth regressions from 2003-16. See Exhibit 36 for results.

Exhibit 36 Fama-MacBeth Cross-Sectional Annualized Premiums (2003-16)



Source: Morningstar, Inc. Data as of Dec. 31, 2016.

About Morningstar® Quantitative Research

Morningstar Quantitative Research is dedicated to developing innovative statistical models and data points, including the Quantitative Equity Ratings and the Global Risk Model.

For More Information

+1 312 244-7541
lee.davidson@morningstar.com

MORNINGSTAR®

22 West Washington Street
Chicago, IL 60602 USA

©2017 Morningstar. All Rights Reserved. Unless otherwise provided in a separate agreement, you may use this report only in the country in which its original distributor is based. The information, data, analyses, and opinions presented herein do not constitute investment advice; are provided solely for informational purposes and therefore are not an offer to buy or sell a security; and are not warranted to be correct, complete, or accurate. The opinions expressed are as of the date written and are subject to change without notice. Except as otherwise required by law, Morningstar shall not be responsible for any trading decisions, damages, or other losses resulting from, or related to, the information, data, analyses, or opinions or their use. References to "Morningstar Credit Ratings" refer to ratings issued by Morningstar Credit Ratings, LLC, a credit rating agency registered with the Securities and Exchange Commission as a nationally recognized statistical rating organization ("NRSRO"). Under its NRSRO registration, Morningstar Credit Ratings issues credit ratings on financial institutions (e.g., banks), corporate issuers, and asset-backed securities. While Morningstar Credit Ratings issues credit ratings on insurance companies, those ratings are not issued under its NRSRO registration. All Morningstar credit ratings and related analysis are solely statements of opinion and not statements of fact or recommendations to purchase, hold, or sell any securities or make any other investment decisions. Morningstar credit ratings and related analysis should not be considered without an understanding and review of our methodologies, disclaimers, disclosures, and other important information found at https://ratingagency.morningstar.com. The information contained herein is the proprietary property of Morningstar and may not be reproduced, in whole or in part, or used in any manner, without the prior written consent of Morningstar. To license the research, call +1 312 696-6869.