Internet Appendix for
"Are Subjective Expectations Formed as in Rational Expectations Models of Active Management?"

Magnus Dahlquist, Markus Ibert, and Felix Wilke

June 2025

A Sample analyst report

Below, we present an anonymized example of an analyst report. The report is for a fund rated under the new methodology and is entitled "Patient process and seasoned managers." Figure A1 shows, for a different fund, how the Analyst Rating is displayed on Morningstar's website.

Summary. *The fund's* experienced team and well-defined approach earn Morningstar Analyst Ratings ranging from Silver to Neutral depending on share class fees. The team invests in dividend-paying stocks for total return, not yield. The fund typically boasts a higher yield than the Russell 1000 Value Index and the S&P 500, but that's not its main objective. *The lead manager* looks for companies with business models and management teams capable of generating enough free cash flow to support and grow dividends, and tries to buy shares when they are undervalued relative to their cash flow. *She/he* buys when *she/he* sees at least 35% upside. The team is well equipped for their task. *The lead manager* started *her/his* career in fixed income and *her/his* experience evaluating company cash flows and liabilities has helped this strategy, which *she/he* started managing in 2002. Three comanagers—*manager A*, *manager B*, and *manager C*—averaging 22 years of industry experience and at least a decade with the team, support *her/him*. A senior analyst with five years' experience rounds out the squad. *The lead manager* and *her/his* team have posted a good risk/return profile. The fund's A shares have captured about three fourths of the Russell 1000 Value's and average large-value Morningstar Category peer's downsides since *the lead manager*'s 2002 start through October 2019. Its annualized return matched the index over that period, but its muted volatility led to superior risk-adjusted performance. The portfolio is not without risk. It has some of the largest sector bets in its category. At the end of September 2019, utilities accounted for 19.3% of the portfolio and consumer defensive stocks made up 27.0%. That's 12.3 and 17.3 percentage points, respectively, above the Russell 1000 Value's stakes. Both positions rank in the top 10 of all large-value peers. The portfolio's average debt-to-capital has also steadily increased over the previous five years. But, its average return on equity and return on invested capital have been consistently above the benchmark's. *The lead manager*, however, has managed those risks over more than one market cycle.

Process. This strategy's well-defined approach earns an Above Average Process rating. Management attempts to balance income, capital appreciation, and capital preservation. *The lead manager* and *her/his* team focus on stocks with steady and increasing dividends, but they look beyond the dividend. Each team member conducts research to project a company's total-return potential during the next two to three years, focusing on companies with strong free cash flows and management teams. *The lead manager* and *her/his* comanagers seek capital appreciation by buying stocks that they determined have at least 35% upside from their current price based on cash flow and dividend discount models and other valuation measures. The team aims to preserve capital by modeling a "bear" case for each stock. They consider the market and company factors that could negatively affect the stock's price and require at least a 3-to-1 upside from the bear case to invest. If a stock's price falls more than 15% from its cost basis, a second analyst reviews the stock to provide a "devil's advocate" point of view. This approach produces a portfolio of 70-85 stocks that covers all sectors, though weightings deviate from the Russell 1000 Value Index. The fund may hold up to 25% of its assets in international stocks, and it has held double-digit cash allocations under *the lead manager*'s tenure. Though it has historically provided protection in tough conditions, the current portfolio is not without risks. First, it's heavily concentrated in two sectors: Utilities accounted for 19.3% and consumer defensive stocks 27.0% of the portfolio at the end of September 2019. That's 12.3 and 17.3 percentage points above the Russell 1000 Value Index's stakes, respectively. The heavy helping of consumer defensive stocks is not new, but the bet on utilities relative to the benchmark has risen steadily over the last five years. Its debt-to-capital ratio has also increased over that span and reached 48% in September 2019—10.0 percentage points above its 2014 level and 6.1 percentage points above the benchmark's ratio at the same period. But the companies in the portfolio have been generating solid returns. The portfolio's average return on equity and return on invested capital are both regularly above the benchmark's—the 19.3% ROIC over the last trailing 12 months through September 2019 was nearly 4.8 percentage points above the benchmark's. It has also kept its yield above the Russell 1000 Value and S&P 500. But *the lead manager* and *her/his* team are also looking for companies with at least 35% upside, such as wide moat brewer Anheuser Busch InBev ABI, which has a low ROE and ROIC but has been acquiring growing brands to increase distribution and hopes to increase margins through cost-cutting.

People. Stable leadership earns this strategy an Above Average People rating. *The lead manager* started on the team in 2002 and took over the fund one year after its inception. *She/he* joined *the fund family* in 1991 as a fixed-income trader and managed bond portfolios before shifting to equities in 1998. *The lead manager* has promoted comanagers from analyst positions, such as April 2016 when *she/he* advanced *manager C*, an analyst since early

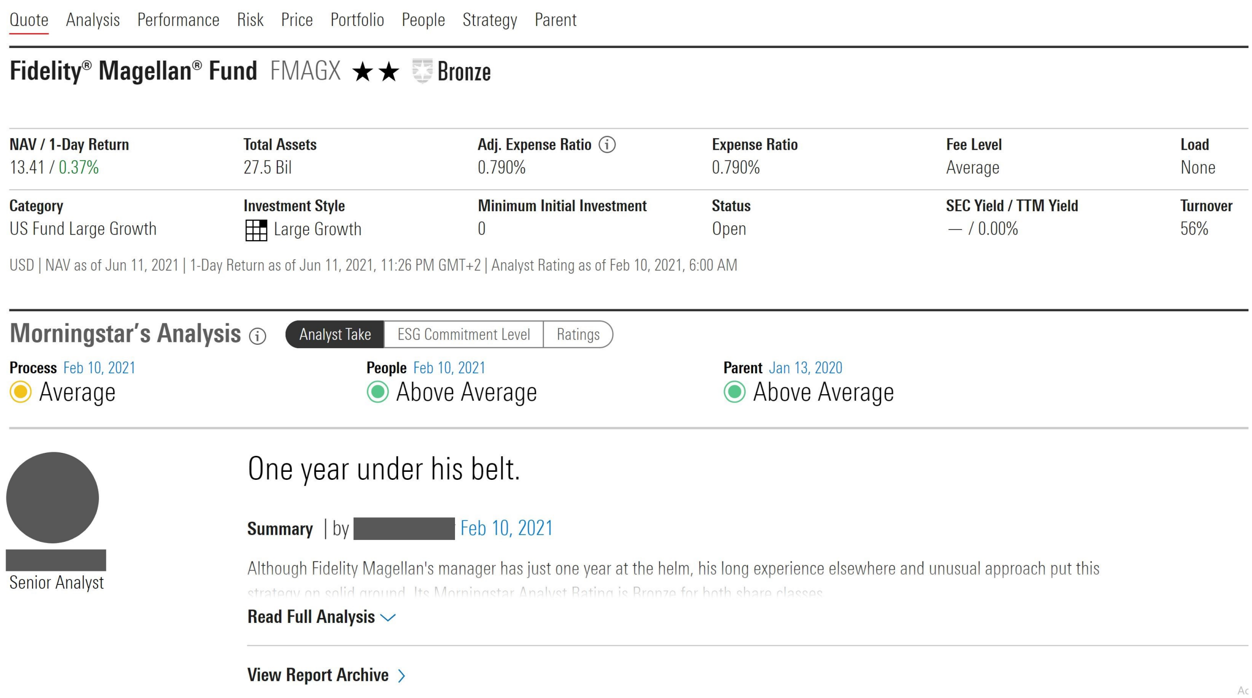
2009. *Manager A* and *manager B* became comanagers in early 2014, a few months before then-portfolio manager *manager D* left the firm. *Manager A* and *manager B* had 10- and eight-years' experience as analysts on the strategy, respectively, before their promotions. In 2014 *the lead manager* hired experienced *analyst A*, who worked closely with *the fund family* veteran *manager E* before *she/he* retired in 2016. Though the team works collaboratively, each member has sector responsibilities. *The lead manager*, for instance, covers financials and industrials. *She/he* also rotates sector responsibilities and tries to give each team member a mix of cyclical and non-cyclical assignments to keep fresh perspectives on companies. *The lead manager* invests more than $1 million in the fund. *His/her* comanagers have smaller investments (between $100,000 and $500,000). Part of the managers' and analysts' deferred compensation is invested in restricted shares of the fund.

Parent. *The fund family* is a vast conglomerate that is growing further by acquiring *fund family B*. Acquisitions are a way of life for *the fund family*: Among them have been *fund family C* in the 1990s, *fund family D* and *fund family E* in the early 2000s, *fund family F* in 2006, *fund family G* in 2010, and the exchange-traded fund business of *fund family H* more recently. The firm's many areas—whether acquired or homegrown—present a mixed picture. In the United States, areas of strength include small-cap U.S. growth funds, dividend-focused funds, and the international funds run by the *specialized* team. The corporate-bond and quantitative equity teams in Europe also stand out. But many U.S.-focused active stock funds have suffered from poor performance and/or manager turnover. Manager turnover has also been an issue with some Hong Kong-based offerings. Various fixed-income teams in the U.S. are well-staffed, but performance has been so-so. Meanwhile, *the fund family*'s passive side has grown nicely, but there are few truly compelling choices. As for *fund family B*, that firm brings some strong international funds with substantial assets, and the *fund family B* addition also allows for cost-cutting. *The fund family CEO A* has plenty of experience in integrations. All told, along with the bright spots there remain many average or underperforming funds and uncertainty how the *fund family B* merger will play out. *The fund family* thus retains its Neutral Parent rating.

Performance. *The fund* has historically given investors some downside protection and outperformed on a risk adjusted basis. From *the lead manager's* December 2002 start through October 2019, the fund's 9.1% annualized return just about matched the Russell 1000 Value Index, but it beat the average large-value peer by 1.0 percentage points. With below-average volatility, the fund's risk-adjusted performance is better than both its benchmark and the typical peer. But it hasn't performed as well against comparable dividend focused indexes. Over the last 10 years through October 2019, the fund trailed the FTSE High Dividend Yield Index's 13.2% annualized return by 2.0 percentage points. The portfolio's posture—with heavy helpings of consumer defensive and utilities stocks—has helped in market drawdowns in the past. The fund captured about three fourths of the Russell 1000 Value's and average large-value Morningstar Category peer's downsides over *the lead manager's* tenure. But the posture hasn't always helped. Underweighting technology stocks and holding a large cash stake, which peaked at 18% in early 2017, have been a drag on recent performance, including in the 2016 and 2017 market rally. The fund's 7.9% annualized return over the last three years through October 2019 trails the Russell 1000 Value by 2.6 percentage points and fell in the bottom decile of the peer category.

Price. It's critical to evaluate expenses, as they come directly out of returns. The share class on this report levies a fee that ranks in its Morningstar category's middle quintile. That's not great, but based on our assessment of the fund's People, Process and Parent pillars in the context of these fees, we think this share class will still be able to deliver positive alpha relative to the category benchmark index, explaining its Morningstar Analyst Rating of Bronze.

Figure A1: An example from the Morningstar website



The figure shows an example of how the Analyst Rating is displayed on Morningstar's website. After searching for a fund on Morningstar's website, the Analyst Rating is shown next to the fund name in the *Quote* section at the very top. In the example of this figure, Fidelity's Magellan fund received an Analyst Rating of "Bronze." The next Section *Analysis* displays the Analyst Rating pillar scores in detail as well as the written analyst report accompanying every Analyst Rating. We have redacted the analyst's identity.

B Data appendix

B.1 Morningstar data

For our main results, we retrieve the universe of worldwide open-end equity mutual funds from Morningstar Direct as of 9 February 2021.[1] The data belong to 416 Morningstar categories, which are exclusively designated "Equity" by the Morningstar variable *Global Broad Category Group* and include live as well as dead funds. We effectively exclude bond funds, money market funds, target-date funds, as well as other non-equity funds and we ensure that all funds have a *Morningstar Category*. The data contain, among other variables, Morningstar's fund and share-class identifiers, the *Global Category*, the *Morningstar Category*, returns, share-class net assets, fund sizes, fees, and monthly Morningstar Analyst and Quantitative Ratings.[2] We download the entire time series from January 1972 to February 2021, but benchmark returns are only available from January 1979 and onwards. In total, we collect data for 195,519 share classes (as identified by SecId) belonging to 59,102 unique funds (as identified by FundId); 44,162 funds have at least one non-missing return.

