Morningstar Fund Shutdown Model
Methodology

Morningstar Quantitative Research
15 April 2019
Version 1.0

Contents

1 Overview
1 Introduction
2 Universe Construction
2 Factor Exposures
3 Model Methodology
4 Probability of Fund Shutdown
5 Conclusion
5 References

Appendix A: Factor Definitions
Appendix B: Input Data FAQ

Harshal Dalvi
Senior Data Scientist
+91 22 61217100
harshal.dalvi@morningstar.com

Som Dubey
Data Scientist
+91 22 61217100
som.dubey@morningstar.com

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

Overview

The Morningstar Quantitative Research team has been researching drivers of fund flows since the 2015 launch of the "What Factors Drive Fund Flows" paper. Subsequently, we conducted a body of work researching fund flows for new funds, how fund flows influence fund closures, and the effects of manager changes on fund flows. In 2017, we released "The Fall of Funds: Why Some Funds Fail," which explores the factors leading to a fund becoming obsolete. In total, the work analyzes fund flows during all phases of a fund's lifecycle: launch, operational periods, closure, and transitional periods. Morningstar is interested in the full lifecycle of fund flows because they represent the aggregate decisions investors make to their portfolios.

In 2018, we released our fund flow models into Morningstar's Investor Pulse tool starting with the U.S. market and expanding globally. Based off "The Fall of Funds" paper, we are implementing the probability of fund shutdown model so these insights can be gleaned in a live setting. The first fund shutdown model implemented is for the U.S. equity market.

Introduction

Morningstar generated a probability of fund shutdown model with three goals. First, determine which characteristics are more prevalent among obsolete funds relative to their surviving counterparts and identify their relative importance. Second, determine the time horizon over which fund closures are more likely given a specific characteristic. Third, provide a framework for thinking about which funds exhibit warning signs.

From a technical perspective, we are modeling the binary variable of whether a fund has become obsolete within the past 24 months against a set of observable characteristics. For the purposes of this model, we define an "obsolete" fund as one for which all share classes have been liquidated or merged. We do not consider a liquidation or a merger of a single share class.

For the remainder of the methodology document, we discuss our coverage universe, factor selection, constructing factor premiums, and the forecast calculation.

Universe Construction

The Morningstar Probability of Fund Shutdown Model uses Morningstar's survivorship-bias-free mutual fund database. In the United States, the fund data set begins in July 2006 because of data availability. Because we are interested in fund characteristics, we rolled up share-class data to the fund level. For funds providing complete asset information for all share classes, we calculated the asset-weighted variables. For those funds where complete asset information was not available, we computed equally weighted variables.

We further restrict our analysis to funds that possessed monthly returns and funds that are categorized into the Morningstar equity broad asset class. We further filter out any funds that have less than $1 million in assets under management or are less than 60 months old. We remove funds that are not in rated Morningstar Categories—for example, trading–inverse equity.

For model estimation purposes, we add additional filters to only include funds from firms with more than $1 billion in AUM. The model parameters are then extended to the larger universe described above.

The model's historical data does not suffer from survivorship bias. Morningstar's global fund databases retain a full history of dead funds, and these funds are included in the historical data set, where appropriate. Moreover, our evaluation technique dynamically incorporates monthly changes in fund universe composition, providing a more holistic and realistic picture of historical performance. Each monthly snapshot captures any funds that were subsequently merged or liquidated away.

Factor Exposures

Several key notions are necessary in understanding the way this model works:

- A **shutdown indicator** is the binary yes or no variable indicating whether a fund became obsolete within the following 24 months.
- A **factor** is an observable data point that appears to influence growth rates, like net expense ratio or fund size.
- A **factor exposure** is a number that measures how much a share class' growth rate is influenced by a factor. Exposures can be positive, negative, or zero. Exposures change through time.
- **Factor odds** is a number that represents how much a particular factor has changed the likelihood of closure for a particular time period.

At first launch, the model uses 14 factors that fall naturally into four distinct groups: performance, price, process, and parent. Exhibit 1 lists the entire set of factors. A more detailed treatment can be found in Appendix A.

Exhibit 1 Set of Fund Shutdown Factor Exposures

Performance	Price	Process
12-Month Growth Rate	Fund Fees	Firm AUM
12-Month Excess Return	Net Expense Ratio	Fund of Funds
12-Month Growth Rate x 12-Month Excess Return		Fund Age
36-Month Category Return		Fund Size
Category Closure Rate		Index Fund
Firm Closure Rate		

Source: Morningstar, Inc.

Model Methodology

To evaluate the specific drivers of fund obsolescence, we employ a panel logistic regression. We regress the 24-month shutdown indicator on a set of contemporaneous explanatory variables. We separate the data into regions to control for regulatory differences by regions. However, for now, we estimate the U.S. equity model only. As constructed, we believe the model offers investors a glimpse at the inherent factors leading to fund closures over a two-year time horizon.