We proceed in two separate steps. First, we describe the data for replicating Analyst and Quantitative Ratings on the share-class level, for which we intend to use the data that Morningstar uses. Second, we describe the data for estimating the rational model of fund performance, for which we intend to use the data that academic research has previously used. In the end, we merge the two datasets to arrive at the final sample for our cross-sectional regressions. As the rational model of fund performance is estimated using annual data and all funds have a rating under the new rating methodology as of the end of December 2020, we primarily use data available as of the end of December 2020.[3]

B.2 Replication of Analyst and Quantitative Ratings

The replication of Analyst and Quantitative Ratings follows the three broad steps laid out in the main text:

[1] For some auxiliary results in this Internet Appendix, we have also retrieved the same data again as of 28 January 2022 to update these auxiliary results.

[2] We noticed that several Morningstar Analyst Ratings originally published from 2011 to 2013 are missing from the data downloaded as of February 2021. However, the ratings are available in data downloaded in January 2020 and corresponding written analyst reports are still available on Morningstar's website as of February 2021. Therefore, we recover the missing ratings from our earlier downloaded data.

[3] For some variables in Morningstar Direct (e.g., manager ownership), only the latest values of a variable (i.e., a snapshot) as opposed to the entire time series are stored, such that these variables are as of 9 February 2021 and not as of the end of December 2020.

1. Estimate the semi-interquartile range (SIQR) as a measure of strategy potential for a given group of funds.

2. Construct the before-fee (i.e., "gross") fund alpha based on the SIQR and pillar scores assigned to individual funds by Morningstar analysts.

3. Subtract share-class fees from gross fund returns and bin the resulting after-fee (i.e., "net") alphas into the final ratings.

B.2.1 Gross returns

To estimate historical gross fund alphas (Equation (1) in the main text), we use a variable for the gross return, which is presumably what Morningstar does too, as opposed to adding fees back to net returns.[4] Morningstar uses the fee variable *Representative Cost* to calculate gross returns from net returns. Hence, using net fund returns and adding back the monthly representative cost should yield similar gross returns.

B.2.2 Benchmark indexes

For the benchmark return in Equation (1) in the main text, we use the return of the *Morningstar Category Index* of a particular *Morningstar Category*. Since a fund's *Morningstar Category* can vary over time, we generally work with the historical *Morningstar Category* as opposed to the snapshot version and we exclude fund-month observations for which the *Morningstar Category* takes on values other than the 416 Morningstar categories that we download (this may happen because historically funds may have belonged to non-equity categories).[5]

[4] We take a value-weighted average of gross share-class returns to form the gross fund return. We do this before our cleaning and imputation procedures for assets under management (AUM), since we do not believe analysts employ these procedures. In the data, gross share-class returns for a given fund are very similar with slight divergences.

[5] The Morningstar Category is mostly unique among all share classes of a fund, with a few exceptions in which a fund's share classes belong to two Morningstar Categories. In all of those cases, one of the two Morningstar Categories is either "EAA Fund Other Equity" or "EAA Fund Property—Indirect Other." Neither category has a designated *Morningstar Category Index* or share classes with Morningstar Quantitative Ratings, but both categories contain some share classes with Morningstar Analyst Ratings. Therefore, we believe that it is likely that in Morningstar's process of awarding the ratings, all share classes of those funds with a Morningstar Analyst Rating and with two Morningstar categories are included in the other *Morningstar Category* we see among the share classes of the respective funds (i.e., the category that is not "EAA Fund Other Equity" or "EAA Fund Property—Indirect Other"). We proceed by setting the *Morningstar Category* to equal that of the other *Morningstar Category* for all share classes of the fund in order to correctly replicate the ratings. Picking the *Morningstar Category* that has most of the fund's AUM leads

B.2.3 Fund strategy potential (SIQR)

Equipped with the time series of gross fund returns and benchmark returns, we estimate all active funds' rolling 36-month gross alphas from January 2000 forward according to Equation (1) in the main text.

To calculate the SIQR for a particular type of strategy, Morningstar groups funds that invest in the same universe of stocks by aggregating Morningstar categories from different fund markets around the world (e.g., funds registered in the U.S. and funds registered in Europe). However, Morningstar is not explicit about the exact mapping of Morningstar categories into such super groups. These super groups are used solely to assess the alpha opportunity of fund strategies and the remainder of the rating setting occurs within Morningstar categories.

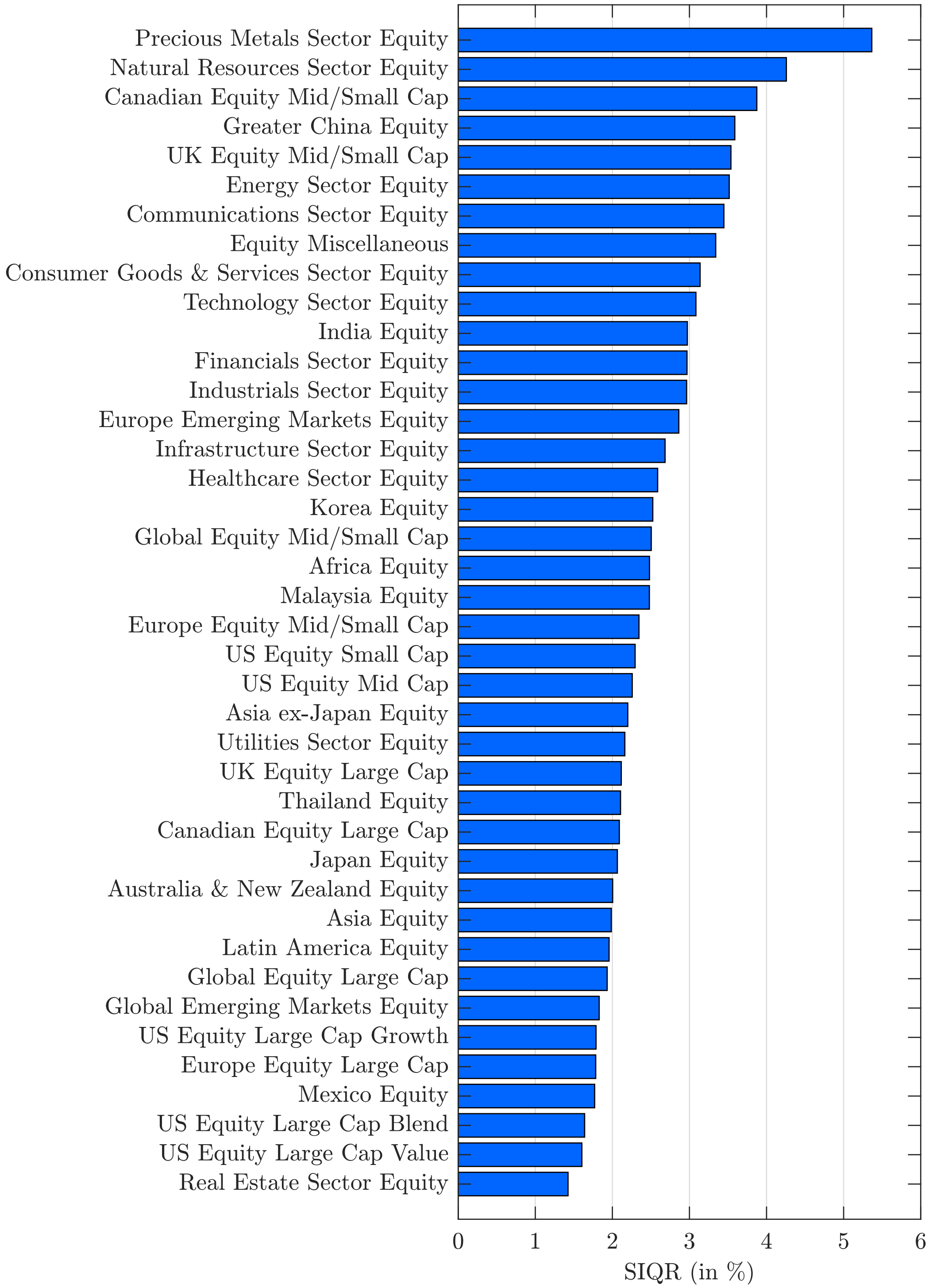
We group Morningstar categories based on the *Global Category* to calculate the SIQR and assign an SIQR to every fund based on its *Morningstar Category* in December 2020. First, we identify the most common *Global Category* among all funds within each *Morningstar Category*. Most funds within a *Morningstar Category* share the same *Global Category*. Then, we bundle all Morningstar categories that have the same most common *Global Category*. In total, we aggregate funds to 40 different strategies based on 40 global categories in our sample.[6] When grouping fund alphas, we exclude index funds (as identified by *Index Fund*), but keep smart beta funds (as identified by *Strategic Beta*), following Morningstar's methodology. Finally, we calculate the SIQR of the resulting distribution of realized alphas, which reflects Morningstar's assessment of the potential of a given strategy.

Figure B1 shows our estimates of the SIQR.

to the same result for 91% of the funds.

[6] For example, the Morningstar categories "US Fund Large Value" and "EAA Fund US Large-Cap Value Equity" are grouped to form the fund strategy "US Equity Large Cap Value."

Figure B1: Semi-interquartile range (SIQR) for global categories



The figure shows our estimates of the semi-interquartile range (SIQR) of different global categories as of December 2020. The SIQR reflects Morningstar's assessment of the potential of a given strategy.

B.2.4 Pillar scores

Morningstar analysts evaluate funds based on three areas that they believe are crucial in order to predict future success: People, Process, and Parent. These pillar scores are available in the database. However, we noticed that pillar scores are missing for some share classes of funds that have a Morningstar Analyst or Quantitative Rating. Since pillar scores are awarded at the fund level, we fill in missing data from other share classes of the same fund.[7]

Next, we set pillar scores to missing if the analyst report that outlines the ratings and justifies the ratings decision is more than one year old. We do so because Analyst Ratings have to be updated once per year according to Morningstar's policies. In a few instances, some share classes have Analyst Ratings as well as Quantitative Ratings, which occurs when analyst coverage of a fund has just ceased. Then, the last Analyst Rating appears alongside the first Quantitative Rating in the data. We keep the more recently published Quantitative Rating and the corresponding pillar scores.

We then calculate the forward-looking gross alpha according to Equation (2) in the main text.

B.2.5 Fees

Under the new methodology, Morningstar deducts share-class-specific fees from gross alphas to arrive at net alphas and awards Analyst Ratings for each share class. Morningstar uses the fee variable Representative Cost, which contains Morningstar's best estimate of the recurring costs charged by funds.

We noticed that fees are still missing for some share classes that have a rating in December 2020. In such cases, we source fees at the end of the sample from other variables to replicate as many ratings as possible. In particular, we fill in missing data using the *Annual Report Net Expense Ratio*, *Ongoing Cost*, *Prospectus Net Expense Ratio*, and the *Semi-Annual Report Net Expense Ratio*, in that order. We set observations less than or equal to zero to missing for all fee variables that we consider before merging data.

[7]Filling in pillar scores allows us to calculate alphas for every share class of a rated fund and to eventually calculate a value-weighted fund-level net alpha reflecting the fee structure of all share classes. However, we do not include alphas of share classes that do not have a Morningstar Analyst or Quantitative Rating in the data when binning net-of-fee alphas into final ratings for our replication exercise.

B.3 Data for estimating the rational model of fund performance

Replicating the Analyst Ratings only requires a historical time series of gross fund and benchmark returns. To estimate the rational model of fund performance, in addition we need historical data on fund sizes. Before estimating the model, we first clean the data in accordance with the literature (e.g., Pástor, Stambaugh, and Taylor, 2015; Berk and van Binsbergen, 2015).

We start from the monthly gross return dataset, which has 11,909,891 share-class-month observations with non-missing returns. Then, we merge in other variables. We merge in only observations of the share-class-month when return data exist (in month t or $t+1$). If a variable is missing, we keep the share-class-month observation and record a missing value for that variable.

B.3.1 Fees

Since we use gross returns in estimating the model, we do not need additional fee data for the model estimation itself, but will use fees as a filter to exclude funds that are unlikely to be actively managed. Our measure of fees is again *Representative Cost*, which is generally populated using a fund's net expense ratio (this can be from the annual report, semi-annual report, or another source) according to Morningstar. At the share-class level, we set fees less than or equal to zero to missing.

Then, we fill in missing data with the annual report net expense ratio. First, we set the net expense ratio to missing if it is less than or equal to zero. Next, we place the net expense ratio at the fiscal year end month if available in Morningstar Direct, and otherwise assume that the fiscal year ends in December. Afterwards, we backward fill missing month ends for up to twelve months (or until the previous reported value) first and then forward fill for up to twelve months. Finally, we use this series to fill in missing monthly fee data.

B.3.2 Cleaning assets under management

Pástor et al. (2015) discover instances of extreme reversal patterns in AUM in the Morningstar data that likely reflect decimal-place mistakes. We adopt their procedure to remove these extreme reversals in monthly fund sizes as well as share-class net assets. First, we create a variable for the fractional change in assets from last month to the current month,

$$\%\text{AUM}_t = \frac{\text{AUM}_t - \text{AUM}_{t-1}}{\text{AUM}_{t-1}}. \tag{B1}$$

Second, we create a reversal variable to capture the reversal pattern,

$$\text{Reversal}_t = \frac{\text{AUM}_{t+1} - \text{AUM}_t}{\text{AUM}_t - \text{AUM}_{t-1}}. \tag{B2}$$

This variable will be approximately –1 if there is a reversal (e.g., 20 million, 2 million, 20 million). Finally, if

$$\text{abs}(\%\text{AUM}_t) >= 0.5, -0.75 > \text{Reversal}_t > -1.25, \text{ and } \text{AUM}_{t-1} >= 10 \text{ million}, \tag{B3}$$

then we set assets at time t (i.e., 2 million in this example) to missing. As a result of this procedure, 0.05% of monthly fund size and 0.02% of monthly share-class net asset observations are set to missing.

We use share-class net assets when aggregating variables such as returns or fees to the fund level and therefore need monthly asset information. However, there are a significant number of missing asset observations. This is in part due to funds reporting at a quarterly or annual frequency, particularly before 1993. We apply the following procedure to fill in missing monthly share-class net assets and fund sizes:

1. We impute missing values in the middle of the data series by using their past values, returns, and a factor adjusted for flow rates as done by Ibert, Kaniel, Van Nieuwerburgh, and Vestman (2018). Specifically, let $[t_0, t]$ and $[t+n, T]$ be periods when asset data are non-missing. The missing values are filled in as follows:

$$\text{AUM}_k = F \times \text{AUM}_{k-1}(1 + r_k), \text{ for } k \in [t+1, t+n-1], \tag{B4}$$

$$F = \left(\frac{1}{\prod_{k=t+1}^{t+n}(1 + r_k)} \frac{\text{AUM}_{t+n}}{\text{AUM}_t} \right)^{\frac{1}{n}}, \tag{B5}$$

where F is the factor adjusted for flow rate and r_k is the return. We implement this step allowing for a maximum gap of twelve months between non-missing observations at times t and $t + n$.

2. When returns are not available for all months with missing asset data between times t and $t + n$, we linearly interpolate the missing observations, again allowing for a maximum gap of twelve months.

3. If assets are missing for the last month in the sample, we forward fill the latest available data going back for a maximum of twelve months from the sample end to account for a time lag in reporting.

4. Finally, we set observations for which assets are zero or negative to missing.

B.3.3 Aggregation of share-class level to fund level

We take value-weighted averages of returns and fees across share classes using lagged share-class assets as weights to form fund-level variables. We take the average across all non-missing share-class values and do not set values to missing at the fund level when one or more share classes have missing data. If all share classes have missing assets, we take an equal-weighted average. We treat the fund size variable as AUM on the fund level and use the sum of share-class net assets if fund size is missing.

B.3.4 Benchmark indexes

A mutual fund's *Morningstar Category* can evolve over time, for example, due to the fund experiencing style drift (e.g., from US Fund Small Cap Growth to US Fund Small Cap Blend). Therefore, we use the *Morningstar Category* time series to assign a benchmark return for every fund-month. We forward and backward fill the *Morningstar Category* for a maximum of twelve months and exclude fund-month observations for which the *Morningstar Category* takes on values other than the 416 Morningstar categories that we download.

B.3.5 Further sample restrictions

Following Pástor et al. (2015), we exclude fund-month observations with fees below 0.1% per year, since it is unlikely that any actively managed fund charges such low fees. In addition, we exclude fund-months with fees above 20% per year. Moreover, we exclude observations before the fund's inflation-adjusted AUM reached USD 5 million, as done by Berk and van Binsbergen (2015) and Fama and French (2010). We keep only funds with twelve monthly observations in a given year and twelve non-missing returns. When going from fund-month to fund-year, we keep the observation in December of each year. Next, we check whether a given fund has a gap in the annual dataset. If a fund has a missing year, we delete all the fund's observations from the sample.

B.3.6 Identifying index funds

To create a dummy variable to identify index funds, as done by Pástor et al. (2015), we use a simple two-step procedure:

1. If Morningstar identifies a fund as an index fund (identified by the variables *Index Fund* or *Enhanced Index*), then we classify it as an index fund. Otherwise, we move to the next step.

2. If the fund name contains "Index" or "index," we classify it as an index fund.

Otherwise, we classify the fund as active. As a result of this procedure, we identify and exclude 5,331 index funds out of 59,102 funds (9.0%).

B.3.7 Inflation adjustment

To make AUM comparable across time, we adjust for inflation using the *Consumer Price Index* from the Federal Reserve Economic Data provided by the St. Louis Fed (FRED). We use the *Consumer Price Index for All Urban Consumers: All Items in U.S. City Average* (CPIAUCSL) series and express all USD items in December 2020 USD.

B.4 Aggregation of analyst alphas from share class to fund level

The replication of ratings is on the share-class level using the data of Section B.2. After validating our replication, for our main analysis we take a value-weighted average of analyst alphas across share classes to arrive at a fund-level alpha using the cleaned share-class assets from above.

We take the average across all non-missing share-class assets and do not set assets to missing at the fund level when one or more share classes have missing data. For value-weighting, we use lagged share-class net assets. If all share classes have missing assets, we take an equal-weighted average.

B.5 Rational learner alpha

Using the data from Section B.3, we estimate the rational model of fund performance. Since we estimate the model using annual data, we use return data up to December 2020 to estimate a fund's perceived skill. Then, we form rational learner alphas at the end of our sample for every fund according to Equation (7) in the main text using perceived skill, fees, and fund sizes measured at the end of December 2020.

The intersection of the fund-level analyst alpha data (Section B.4) and the data for the model estimation (Section B.3) is the sample for the main regressions in the paper.

Lastly, in our empirical implementation of the model, the forecast horizon is one year. Morningstar states, for example, that the medalist ratings indicate an expected outperformance "over the long term, meaning a period of at least five years." To compare analyst alphas with those of our model, we assume that analysts' five-year forecasts equal their unobserved one-year forecasts. We also assume that the fee is contracted in advance such that $\mathrm{E}_t[f_{i,t+1}] = f_{i,t+1}$ and that it can be approximated by the current fee.[8] An alternative to these assumptions would be to iterate Equation (3) in the main text forward using a law of motion for AUM and the expected path of fees. However, modeling the path of fees and a law of motion for AUM would significantly complicate the model; it would require additional assumptions as to the fee-setting behavior of the fund over time and as to how investors' money flows into and out of funds in response to past performance.

[8]The management fee, the largest part of a fund's overall fee (also known as the expense ratio), is indeed known in advance. However, other parts of the overall fee are not necessarily known in advance (e.g., distribution costs). In any case, funds' overall fees are extremely persistent (see, e.g., Cooper, Halling, and Yang, 2020) and the fee this year is a reasonable forecast of the fee next year.

C Perturbations of rational expectations learning model

C.1 Flexible decreasing returns to scale technology

Our initial cost function for the impact of size on returns, c, was the logarithmic function. A motivation for this assumption is given in Panel (a) of Figure C1, which plots realized before-fee alphas against the logarithm of a fund's AUM at the end of the previous year. The relationship between fund returns and the logarithm of AUM is approximately linear, consistent with our assumption in the main text (see Equation (3) in the main text).

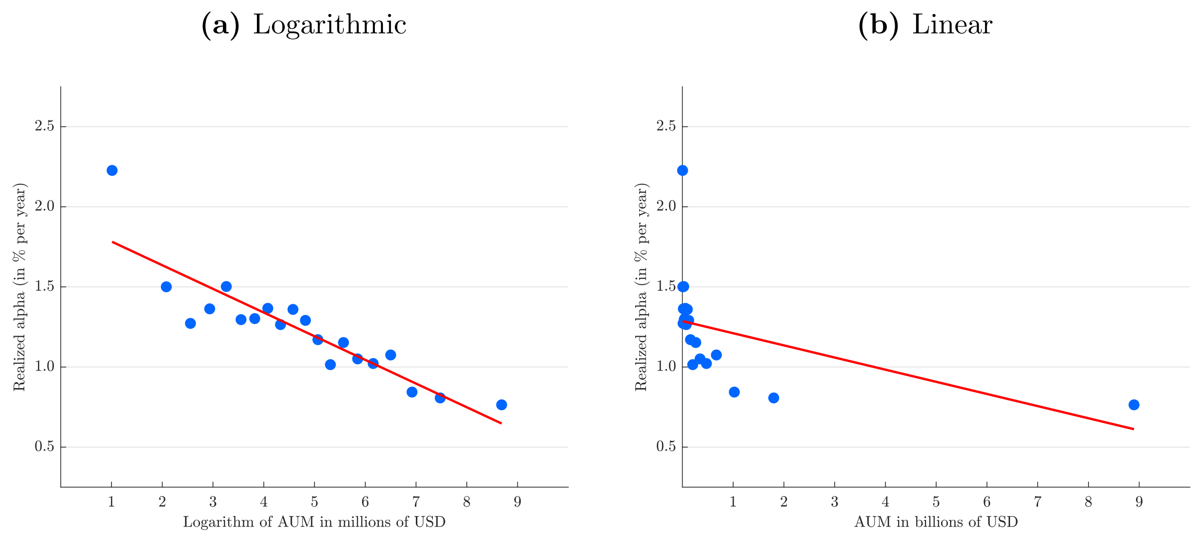
In contrast, Panel (b) shows the relationship between realized before-fee alphas and the level of AUM, corresponding to a linear cost function, c. This relationship is not well approximated by a linear function. In particular, the distribution of AUM has positive skewness such that the relationship between the returns and AUMs of the largest funds drives the average relationship.

To formalize this argument, we allow for a more flexible impact of AUM on returns in the rational expectations learning model: $c(AUM) = \eta\frac{(AUM)^{\gamma}-1}{\gamma}$, with $\gamma \in (0,1)$ and γ being an additional parameter to estimate (as in Roussanov, Ruan, and Wei, 2020). If $\gamma = 1$, the cost function is linear in AUM, as in Panel (b) of Figure C1. As γ approaches zero, the cost function converges to the logarithmic function, as in Panel (a) of Figure C1.

Table C1 presents the parameter estimates and their standard errors for the rational expectations learning model with the flexible functional form of the decreasing returns to scale technology. The parameter estimates are similar to our baseline estimates and the shape parameter γ is 0.12, indicating that a logarithmic functional form fits the data better than does a linear functional form, consistent with the intuition obtained from Figure C1.

For completeness, we recalculate perceived skill using the parameter estimates in this subsection and rerun our main regressions, corresponding to Tables 5 and 6 in the main text. Tables C2 and C3 present the results, which are similar to our baseline results in the main text. In specification (4) of Table C3, the coefficient estimate on size is positive, but not statistically different from zero. The coefficient estimate not statistically different from zero does not affect our main conclusions, since the estimate is still significantly different from the model-implied effect of size on returns (p-value of 0.00).

Figure C1: Realized before-fee alphas against lagged fund size



The figure shows realized before-fee alphas relative to each fund's Morningstar Category benchmark against lagged fund size using fund-year observations from 1979 to 2020. Fund size is measured as the logarithm of AUM in millions of USD in Panel (a) and as AUM in billions of USD in Panel (b), corresponding to a logarithmic and a linear cost function, respectively. We group lagged fund size into 20 equal-sized bins, compute the mean of lagged fund size and realized alphas within each bin, and then create a scatterplot of these data points.