Factor Odds Estimation

We apply the following framework to the data across regions:

Panel Regression:

$$Obsolete\ Status_i = \alpha + \lambda_t X_{i,t} + \sigma_t B_{i,t} + \varepsilon_{i,t+1}$$

$Obsolete\ Status_i$ is a binary variable where 1 indicates the fund closes within the next 24 months and 0 indicates the fund is still open in 24 months. $\mathrm{X_{i,t}}$ is a vector of explanatory characteristics at time t, and $\mathrm{B_{i,t}}$ is a set of indicator characteristics.

The panel regression, as specified above, is run for each model. As a result, we are left with a vector of coefficients. The output of a logistic regression is the change in log odds of a fund becoming obsolete given a certain characteristic. For simplicity, we convert from log odds to the change in odds. Given the below regression,

$$logit(Obsolete) = \propto_0 + \beta_i X_i$$

We calculate the percentage change in odds for a given variable as follows:

$$\%\ change\ in\ odds\ given\ X_j = e^{\beta_j} - 1$$

In Exhibit 2, we show the resultant odds from the model. The odds measure the likelihood of a fund becoming obsolete among each independent variable.

Exhibit 2 Effect of Factor on Shutdown for U.S. Equity Funds

Factor	12-Month Growth Rate	12-Month Excess Return	Fund Size	Category Closure Rate
Change in Odds	-1.5%	-7.2%	-19.8%	9.6%

Source: Morningstar, Inc.

Probability of Fund Shutdown

Our fund shutdown categorization is a two-step process. First, we calculate the forecast of fund shutdown. Second, we combine the fund shutdown probability with the fund flow forecast to categorize funds into different shutdown groups.

Forecast of Fund Shutdown

We apply the odds generated from the logistic regression to the fund factor exposures. The likelihood of closure is then the average of the factor odds weighted by its exposure to each component.

$$\boldsymbol{p_t} = \boldsymbol{w_t^T}(\mathrm{X_t f_t} + \epsilon_\mathrm{t})$$

Fund Shutdown Categorization

The fund shutdown categorization summarizes a fund's shutdown relative to the fund flow potential. To determine, we utilize the SWOT score produced by the Morningstar Fund Flow models. For each fund, we take the SWOT score of the largest share class, which represents the peer-relative growth potential. If the SWOT score is high, then the fund is positioned for growth in the future. Conversely, if the SWOT score is low, then the fund does not currently have the attributes investors prefer. We combine the two to classify funds into four buckets:

- Unlikely: Fund AUM > 0.5 z-score
- Likely: Probability of Shutdown > 15% and SWOT score < 33
- Wait & See: Probability of Shutdown > 15% and SWOT score > 33
- Unlikely: All Else

Conclusion

The probability of fund shutdown model for the U.S. equity market is Morningstar's first statistical model to understand the drivers behind fund closures. It complements the suite of fund flow models to understand investors' decision-making process for when and why they invest. In the coming months, we will be releasing other lifecycle models and improving the existing models.

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

References

Davidson, L., Sargis, M., & Strauts, T. 2017. The Fall of Funds: Why Some Funds Fail.
https://www-prd.morningstar.com/lp/fall-of-funds

Davidson, L., Sargis, M., & Strauts, T. 2016. The Rise and Fall of New Funds.
https://www.morningstar.com/lp/rise-and-fall-of-funds

Davidson, L. & Strauts, T. 2015. What Factors Drive Fund Flows?
https://www.morningstar.com/lp/what-factors-drive-investment-flows?

Morningstar Category Classifications. 2016.
https://www-prd.morningstar.com/content/dam/marketing/shared/research/foundational/MorningstarCategoryClassificationEffectiveApril2018.pdf

Morningstar Cash Flow Methodology. Oct. 31, 2011.
https://www-prd.morningstar.com/content/dam/marketing/shared/research/methodology/765555_Estimated_Net_Cash_Flow_Methodology.pdf

Morningstar Quantitative Rating for Funds Input Data Methodology. June 26, 2017.
https://www-prd.morningstar.com/content/dam/marketing/shared/research/methodology/813699_QuantitativeFundsInput.pdf

Morningstar Rating for Funds. Sept. 30, 2016.
https://www-prd.morningstar.com/content/dam/marketing/shared/research/methodology/771945_Morningstar_Rating_for_Funds_Methodology.pdf

Appendix A: Factor Definitions

Category Average Return

This is the trailing 36-month return of the fund's category. A higher score is indicative of a higher-performing category. Typically, fund closures are negatively correlated with long-term performance.

Category Closure Rate

This is a numerical variable indicating the rate of fund closure within a category. The calculation is the number of funds closed within the trailing five years, divided by the number of funds alive in the category, plus any subsequently incepted funds.

Cumulative Trailing-12-Month Growth Rate

For each share class, we sum the trailing flows over the past 12 months and then divide by the starting assets. We convert assets and flows into the model's denominated currency.

$$Growth\ Rate_{12\ Mo} = \frac{\sum_{t=1}^{12} flows_t}{Net\ Assets_{t=0}}$$

A higher score indicates a fund with a higher 12-month growth rate.

Fee Revenue

Fee Revenue is the summation of the share class net assets multiplied by the share class net expense ratio. Net assets are expressed in U.S. dollars. The net expense ratio used is the Global Net Expense Ratio Equivalent listed below.