Table C1: Parameter estimates of the rational fund performance model with flexible decreasing returns to scale technology

Parameter	Description	Estimate
η	Decreasing returns to scale (%)	0.141***
		(0.023)
γ	Shape of DRS	0.117***
		(0.029)
a_0	Prior mean (%)	2.043***
		(0.079)
$\sigma_{a,0}$	Prior standard deviation (%)	2.092***
		(0.041)
σ_ϵ	Residual standard deviation (%)	8.112***
		(0.015)
ρ	Skill persistence	0.950***
		(0.006)

The table shows the parameter estimates of the rational fund performance model with flexible decreasing returns to scale technology in % per year. Standard errors are shown in parentheses. The model is estimated using fund-year observations from 1979 to 2020. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

Table C2: Cross-sectional regressions of alphas on fund characteristics with flexible decreasing returns to scale technology

	Analyst Ratings	Analyst and Quantitative Ratings
	(1)	(2)
Perceived skill	0.377***	0.695***
	(0.066)	(0.040)
	[0.000]	[0.000]
Size × η	0.192**	0.550***
	(0.095)	(0.088)
	[0.000]	[0.000]
Fees	−0.943***	−1.525***
	(0.151)	(0.059)
	[0.708]	[0.000]
Constant (× 100)	0.280	−1.258***
	(0.235)	(0.150)
	[0.234]	[0.000]
N	1454	13934
Adj. R^2	0.15	0.32

The table shows regressions of Morningstar analyst alphas on skill as perceived by a rational learner, fund size × η, and fees for cross-sections of funds in December 2020. Fund size is measured as $\frac{\text{AUM}^{\gamma}-1}{\gamma}$, where AUM refers to a fund's total assets under management in millions of USD and γ is given in Table C1. Specification (1) uses funds with an Analyst Rating. Specification (2) uses funds with an Analyst Rating or a Quantitative Rating. Alphas are relative to each fund's Morningstar Category benchmark. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient. In brackets are p-values for the null hypothesis that the coefficients of skill, size × η, fees, and the constant equal the model-predicted parameters of +1, −1, −1, and 0, respectively.

Table C3: Cross-sectional regressions of alphas on additional fund characteristics with flexible decreasing returns to scale technology

	Analyst Ratings		Analyst and Quantitative Ratings	
	(1)	(2)	(3)	(4)
Rational learner				
Perceived skill	0.266***	0.096	0.852***	0.338***
	(0.066)	(0.072)	(0.080)	(0.057)
Size × η	0.461***	0.219**	0.354***	0.139
	(0.118)	(0.108)	(0.090)	(0.105)
Fees	−1.325***	−0.950***	−1.761***	−0.968***
	(0.142)	(0.116)	(0.197)	(0.212)
People				
Manager tenure		0.109***		0.248***
		(0.040)		(0.033)
Manager ownership		0.115**		0.193***
		(0.053)		(0.041)
Managerial multitasking		0.649***		0.580***
		(0.204)		(0.193)
Management team		0.096		0.496***
		(0.110)		(0.114)
Process				
Top 10 assets (%)		0.128		−0.027
		(0.130)		(0.091)
Tracking error		−0.005		−0.153
		(0.065)		(0.093)
Turnover ratio		−0.485***		−0.111
		(0.155)		(0.081)
Retail		−0.289***		−0.157*
		(0.092)		(0.089)
Broker-sold		−0.267**		−0.063
		(0.117)		(0.105)
N	698	650	2830	2626
Adj. R^2	0.26	0.62	0.29	0.64
Sustainability FE	No	Yes	No	Yes
Star FE	No	Yes	No	Yes
Morningstar Category FE	No	Yes	No	Yes
Fund Family FE	No	Yes	No	Yes

The table shows regressions of Morningstar analyst alphas on fund and manager characteristics for cross-sections of funds in December 2020. Specifications (1) and (2) use U.S.-domiciled funds with an Analyst Rating. Specifications (3) and (4) use U.S.-domiciled funds with an Analyst Rating or a Quantitative Rating. Alphas are relative to each fund's Morningstar Category benchmark. Manager tenure is the maximum tenure (in months) taken over all managers, manager ownership is the average amount managers of a fund personally invest in the fund, managerial multitasking is the average number of additional funds that managers of a particular fund manage, and management team is a dummy for team-managed funds. Top 10 assets is the percentage of AUM in the ten largest positions, tracking error is the standard deviation of returns in excess of the benchmark over the life of the fund, turnover is a fund's trading activity as reported to the SEC, retail is a dummy for whether a fund is primarily held by retail investors, and broker-sold is a dummy for whether a fund is primarily sold through brokers. "People" and "Process" variables are standardized to zero mean and unit standard deviation (except for the dummy variables), and the coefficient estimates are multiplied by 100. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

C.2 Heterogeneity in decreasing returns to scale

C.2.1 Active share

In an extension of their model, Berk and Green (2004) capture the idea that more active funds are subject to steeper decreasing returns to scale by allowing funds to index part of their assets, that is, to directly invest in the passive benchmark. Investors still pay the fee on this part, but since it is not actively managed it does not affect returns through the cost function c in Equation (3) in the main text.

If active funds are allowed to index part of their assets, following Equation (11) from Berk and Green (2004), the measurement equation becomes

$$r_{i,t+1} + f_{i,t+1} = h_{i,t}a_{i,t} - c(h_{i,t}AUM_{i,t}) + h_{i,t}\epsilon_{i,t+1}, \tag{C1}$$

where $h_{i,t}$ refers to a fund's fraction of assets that are actively managed. The state transition equation is the same as before, that is, Equation (4) in the main text. The updating equations become

$$\widehat{a}_{i,t+1} = \rho\left(\widehat{a}_{i,t} + \frac{\widehat{\sigma}^2_{a,t}}{h_{i,t}(\widehat{\sigma}^2_{a,t} + \sigma^2_{\epsilon})}(r_{i,t+1} - h_{i,t}\widehat{a}_{i,t} + c(h_{i,t}\text{AUM}_{i,t}) + f_{i,t+1})\right) + (1-\rho)a_0, \tag{C2}$$

$$\widehat{\sigma}^2_{a,t+1} = \rho^2\widehat{\sigma}^2_{a,t}\left(1 - \frac{\widehat{\sigma}^2_{a,t}}{\widehat{\sigma}^2_{a,t} + \sigma^2_{\epsilon}}\right) + (1-\rho^2)\sigma^2_{a,0}. \tag{C3}$$

Our original cost function, c, was the logarithmic function. Theoretically, a fund could index all of its assets so that the log of actively managed assets is undefined. Therefore, we choose the more flexible form of the impact of scale on returns from the previous subsection: $c(hAUM) = \eta\frac{(hAUM)^{\gamma}-1}{\gamma}$ with $\gamma \in (0, 1]$.

We estimate a fund's three-year rolling-window R^2 relative to the benchmark and compute the active share as $1 - R^2$ (Amihud and Goyenko, 2013). We estimate the model using maximum likelihood, recalculate perceived skill using Equation (C2) at the end of our sample, and reproduce Tables 5 and 6 in the main text using the corresponding variables from the measurement equation, Equation (C1). We winsorize the R^2 values at the 1st and 99th percentiles to estimate the model using data from 1979 to 2020, and use the values at the end of our sample in December 2020 to calculate active perceived skill and active fund size for our cross-sectional regressions.

Table C4 shows the parameter estimates. To compare these estimates to those in Table 4 in the main text, the parameter estimates for the prior mean, the prior standard deviation, and the residual standard deviation need to be multiplied by the active share. The average active share in the data is 13%. The estimate close to zero for γ again suggests that the log functional form of the cost function fits the data well. As before, we find a parameter estimate, η, that is significantly positive, indicating decreasing returns to scale in actual fund returns.

Table C5 reproduces Table 5 in the main text based on Equation (C1). If the rational expectations model was the model analysts use to form their expectations, the coefficient estimates should be 1 on active share times perceived skill, -1 on active fund size $\times$ η, and -1 on fees. As before, the coefficient estimates on scale, this time measured as actively managed size $\times$ η, are significantly positive and significantly different from the model-implied coefficient estimate of -1.

Table C6 reproduces Table 6 in the main text. The coefficient estimates on actively managed size are significantly positive in all specifications.

Table C4: Parameter estimates of the rational fund performance model with indexing

Parameter	Description	Estimate
η	Decreasing returns to scale (DRS) (%)	0.115***
		(0.005)
γ	Shape of DRS ($\times$ 10^6)	0.004
		(0.007)
a_0	Prior mean (%)	19.369***
		(0.382)
$\sigma_{a,0}$	Prior standard deviation (%)	44.614***
		(0.627)
σ_ϵ	Residual standard deviation (%)	116.533***
		(0.278)
ρ	Skill persistence	0.845***
		(0.008)

The table shows the parameter estimates of the rational fund performance model with indexing in % per year. Standard errors are shown in parentheses. The model is estimated using fund-year observations from 1979 to 2020. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

Table C5: Cross-sectional regressions of alphas on fund characteristics with indexing

	Analyst Ratings	Analyst and Quantitative Ratings
	(1)	(2)
Perceived skill × h	0.130***	0.094***
	(0.037)	(0.027)
	[0.000]	[0.000]
Active fund size × η	1.002***	1.914***
	(0.248)	(0.178)
	[0.000]	[0.000]
Fees	−0.888***	−1.565***
	(0.166)	(0.063)
	[0.501]	[0.000]
Constant (× 100)	0.862***	0.203*
	(0.212)	(0.113)
	[0.000]	[0.073]
N	1451	13627
Adj. R^2	0.13	0.26

The table shows regressions of Morningstar analyst alphas on skill as perceived by a rational learner, active fund size × η, and fees following Equation (C1) for cross-sections of funds in December 2020. Active fund size is measured as $\frac{(h\text{AUM})^{\gamma}-1}{\gamma}$, where AUM refers to a fund's total assets under management in millions of USD, h refers to a fund's active share, and γ is given in Table C4. Specification (1) uses funds with an Analyst Rating. Specification (2) uses funds with an Analyst Rating or a Quantitative Rating. Alphas are relative to each fund's Morningstar Category benchmark. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient. In brackets are p-values for the null hypothesis that the coefficients of skill, size × η, fees, and the constant equal the model-predicted parameters of +1, −1, −1, and 0, respectively.

Table C6: Cross-sectional regressions of alphas on additional fund characteristics with indexing

	Analyst Ratings		Analyst and Quantitative Ratings	
	(1)	(2)	(3)	(4)
Rational learner				
Perceived skill $\times h$	0.023	−0.018	0.194***	0.055*
	(0.038)	(0.043)	(0.047)	(0.033)
Active fund size $\times$ η	2.035***	1.044***	2.075***	0.924***
	(0.278)	(0.282)	(0.216)	(0.215)
Fees	−1.301***	−0.924***	−1.862***	−0.912***
	(0.153)	(0.109)	(0.220)	(0.191)
People				
Manager tenure		0.106***		0.264***
		(0.037)		(0.034)
Manager ownership		0.098**		0.193***
		(0.048)		(0.042)
Managerial multitasking		0.670***		0.548***
		(0.206)		(0.175)
Management team		0.084		0.439***
		(0.105)		(0.102)
Process				
Top 10 assets (%)		0.131		−0.048
		(0.125)		(0.098)
Tracking error		−0.012		−0.155
		(0.064)		(0.100)
Turnover ratio		−0.472***		−0.104
		(0.154)		(0.077)
Retail		−0.276***		−0.148*
		(0.092)		(0.087)
Broker-sold		−0.296**		−0.094
		(0.120)		(0.115)
N	697	648	2808	2603
Adj. R^2	0.26	0.63	0.21	0.63
Sustainability FE	No	Yes	No	Yes
Star FE	No	Yes	No	Yes
Morningstar Category FE	No	Yes	No	Yes
Fund Family FE	No	Yes	No	Yes

The table shows regressions of Morningstar analyst alphas on fund and manager characteristics for cross-sections of funds in December 2020. Specifications (1) and (2) use U.S.-domiciled funds with an Analyst Rating. Specifications (3) and (4) use U.S.-domiciled funds with an Analyst Rating or a Quantitative Rating. Alphas are relative to each fund's Morningstar Category benchmark. Manager tenure is the maximum tenure (in months) taken over all managers, manager ownership is the average amount managers of a fund personally invest in the fund, managerial multitasking is the average number of additional funds that managers of a particular fund manage, and management team is a dummy for team-managed funds. Top 10 assets is the percentage of AUM in the ten largest positions, tracking error is the standard deviation of returns in excess of the benchmark over the life of the fund, turnover is a fund's trading activity as reported to the SEC, retail is a dummy for whether a fund is primarily held by retail investors, and broker-sold is a dummy for whether a fund is primarily sold through brokers. "People" and "Process" variables are standardized to zero mean and unit standard deviation (except for the dummy variables), and the coefficient estimates are multiplied by 100. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

C.2.2 Fund turnover

An alternative measure of activeness is a fund's turnover. To incorporate the idea that funds that trade more face steeper decreasing returns to scale in the model, we add the interaction of turnover and size as well as base effects to the measurement equation. We construct the turnover variable as done by Pástor et al. (2015). The new measurement equation reads

$$r_{i,t+1} + f_{i,t+1} = a_{i,t} - \eta \log(\text{AUM}_{i,t}) - \theta \text{Turnover}_{i,t} - \lambda \log(\text{AUM}_{i,t}) \times \text{Turnover}_{i,t} + \epsilon_{i,t+1}. \tag{C4}$$

We re-estimate the model using this measurement equation.