Firm AUM

The Firm AUM number is a simple summation of each fund’s AUM assigned to the firm. Firm-level AUM is expressed in U.S. dollars. A higher score indicates a larger firm. With respect to investment flows, firm size matters. Larger firms attract a larger share of the assets.

Firm Closure Rate

This is a numerical variable indicating the rate of fund closure at a firm. The calculation is the number of funds closed within the trailing five years, divided by the number of funds alive in the category, plus any subsequently incepted funds.

Fund Age

Fund Age is measured as the number of months from inception to time t. Fund Age was also similarly right-skewed, and therefore it was necessary to log-transform it. A higher score indicates an older fund, whereas a negative score indicates a newer fund.

Fund of Funds

This is a categorical dummy variable that indicates whether a fund is structured as a fund of funds—a fund that specializes in buying shares in other mutual funds rather than individual securities. Quite often, this type of fund is not discernible from its name alone but rather through prospectus wording (that is, the fund's charter).

Fund Size (AUM)

Fund Size is measured as the fund's total market value of investments in USD. Not surprisingly, this data point had a heavy right-skewed distribution–there were much larger AUM funds than would be expected under a normal distribution. In order to better prepare these data for an OLS regression, we performed log-transformations on AUM. A higher score indicates a larger fund size.

Index Fund

This is a categorical dummy variable that indicates whether a fund tracks an index. While an index typically has a much larger portfolio than a mutual fund, the fund's management may study the index's movements to develop a representative sampling and match sectors proportionately.

Global Net Expense Ratio Equivalent

The Global Net Expense Ratio Equivalent data point definition should be consistent with Morningstar's Expense Ratio definition: "the annual fee that all funds or ETFs charge their shareholders. It expresses the percentage of assets deducted each fiscal year for fund expenses, including 12b-1 fees, management fees, administrative fees, operating costs, and all other asset-based costs incurred by the fund."

Moreover, this data point definition should hold across boundaries. The following is the logic for calculating a globally consistent expense ratio. Exhibit 3 outlines the calculation for an individual fund's Global Net Expense Ratio Equivalent (NER_{eq}).

Exhibit 3 Global Net Expense Ratio Equivalent

Single Fund:

$$Net\ Expense\ Ratio\ Equivalent\ (NER_{eq}) = \begin{cases} \textbf{Calculation} & \textbf{Domicile} \\ Indirect\ Cost\ Ratio\ (ICR)\ or\ Management\ Expense\ Ratio\ (MER) & Australia \\ MER & Canada\ or\ New\ Zealand \\ Ongoing\ Charge + Performance\ Fee & Europe\ or\ United\ Kingdom \\ Japan\ After\ Tax\ Total\ Expense\ Ratio & Japan \\ Net\ Expense\ Ratio\ (NER) & United\ States \\ NER & Else \end{cases}$$

Fund of Funds:

$Net\ Expense\ Ratio\ Equivalent\ (NER_{eq}) = FoFexp_i + \sum_{i=1}^{N} w_i NER_{eq}$ when Fund of Funds = TRUE

Source: Morningstar, Inc.

A high score indicates a more expensive fund. Investors, on average, prefer lower-fee funds, so there is a negative correlation between fees and fund flows.

Trailing-12-Month Excess Return

For each share class, we calculate the trailing-12-month cumulative U.S.-dollar return. For each category, we calculate the simple average of the trailing-12-month cumulative U.S.-dollar return for all share classes in the category. We then subtract the category average from each share class' return. A higher score indicates a fund with a higher excess return.

Appendix B: Input Data FAQ

How do we handle missing data?

In the case of missing data, we have a three-step process. First, we check if the fee data is missing. If so, we carry forward the fund's last calculated global net expense ratio equivalent for the trailing 12 months. Second, for share classes still missing fee data and all other data points, we cross-sectionally impute the median value of the Morningstar Category to which the fund is assigned. Third, in the rare case an entire category is missing data, we impute using the region-asset class group. We believe Morningstar Categories represents the appropriate peer group to determine similar characteristics. Furthermore, the model aims to forecast growth rates within categories. Therefore, imputed values will be treated as the "average" and hence unlikely to sway the forecast output.

How do we handle outlier data?

All continuous explanatory variables are winsorized between the 98% and 99% level and standardized to standard deviation units (mean 0, standard deviation 1) cross-sectionally by date and asset class. The dependent variable is winsorized at the 98% level but is not standardized. Winsorizing the outlier growth-rate data reduces the leverage of extreme observations and increases the variance captured within the monthly cross-sectional regressions.

How do we normalize the input data?

After all data is calculated and collected, we cross-sectionally normalize the data by asset class 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.

What are the investment types covered by the model?

The model covers exchange-traded funds, insurance funds, open-end funds, unit trusts, UK LP subaccounts, and variable annuity subaccounts.

What does "average" stand for?

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

About Morningstar® Quantitative Research

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

For More Information

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

22 West Washington Street
Chicago, IL 60602 USA