Table C7 presents parameter estimates for this model. As economic intuition and previous research suggest, the interaction between size and turnover is positive, showing that funds with higher turnover face steeper decreasing returns to scale. Note that, as in the main text, the measurement equation has negative signs in front of the coefficients. Thus, positive coefficient estimates for, say, size in the tables imply the presence of decreasing returns to scale in actual fund returns.

As in the main text, we then simply recompute perceived skill as implied by the model in December 2020 and run cross-sectional regressions of analyst alphas on the fund characteristics implied by the model. Table C8 shows the results.

Consistent with our main results, the coefficient estimate on size $\times\, \eta$ remains positive and far away from the -1 estimate implied by the model. The point estimate on the interaction between turnover and size is negative (-0.038 in specification (1)), but is not statistically different from zero and is economically small. Even for extreme turnover values of around 200% per year, corresponding to the 99th percentile of the turnover distribution, the marginal effect of size on analyst alphas does not become negative (not tabulated).

Table C7: Parameter estimates of the rational fund performance model with turnover

Parameter	Description	Estimate
η	Decreasing returns to scale (%)	0.240***
		(0.021)
a_0	Prior mean (%)	2.270***
		(0.107)
θ	Turnover (%)	−0.706***
		(0.070)
λ	Turnover × Size (%)	0.086***
		(0.015)
$\sigma_{a,0}$	Prior standard deviation (%)	2.017***
		(0.058)
σ_ϵ	Residual standard deviation (%)	8.330***
		(0.021)
ρ	Skill persistence	0.953***
		(0.009)

The table shows the parameter estimates of the rational fund performance model with turnover in % per year. Standard errors are shown in parentheses. The model is estimated using fund-year observations from 1979 to 2020. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

Table C8: Cross-sectional regressions of alphas on fund characteristics with turnover

	Analyst Ratings	Analyst and Quantitative Ratings
	(1)	(2)
Perceived skill	0.269***	0.790***
	(0.072)	(0.067)
	[0.000]	[0.002]
Size $\times$ η	0.487*	0.602***
	(0.281)	(0.130)
	[0.000]	[0.000]
Fees	−0.776***	−1.261***
	(0.241)	(0.123)
	[0.354]	[0.034]
Turnover ($\times$ 100)	−0.351	0.059
	(0.703)	(0.152)
	[0.135]	[0.000]
Turnover $\times$ Size ($\times$ 100)	−0.038	−0.055
	(0.100)	(0.034)
	[0.628]	[0.359]
Constant ($\times$ 100)	0.011	−1.778***
	(0.488)	(0.245)
	[0.982]	[0.000]
N	836	5060
Adj. R^2	0.16	0.28

The table shows regressions of Morningstar analyst alphas on skill as perceived by a rational learner, fund size (logarithm of assets under management in millions of USD) $\times$ η, and fees for cross-sections of funds in December 2020. Specification (1) uses funds with an Analyst Rating. Specification (2) uses funds with an Analyst Rating or a Quantitative Rating. Perceived skill is estimated using a rational model including turnover. Alphas are relative to each fund's Morningstar Category benchmark. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient. In brackets are p-values for the null hypothesis that the coefficients of skill, size $\times$ η, fees, turnover, turnover times size, and the constant equal the model-predicted parameters of +1, −1, −1, 0.706 (the estimate of $-\theta$ in Table C7), −0.086 and 0, respectively.

C.2.3 Small-cap fund indicator

Economic intuition also suggests that funds trading less liquid stocks may face steeper decreasing returns to scale. To capture this idea, as with fund turnover, we add a small-cap fund indicator to the measurement equation together with an interaction of the small-cap fund indicator and fund size. As usual, we then re-estimate the model and our cross-sectional regressions.

Funds investing in small-cap stocks do face steeper decreasing returns to scale in actual fund returns (see Table C9), but this pattern is not mirrored in analysts' expectations: in Table C10, the interaction between the small-cap fund indicator and size is positive.

Overall, we find little evidence that analysts expect returns to scale to vary across funds according to fund characteristics that have been used in the literature. This stands in contrast to the heterogeneity in decreasing returns to scale that has been documented using actual fund returns.

Table C9: Parameter estimates of the rational fund performance model with a small-cap indicator variable

Parameter	Description	Estimate
η	Decreasing returns to scale (%)	0.186^{***}
		(0.013)
a_0	Prior mean (%)	1.818^{***}
		(0.063)
θ	SmlCap (%)	-3.818^{***}
		(0.188)
λ	SmlCap $\times$ Size (%)	0.431^{***}
		(0.038)
$\sigma_{a,0}$	Prior standard deviation (%)	1.957^{***}
		(0.044)
σ_{ϵ}	Residual standard deviation (%)	8.117^{***}
		(0.015)
ρ	Skill persistence	0.937^{***}
		(0.008)

The table shows the parameter estimates of the rational fund performance model with a small-cap indicator variable in % per year. Standard errors are shown in parentheses. The model is estimated using fund-year observations from 1979 to 2020. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

Table C10: Cross-sectional regressions of alphas on fund characteristics with a small-cap indicator variable

	Analyst Ratings	Analyst and Quantitative Ratings
	(1)	(2)
Perceived skill	0.444***	0.821***
	(0.078)	(0.047)
	[0.000]	[0.000]
Size × η	0.305*	0.712***
	(0.181)	(0.135)
	[0.000]	[0.000]
Fees	−0.948***	−1.536***
	(0.150)	(0.060)
	[0.731]	[0.000]
SmlCap (× 100)	−0.885	−0.078
	(0.540)	(0.218)
	[0.000]	[0.000]
SmlCap × Size (× 100)	0.131*	0.061
	(0.075)	(0.038)
	[0.000]	[0.000]
Constant (× 100)	0.178	−1.437***
	(0.268)	(0.174)
	[0.508]	[0.000]
N	1454	13934
Adj. R^2	0.15	0.32

The table shows regressions of Morningstar analyst alphas on skill as perceived by a rational learner, fund size (logarithm of assets under management in millions of USD) × η, and fees for cross-sections of funds in December 2020. Specification (1) uses funds with an Analyst Rating. Specification (2) uses funds with an Analyst Rating or a Quantitative Rating. Perceived skill is estimated using a rational model including a small-cap indicator variable. Alphas are relative to each fund's Morningstar Category benchmark. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient. In brackets are p-values for the null hypothesis that the coefficients of skill, size × η, fees, small-cap, small-cap times size, and the constant equal the model-predicted parameters of +1, −1, −1, 3.818 (the estimate of $-\theta$ in Table C9), −0.431, and 0, respectively.

C.3 Time-variation and uncertainty regarding the decreasing returns to scale parameter

In this subsection, we modify our baseline model and allow for uncertainty about the decreasing returns to scale parameter in addition to uncertainty about managerial skill. The measurement equation is the same as before (omitting fund i subscripts):

$$r_{t+1} + f_{t+1} = \begin{pmatrix} 1 & -c(\text{AUM}_t) \end{pmatrix} \begin{pmatrix} a_t \\ \eta_t \end{pmatrix} + \epsilon_{t+1}, \tag{C5}$$

where $c(\text{AUM}_t) = \log(\text{AUM}_t)$ as the decreasing returns to scale parameter, η_t, is now included in the vector of state variables. With a decreasing returns to scale parameter that varies over time following an AR(1), the state transition equations are:

$$\begin{pmatrix} a_t \\ \eta_t \end{pmatrix} = \begin{pmatrix} \rho & 0 \\ 0 & q \end{pmatrix} \begin{pmatrix} a_{t-1} \\ \eta_{t-1} \end{pmatrix} + \begin{pmatrix} 1-\rho & 0 \\ 0 & 1-q \end{pmatrix} \begin{pmatrix} a_0 \\ \eta_0 \end{pmatrix} + \begin{pmatrix} \sqrt{(1-\rho^2)} & 0 \\ 0 & \sqrt{(1-q^2)} \end{pmatrix} \nu_t, \tag{C6}$$

where $q \in [0,1]$ captures persistence in scalability, and $\nu_t \sim N(0,V)$ is a 2×1 column vector of correlated shocks following a multivariate normal distribution with mean zero and variance-covariance matrix

$$V = \begin{pmatrix} \sigma_{a,0}^2 & \sigma_{a,0}\sigma_{\eta,0}\rho_{\eta_0,a_0} \\ \sigma_{a,0}\sigma_{\eta,0}\rho_{\eta_0,a_0} & \sigma_{\eta,0}^2 \end{pmatrix}. \tag{C7}$$

Skill and scale when a fund is born are distributed as $N((a_0, \eta_0)^T, V)$. Relative to the baseline model, the new parameters are persistence in scalability (q), the prior mean of the decreasing returns to scale parameter (η_0), the prior standard deviation of the decreasing returns to scale parameter ($\sigma_{\eta,0}$), and the correlation between the prior mean of skill and the prior mean of scalability (ρ_{η_0,a_0}).

The updating equations for managerial skill and the decreasing returns to scale parameter again follow directly from the Kalman filter. After solving the model, we again estimate the model using an Expectation Maximization algorithm (see, e.g., Dempster, Laird, and Rubin, 1977; Shumway and Stoffer, 1982; Watson and Engle, 1983).

With uncertainty about the decreasing returns to scale parameter, the decreasing returns to scale parameter is naturally different for every individual fund, just like managerial skill is. That, is as opposed to assuming before-fee alphas are given by $\alpha_i(q_i) = a_i - b \times q_i$, we now assume that before-fee alphas are given by $\alpha_i(q_i) = a_i - b_i \times q_i$. This assumption is similar to the assumption in Barras, Gagliardini, and Scaillet (2022), except that we use an empirical Bayes approach to estimate a_i and b_i as opposed to their purely frequentist fund-by-fund regressions.

Table C11 shows the parameter estimates. The parameter estimate for the prior mean of the decreasing returns to scale parameter is 0.22 with a standard deviation of 0.86. As in Barras et al. (2022), skill, a_0, and scale, η_0, are strongly positively correlated. The correlation is 0.997.

Similar to the main text, we then estimate cross-sectional regressions of analyst alphas on perceived skill, $\text{size}_{i,t} \times \widehat{\eta}_{i,t}$, and fees. Importantly, $\widehat{\eta}$ now captures perceived scalability at a given point in time for a given fund. Under the null that analysts form their expectations as in this rational model, the coefficient estimate on the product of size and perceived scalability is -1.

Table C12 shows the results of these cross-sectional regressions. The estimated coefficients on $\text{size}_{i,t} \times \widehat{\eta}_{i,t}$ are always positive, though for the sample of only analyst-rated funds not significantly different from zero. That said, all the coefficient estimates are significantly different from the -1 coefficient under the null of the rational model. The p-values for these tests are 0.00 throughout.

Similar to the main text, Panel (a) of Figure C2 shows that before-fee alphas derived from the model that allows for time-variation and uncertainty regarding the decreasing returns to scale parameter are notably shifted to the left relative to analyst before-fee alphas for the largest decile of funds with an Analyst Rating, the funds that have grown to be the largest. The figure illustrates before-fee alphas to rule out that cross-sectional differences in fees drive the results. For comparison, Figure C3 corresponds to Figure 1 in the main text, except that it shows before-fee as opposed to net-of-fee alphas.

Another way to see that analyst alphas for the largest funds appear too large is to consider the value added measure of Berk and van Binsbergen (2015), the product of the before-fee alpha and AUM. Panel (b) of Figure C2 illustrates a mismatch between realized value added and value added as implied by analysts' expectations. Intuitively, since realized value added—as opposed to realized alpha—is well known to be highly persistent (see, e.g., Berk and van Binsbergen, 2015; Gerakos, Linnainmaa, and Morse, 2021), one would expect a rational forecast of value added to be close to past realized value added. Indeed, as one would expect of a rational forecast, the distribution of model-implied value added matches the distribution of realized value added closely (these value added results are the same for the baseline model in the main text).

A similar conclusion arises from Figure C4, which compares to Figure 3 in the main text. As in the main text, analyst alphas are too large relative to model-implied expectations for the largest funds, including all funds with an Analyst Rating, but too small for the vast majority of other funds, including most funds with a Quantitative Rating.

Table C11: Parameter estimates of the rational fund performance model with time-variation and uncertainty regarding the decreasing returns to scale parameter

Parameter	Description	Estimate
η_0	Prior mean decreasing returns to scale (DRS) (%)	0.222***
		(0.013)
$\sigma_{\eta,0}$	Prior standard deviation of DRS (%)	0.858***
		(0.024)
a_0	Prior skill mean (%)	2.281***
		(0.071)
$\sigma_{a,0}$	Prior skill standard deviation (%)	6.266***
		(0.143)
σ_ϵ	Residual standard deviation (%)	8.151***
		(0.018)
ρ	Skill persistence	0.926***
		(0.007)
q	Scale persistence	0.931***
		(0.008)
ρ_{η_0,a_0}	Correlation prior skill mean and prior mean DRS	0.997***
		(0.002)

The table shows the parameter estimates of the rational fund performance model with time-variation and uncertainty about the decreasing returns to scale parameter in % per year. Standard errors are shown in parentheses. The model is estimated using fund-year observations from 1979 to 2020. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

Table C12: Cross-sectional regressions of alphas on fund characteristics with time-variation and uncertainty regarding the decreasing returns to scale parameter

	Analyst Ratings	Analyst and Quantitative Ratings
	(1)	(2)
Perceived skill	−0.021	0.022
	(0.085)	(0.028)
	[0.000]	[0.000]
Size $\times$ η	0.080	0.320***
	(0.085)	(0.042)
	[0.000]	[0.000]
Fees	−0.934***	−1.632***
	(0.170)	(0.061)
	[0.696]	[0.000]
Constant ($\times$ 100)	1.472***	0.417***
	(0.180)	(0.110)
	[0.000]	[0.000]
N	1454	13934
Adj. R^2	0.07	0.29

The table shows regressions of Morningstar analyst alphas on skill as perceived by a rational learner, fund size (logarithm of assets under management in millions of USD) $\times$ η, and fees for cross-sections of funds in December 2020. Specification (1) uses funds with an Analyst Rating. Specification (2) uses funds with an Analyst Rating or a Quantitative Rating. Alphas are relative to each fund's Morningstar Category benchmark. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient. In brackets are p-values for the null hypothesis that the coefficients of skill, size $\times$ η, fees, and the constant equal the model-predicted parameters of +1, −1, −1, and 0, respectively.

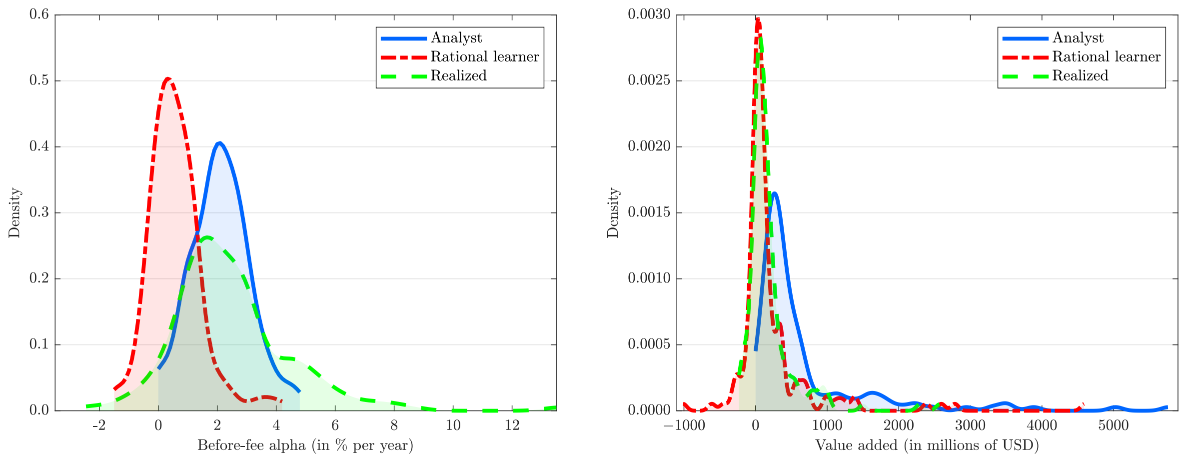
Table C13: Cross-sectional regressions of alphas on additional fund characteristics with time-variation and uncertainty regarding the decreasing returns to scale parameter

	Analyst Ratings		Analyst and Quantitative Ratings	
	(1)	(2)	(3)	(4)
Rational learner				
Perceived skill	−0.142	−0.187*	0.197***	0.006
	(0.151)	(0.099)	(0.056)	(0.041)
Size × η	0.120	0.171*	0.101	0.098**
	(0.130)	(0.088)	(0.084)	(0.046)
Fees	−1.520***	−0.967***	−2.104***	−0.991***
	(0.173)	(0.110)	(0.204)	(0.200)
People				
Manager tenure		0.133***		0.293***
		(0.042)		(0.034)
Manager ownership		0.114**		0.215***
		(0.057)		(0.042)
Managerial multitasking		0.651***		0.577***
		(0.198)		(0.189)
Management team		0.068		0.489***
		(0.116)		(0.115)
Process				
Top 10 assets (%)		0.150		−0.036
		(0.126)		(0.092)
Tracking error		0.039		−0.117
		(0.056)		(0.094)
Turnover ratio		−0.550***		−0.143*
		(0.157)		(0.081)
Retail		−0.276***		−0.142
		(0.092)		(0.090)
Broker-sold		−0.295**		−0.095
		(0.115)		(0.109)
N	698	650	2830	2626
Adj. R^2	0.16	0.62	0.21	0.63
Sustainability FE	No	Yes	No	Yes
Star FE	No	Yes	No	Yes
Morningstar Category FE	No	Yes	No	Yes
Fund Family FE	No	Yes	No	Yes

The table shows regressions of Morningstar analyst alphas on fund and manager characteristics for cross-sections of funds in December 2020. Specifications (1) and (2) use U.S.-domiciled funds with an Analyst Rating. Specifications (3) and (4) use U.S.-domiciled funds with an Analyst Rating or a Quantitative Rating. Alphas are relative to each fund's Morningstar Category benchmark. Manager tenure is the maximum tenure (in months) taken over all managers, manager ownership is the average amount managers of a fund personally invest in the fund, managerial multitasking is the average number of additional funds that managers of a particular fund manage, and management team is a dummy for team-managed funds. Top 10 assets is the percentage of AUM in the ten largest positions, tracking error is the standard deviation of returns in excess of the benchmark over the life of the fund, turnover is a fund's trading activity as reported to the SEC, retail is a dummy for whether a fund is primarily held by retail investors, and broker-sold is a dummy for whether a fund is primarily sold through brokers. "People" and "Process" variables are standardized to zero mean and unit standard deviation (except for the dummy variables), and the coefficient estimates are multiplied by 100. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

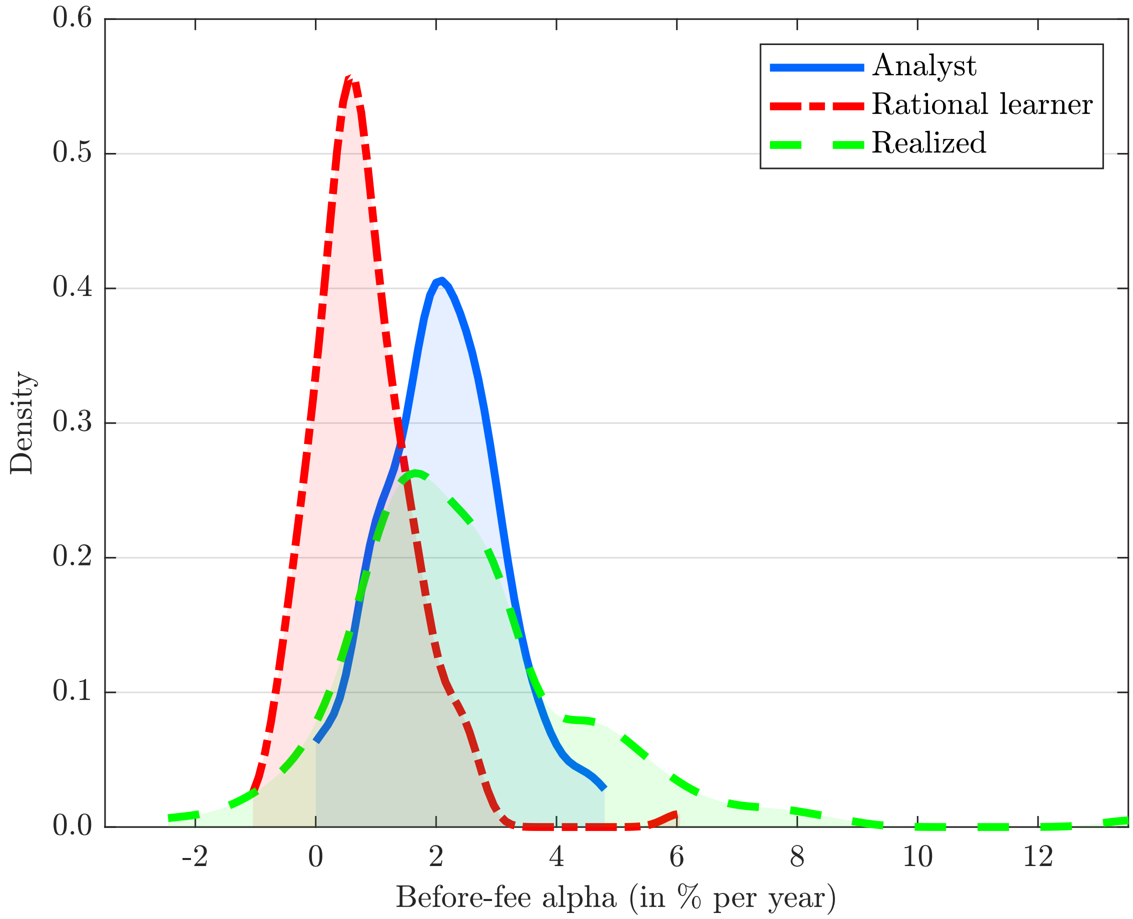
Figure C2: Before-fee alphas and value added of the ten percent largest analyst-rated funds with time-variation and uncertainty regarding the decreasing returns to scale parameter

(a) Before-fee alpha **(b)** Value added



Panel (a) shows the cross-sectional distributions of analyst before-fee alphas (in blue) and before-fee alphas implied by a rational expectations learning model (in red), as well as backward-looking historically realized before-fee alphas (in green), all as of December 2020. Realized alphas are computed over the lifetime of a fund. Panel (b) shows expected value added, which is computed as the product of before-fee alphas and assets under management (AUM) in December 2020 together with historically realized value added, which is the average of the product of annual before-fee realized alphas times lagged AUM over the lifetime of a fund. The sample is restricted to the ten percent largest funds with an Analyst Rating as of December 2020. On average, these 145 funds have existed for 30 years and grown their assets under management (AUM) from USD 1 billion to USD 30 billion, managing about 30% of worldwide AUM in the active equity mutual fund industry as of December 2020. Alphas are relative to each fund's Morningstar Category benchmark. The rational learner's expectations are derived from a model that allows for time-variation and uncertainty regarding both managerial skill and the decreasing returns to scale parameter.

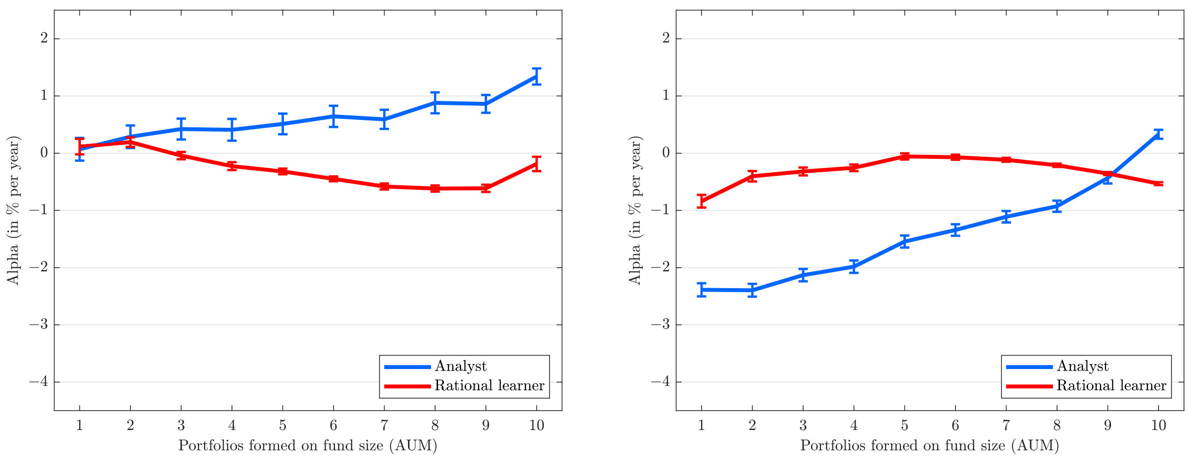
Figure C3: Before-fee alphas of the ten percent largest analyst-rated funds



The figure shows the cross-sectional distributions of analyst before-fee alphas (in blue) and before-fee alphas implied by a rational expectations learning model (in red), as well as backward-looking historically realized before-fee alphas (in green), all as of December 2020. Realized alphas are computed over the lifetime of a fund. The sample is restricted to the ten percent largest funds with an Analyst Rating as of December 2020. On average, these 145 funds have existed for 30 years and grown their assets under management (AUM) from USD 1 billion to USD 30 billion, managing about 30% of worldwide AUM in the active equity mutual fund industry as of December 2020. Alphas are relative to each fund's Morningstar Category benchmark.

Figure C4: Alphas against fund size

(a) Analyst Ratings **(b)** Analyst and Quantitative Ratings



The figure shows expected net-of-fee abnormal returns (alphas) against fund size (AUM) as of December 2020 for analysts (in blue) and for a rational learner (in red). Panel (a) includes funds with an Analyst Rating. Panel (b) includes funds with an Analyst Rating or a Quantitative Rating. Alphas are relative to each fund's Morningstar Category benchmark. The bars indicate 90% confidence bands. The rational learner's expectations are derived from a model that allows for time-variation and uncertainty regarding both managerial skill and the decreasing returns to scale parameter.

Table C14: Parameter estimates of the rational fund performance model with industry size

Parameter	Description	Estimate
η	Decreasing returns to scale (%)	0.248***
		(0.013)
θ	Industry size (%)	13.543***
		(1.881)
a_0	Prior mean (%)	3.635***
		(0.214)
$\sigma_{a,0}$	Prior standard deviation (%)	2.166***
		(0.040)
σ_ϵ	Residual standard deviation (%)	7.754***
		(0.015)
ρ	Skill persistence	0.962***
		(0.005)

The table shows the parameter estimates of the rational fund performance model with industry size in % per year. Standard errors are shown in parentheses. The model is estimated using fund-year observations from 2000 to 2020. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

C.4 Industry size

Pástor et al. (2015) show that industry size is a determinant of fund returns. We include industry size in our baseline measurement equation and re-estimate the structural model to obtain another measure of perceived skill. The resulting parameter estimates are shown in Table C14. Consistent with their results, we find a positive coefficient for industry size, showing that an increase in industry size decreases fund returns. We compute perceived skill in December 2020 according to this model and then re-run our cross-sectional regressions. As shown in Table C15, we obtain results similar to those in the main text. Note that industry size drops out as a regressor in our cross-sectional regressions since it is constant for a given cross-section (it does, however, affect the measure of perceived skill in December 2020).

Table C15: Cross-sectional regressions of alphas on fund characteristics with industry size in estimation of the rational fund performance model

	Analyst Ratings	Analyst and Quantitative Ratings
	(1)	(2)
Perceived skill	0.354***	0.608***
	(0.058)	(0.036)
	[0.000]	[0.000]
Size × η	0.260**	0.543***
	(0.129)	(0.100)
	[0.000]	[0.000]
Fees	−0.985***	−1.537***
	(0.148)	(0.059)
	[0.920]	[0.000]
Constant (× 100)	−0.362	−2.173***
	(0.299)	(0.184)
	[0.228]	[0.000]
N	1454	13934
Adj. R^2	0.16	0.32

The table shows regressions of Morningstar analyst alphas on skill computed as perceived by a rational learner, fund size (logarithm of assets under management in millions of USD) × η, and fees for cross-sections of funds in December 2020. Specification (1) uses funds with an Analyst Rating. Specification (2) uses funds with an Analyst Rating or a Quantitative Rating. Alphas are relative to each fund's Morningstar Category benchmark. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient. In brackets are p-values for the null hypothesis that the coefficients of skill, size × η, fees, and the constant equal the model-predicted parameters of +1, −1, −1, and 0, respectively.

C.5 Value added

While an important benchmark, one caveat to the rational expectations learning model is that the resulting measure of perceived skill depends on the assumed functional form of the decreasing returns to scale technology. Berk and van Binsbergen (2015) propose value added as measure of skill. In contrast to the measure of perceived skill from the rational expectations learning model, value added does not depend on the assumed functional form of the decreasing returns to scale technology. Perhaps analysts use value added as a measure of skill, so that once value added is controlled for, the coefficient estimate on size becomes negative. This would be broadly consistent with the rational expectations paradigm for mutual funds.

Table C16 presents regressions of analyst alphas on value added as defined by Berk and van Binsbergen (2015), size, and fees. While value added does not depend on the decreasing returns to scale technology, our cross-sectional regressions of course still need to assume a functional form between analyst alphas and AUM. In specification (1), as in most of our other analyses, we assume a log-linear functional form (which, as shown above, describes the relationship between *actual* returns and size well). Perhaps analysts believe in decreasing returns to scale, but use a different functional form. In specification (2), we assume a linear functional form. Interestingly, the point estimate in (2) is not statistically different from zero—however, it is still far from the decreasing returns to scale estimates in our various models. Specifications (3) and (4) include the value added estimated over the last 10% of observations to account for possible slow learning by investors, as suggested by Barras et al. (2022). Specifications (5) to (8) include funds with a Quantitative Rating.

In short, the coefficient estimate of size on returns is positive in all specifications, so our conclusions remain unchanged when we control for value added rather than perceived skill as implied by the rational expectations learning model.

Table C16: Cross-sectional regressions of alphas on fund characteristics—value added

	Analyst Ratings				Analyst and Quantitative Ratings			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Value added (× 100)	0.006***	0.007***			0.014***	0.021***		
	(0.001)	(0.001)			(0.002)	(0.002)		
Value added 10% (× 100)			0.001**	0.001***			0.005***	0.006***
			(0.000)	(0.000)			(0.001)	(0.001)
Size (× 100)	0.066*		0.120***		0.210***		0.228***	
	(0.034)		(0.031)		(0.025)		(0.025)	
Size, linear (× 10^6)		0.028		0.086***		0.078		0.230***
		(0.020)		(0.023)		(0.072)		(0.084)
Fees	−0.808***	−0.871***	−0.742***	−0.827***	−1.442***	−1.606***	−1.445***	−1.622***
	(0.167)	(0.169)	(0.178)	(0.177)	(0.064)	(0.062)	(0.065)	(0.062)
Constant (× 100)	0.855***	1.359***	0.493*	1.385***	−0.457**	0.763***	−0.519***	0.804***
	(0.280)	(0.160)	(0.278)	(0.165)	(0.188)	(0.106)	(0.188)	(0.108)
N	1449	1449	1449	1449	13580	13580	13580	13580
Adj. R^2	0.12	0.11	0.10	0.09	0.27	0.25	0.27	0.24

The table shows regressions of Morningstar analyst alphas on value added as defined by Berk and van Binsbergen (2015), fund size, and fees for cross-sections of funds in December 2020. Specifications (1) to (4) use funds with an Analyst Rating. Specifications (5) to (8) use funds with an Analyst Rating or a Quantitative Rating. Specifications (3), (4), (7), and (8) compute value added over the last 10% of observations to account for possible slow learning by investors. Specifications (1), (3), (5), and (7) use a log-linear functional form of decreasing returns to scale, and specifications (2), (4), (6), and (8) use a linear functional form. Alphas are relative to each fund's Morningstar Category benchmark. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

C.6 Estimation by Global Category

We also estimate our baseline rational expectations model by Global Category, as the impact of size on returns may vary across categories. We do not report summary statistics of the parameter estimates, but simply recalculate perceived skill at the end of our sample in December 2020 using the parameter estimates that vary by Global Category, and rerun our main regressions in Tables 5 and 6 in the main text. Tables C17 and C18 show the results. The coefficient estimates on size are significantly positive in the specifications including all control variables and fixed effects.

Table C17: Cross-sectional regressions of alphas on fund and manager characteristics with estimation of the rational fund performance model by Global Category

	Analyst Ratings	Analyst and Quantitative Ratings
	(1)	(2)
Perceived skill	0.167***	0.194***
	(0.044)	(0.026)
	[0.000]	[0.000]
Size × η	−0.063	0.053
	(0.040)	(0.036)
	[0.000]	[0.000]
Fees	−1.125***	−1.693***
	(0.143)	(0.060)
	[0.384]	[0.000]
Constant (× 100)	1.461***	−0.536***
	(0.155)	(0.111)
	[0.000]	[0.000]
N	1454	13934
Adj. R^2	0.11	0.25

The table shows regressions of Morningstar analyst alphas on skill as perceived by a rational learner, fund size × η, and fees for cross-sections of funds in December 2020. Specification (1) uses funds with an Analyst Rating under the new methodology. Specification (2) uses funds with an Analyst Rating or a Quantitative Rating. Alphas are relative to each fund's Morningstar Category benchmark. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient. In brackets are p-values for the null hypothesis that the coefficients of skill, size × η, fees, and the constant equal the model-predicted parameters of +1, −1, −1, and 0, respectively.

Table C18: Cross-sectional regressions of alphas on additional characteristics with estimation of the rational fund performance model by Global Category

	Analyst Ratings		Analyst and Quantitative Ratings	
	(1)	(2)	(3)	(4)
Rational learner				
Perceived skill	0.152***	−0.023	0.385***	0.089**
	(0.055)	(0.052)	(0.042)	(0.037)
Size × η	−0.064	0.395***	−0.204***	0.225***
	(0.051)	(0.092)	(0.051)	(0.063)
Fees	−1.782***	−0.933***	−2.377***	−0.980***
	(0.234)	(0.112)	(0.213)	(0.200)
People				
Manager tenure		0.114***		0.276***
		(0.039)		(0.034)
Manager ownership		0.123**		0.204***
		(0.052)		(0.042)
Managerial multitasking		0.650***		0.577***
		(0.200)		(0.192)
Management team		0.097		0.476***
		(0.107)		(0.112)
Process				
Top 10 assets (%)		0.136		−0.020
		(0.127)		(0.093)
Tracking error		0.017		−0.070
		(0.056)		(0.104)
Turnover ratio		−0.486***		−0.119
		(0.146)		(0.079)
Retail		−0.251***		−0.154*
		(0.088)		(0.090)
Broker-sold		−0.308***		−0.078
		(0.114)		(0.105)
N	698	650	2830	2626
Adj. R^2	0.19	0.63	0.19	0.63
Sustainability FE	No	Yes	No	Yes
Star FE	No	Yes	No	Yes
Morningstar Category FE	No	Yes	No	Yes
Fund Family FE	No	Yes	No	Yes

The table shows regressions of Morningstar analyst alphas on fund and manager characteristics for cross-sections of funds in December 2020. Specifications (1) and (2) use U.S.-domiciled funds with an Analyst Rating. Specifications (3) and (4) use U.S.-domiciled funds with an Analyst Rating or a Quantitative Rating. Alphas are relative to each fund's Morningstar Category benchmark. Manager tenure is the maximum tenure (in months) taken over all managers, manager ownership is the average amount managers of a fund personally invest in the fund, managerial multitasking is the average number of additional funds that managers of a particular fund manage, and management team is a dummy for team-managed funds. Top 10 assets is the percentage of AUM in the ten largest positions, tracking error is the standard deviation of returns in excess of the benchmark over the life of the fund, turnover is a fund's trading activity as reported to the SEC, retail is a dummy for whether a fund is primarily held by retail investors, and broker-sold is a dummy for whether a fund is primarily sold through brokers. "People" and "Process" variables are standardized to zero mean and unit standard deviation (except for the dummy variables), and the coefficient estimates are multiplied by 100. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

Table C19: Parameter estimates of the rational fund performance model restricted to funds incepted since 2000

Parameter	Description	Estimate
η	Decreasing returns to scale (%)	0.223***
		(0.018)
a_0	Prior mean (%)	1.981***
		(0.082)
$\sigma_{a,0}$	Prior standard deviation (%)	2.429***
		(0.054)
σ_ϵ	Residual standard deviation (%)	8.394***
		(0.022)
ρ	Skill persistence	0.955***
		(0.009)

The table shows the parameter estimates of the rational fund performance model restricted to funds incepted since 2000 in % per year. Standard errors are shown in parentheses. The model is estimated using fund-year observations from 2000 to 2020. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

C.7 Estimation with funds incepted since 2000

For our baseline analysis, we estimate the rational expectations model using the entire time series of data available since 1979. Since Morningstar only uses data since 2000 to construct the Analyst Ratings, one may worry that the relationships between returns, skill, size, and fees before 2000 may differ from the comparable relationships since 2000, for instance, because of a structural break. Table C19 alleviates such concerns. The table reports parameter estimates when only funds incepted since 2000 are included in the estimation, and the parameter estimates are similar to the ones in Table 4 in the main text. It is important to use only funds incepted since 2000 when splitting the sample, as otherwise funds incepted before 2000 would be assigned a wrong prior mean in the estimation.

Table C20: Parameter estimates of the rational fund performance model with fund returns estimated in rolling window factor regressions

Parameter	Description	Estimate
η	Decreasing returns to scale (%)	0.222***
		(0.013)
a_0	Prior mean (%)	2.069***
		(0.066)
$\sigma_{a,0}$	Prior standard deviation (%)	2.452***
		(0.053)
σ_ϵ	Residual standard deviation (%)	8.066***
		(0.019)
ρ	Skill persistence	0.872***
		(0.012)

The table shows the parameter estimates of the rational fund performance model with fund returns estimated in rolling window factor regressions in % per year. Standard errors are shown in parentheses. The model is estimated using fund-year observations from 1979 to 2020. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

C.8 Estimation with abnormal fund returns estimated in rolling window factor regressions

We estimate our baseline rational expectations model with abnormal fund returns relative to each fund's Morningstar Category benchmark estimated using a single factor regression over the entire life of a fund. This procedure follows Roussanov, Ruan, and Wei (2021), but one concern is that computing abnormal returns this way could create a bias towards finding decreasing returns to scale in actual fund returns similar to the bias that troubles finite-sample fixed effects regressions (see, e.g., Pástor et al., 2015). To alleviate such concerns we also estimate the rational expectations model with abnormal fund returns computed using three-year rolling window factor regressions. Table C20 reports parameter estimates of the rational model, which are similar to the ones in Table 4 in the main text.

C.9 Diagnostic expectations and past returns

This section forms alternative versions of perceived skill that give more weight to recent past performance. We do so to verify that the coefficient estimate on size $\times\ \eta$ in our main regression does not flip when including such alternative measures of perceived skill. If it did, analysts' would excessively weight past returns relative to the rational benchmark, but not necessarily misunderstand decreasing returns to scale.

To construct the alternative version of perceived skill, we consider the diagnostic Kalman filter of Bordalo, Gennaioli, La Porta, and Shleifer (2019). The diagnostic Kalman filter allows a learner to exaggerate the signal-to-noise ratio relative to the standard Kalman filter and thereby to increase the weight on past returns relative to a rational learner:

$$\widehat{a}_{i,t+1} = \rho \left(\widehat{a}_{i,t} + (1+\theta) \frac{\widehat{\sigma}^2_{a,t}}{\widehat{\sigma}^2_{a,t} + \sigma^2_{\epsilon}} (r_{i,t+1} - \widehat{a}_{i,t} + c(\text{AUM}_{i,t}) + f_{i,t+1}) \right) + (1-\rho) a_0. \quad \text{(C8)}$$

The parameter $\theta \geq 0$ controls the extent of overreaction to information. The larger θ, the more overreaction there is. If $\theta = 0$, the equation collapses to the standard Kalman filter updating equation, i.e., Equation (5) in the paper.

We compute perceived skill at the end of 2020 using our estimated model parameters from Table 4 in the main text and consider θ values ranging from 0 to 2. As mentioned above, as we increase θ, perceived skill loads more on recent returns.

To verify that perceived skill measures that use $\theta > 0$ load more on past returns than the rational benchmark, we regress perceived skill for a given θ on past returns. Table C21 reports the results. At $\theta = 0$ the model reduces to rational learning, our baseline model in the main text. As expected, the coefficient estimates on past performance increase with θ. Thus, we have constructed measures of perceived skill that load more on past performance than the rational benchmark.

Next, we regress analyst alphas on the perceived skill measures computed using the diagnostic Kalman filter. Table C22 shows the results. The estimated coefficients on size $\times\ \eta$ do not turn zero or negative. All coefficient estimates are positive and significantly different from zero.

The regressions in Table 6 in the main text control for Star Rating fixed effects, a comprehensive measure of past performance. We also control for several lags of past performance explicitly. Table C23 shows the results. The estimated coefficients on size $\times\ \eta$ are positive.

Table C21: Cross-sectional regressions of diagnostic perceived skill on past returns

	(1)	(2)	(3)	(4)	(5)	(6)
θ	0.0	0.4	0.8	1.2	1.6	2.0
Net abn. return	0.050***	0.070***	0.089***	0.109***	0.128***	0.147***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Net abn. return (t–1)	0.064***	0.085***	0.104***	0.120***	0.135***	0.148***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Net abn. return (t–2)	0.035***	0.046***	0.057***	0.066***	0.075***	0.082***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Constant (× 100)	2.245***	2.269***	2.285***	2.297***	2.306***	2.313***
	(0.006)	(0.007)	(0.009)	(0.009)	(0.010)	(0.010)
N	11994	11994	11994	11994	11994	11994
Adj. R^2	0.681	0.725	0.763	0.796	0.825	0.849

The table shows regressions of skill as perceived by a diagnostic learner on three lags of annual abnormal net-of-fee returns for cross-sections of funds in December 2020. The sample includes funds with an Analyst Rating or a Quantitative Rating. Parameter θ controls the extent of overreaction to recent returns. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

Table C22: Cross-sectional regressions of alphas on fund characteristics with perceived skill from diagnostic Kalman filter

	(1)	(2)	(3)	(4)	(5)	(6)
θ	0.0	0.4	0.8	1.2	1.6	2.0
Panel A: Analyst Ratings						
Perceived skill	0.395***	0.299***	0.241***	0.200***	0.170***	0.147***
	(0.067)	(0.051)	(0.041)	(0.035)	(0.030)	(0.027)
Size × η	0.321**	0.334**	0.350**	0.368**	0.386***	0.404***
	(0.149)	(0.148)	(0.147)	(0.147)	(0.146)	(0.145)
Fees	−0.962***	−0.953***	−0.941***	−0.926***	−0.911***	−0.897***
	(0.150)	(0.151)	(0.153)	(0.154)	(0.156)	(0.158)
Constant (× 100)	0.062	0.256	0.359	0.420	0.455*	0.477*
	(0.274)	(0.266)	(0.264)	(0.264)	(0.264)	(0.264)
N	1454	1454	1454	1454	1454	1454
Adj. R^2	0.152	0.152	0.150	0.147	0.144	0.140
Panel B: Analyst and Quantitative Ratings						
Perceived skill	0.729***	0.546***	0.440***	0.369***	0.318***	0.279***
	(0.042)	(0.031)	(0.026)	(0.022)	(0.019)	(0.017)
Size × η	0.644***	0.662***	0.684***	0.706***	0.730***	0.753***
	(0.114)	(0.114)	(0.114)	(0.114)	(0.114)	(0.114)
Fees	−1.536***	−1.539***	−1.541***	−1.543***	−1.543***	−1.543***
	(0.059)	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)
Constant (× 100)	−1.509***	−1.131***	−0.921***	−0.791***	−0.704***	−0.645***
	(0.174)	(0.171)	(0.171)	(0.172)	(0.172)	(0.173)
N	13934	13934	13934	13934	13934	13934
Adj. R^2	0.320	0.320	0.319	0.318	0.316	0.314

The table shows regressions of Morningstar analyst alphas on skill as perceived by a diagnostic learner, fund size (logarithm of assets under management in millions of USD) × η, and fees for cross-sections of funds in December 2020. Parameter θ controls the extent of overreaction to recent returns. Panel A uses funds with an Analyst Rating. Panel B uses funds with an Analyst Rating or a Quantitative Rating. Alphas are relative to each fund's Morningstar Category benchmark. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

Table C23: Cross-sectional regressions of alphas on fund characteristics and multiple lags of past returns

	Analyst Ratings				Analyst and Quantitative Ratings			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Perceived skill	0.395***	0.705***	0.566***	0.232	0.729***	0.758***	0.579***	0.081
	(0.067)	(0.116)	(0.119)	(0.153)	(0.042)	(0.074)	(0.088)	(0.172)
Size × η	0.321**	0.206	0.322*	0.513***	0.644***	0.491***	0.462***	0.622***
	(0.149)	(0.158)	(0.165)	(0.175)	(0.114)	(0.121)	(0.119)	(0.139)
Fees	−0.962***	−1.029***	−0.941***	−0.834***	−1.536***	−1.482***	−1.452***	−1.181***
	(0.150)	(0.150)	(0.158)	(0.177)	(0.059)	(0.066)	(0.070)	(0.092)
Net abn. return		−0.032***	−0.024***	−0.007		0.000	0.012**	0.039***
		(0.007)	(0.008)	(0.009)		(0.005)	(0.006)	(0.009)
Net abn. return (t–1)		−0.023*	−0.025**	−0.018		−0.001	0.009	0.024***
		(0.012)	(0.012)	(0.012)		(0.007)	(0.007)	(0.009)
Net abn. return (t–2)		0.022**	0.022**	0.027**		0.050***	0.067***	0.086***
		(0.009)	(0.010)	(0.011)		(0.007)	(0.007)	(0.010)
Net abn. return (t–3)			0.030***	0.040***			0.025***	0.046***
			(0.008)	(0.012)			(0.009)	(0.009)
Net abn. return (t–4)			0.039***	0.052***			0.025***	0.042***
			(0.010)	(0.011)			(0.009)	(0.011)
Net abn. return (t–5)				0.034***				0.016**
				(0.009)				(0.008)
Net abn. return (t–6)				0.020**				0.035***
				(0.010)				(0.009)
Net abn. return (t–7)				−0.001				0.008
				(0.008)				(0.007)
Net abn. return (t–8)				0.010				0.016*
				(0.011)				(0.009)
Net abn. return (t–9)				0.021***				0.036***
				(0.008)				(0.008)
Constant (× 100)	0.062	−0.430	−0.374	−0.007	−1.509***	−1.279***	−0.785***	−0.066
	(0.274)	(0.318)	(0.334)	(0.439)	(0.174)	(0.229)	(0.236)	(0.342)
N	1454	1415	1355	1180	13934	11994	10483	7396
Adj. R^2	0.152	0.189	0.217	0.238	0.320	0.348	0.366	0.376

The table shows regressions of Morningstar analyst alphas on skill as perceived by a rational learner, fund size (logarithm of assets under management in millions of USD) × η, fees, and multiple lags of past returns for cross-sections of funds in December 2020. Specificationa (1)–(4) use funds with an Analyst Rating. Specifications (5)–(8) use funds with an Analyst Rating or a Quantitative Rating. Alphas are relative to each fund's Morningstar Category benchmark. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

D Regressions of pillar scores on fund and manager characteristics

In this section, we regress the three pillar scores "Parent," "People," and "Process" (as opposed to analyst alphas) on fund and manager characteristics. The pillar scores take on the values –2, –1, 0, +1, and +2, and we simply run linear regressions.

We are particularly interested in the effect of size on the pillar scores. Perhaps the positive relationship between analyst alphas and fund size is due to a positive relationship between one particular pillar (e.g., the "Process" pillar) and size.

However, Table D1 shows that this is not the case. The estimate on size $\times\ \eta$ is significantly positive and similar in magnitude for all pillars. Larger funds receive higher "Parent" scores, higher "People" scores, and higher "Process" scores.

Table D1: Cross-sectional regressions of pillar scores on fund characteristics

	Analyst Ratings		Analyst and Quantitative Ratings	
	(1)	(2)	(3)	(4)
Panel A: Parent				
Perceived skill	2.884	−6.991	19.602***	9.967***
	(4.900)	(6.995)	(3.035)	(3.695)
Size × η	41.466**	54.163***	33.985***	33.176***
	(19.446)	(20.238)	(7.976)	(8.788)
Fees	−73.182**	−58.507*	−81.550***	−75.872***
	(30.569)	(30.949)	(12.455)	(13.943)
N	698	650	2830	2626
Adj. R^2	0.19	0.30	0.24	0.30
Controls and FEs	No	Yes	No	Yes
Panel B: People				
Perceived skill	5.666*	3.286	29.779***	11.184***
	(3.374)	(3.980)	(3.523)	(3.243)
Size × η	58.140***	22.176**	37.454***	18.483**
	(12.459)	(10.066)	(7.493)	(7.242)
Fees	−18.350***	13.448*	−30.233***	2.734
	(6.802)	(7.433)	(8.706)	(9.244)
N	698	650	2830	2626
Adj. R^2	0.13	0.50	0.15	0.51
Controls and FEs	No	Yes	No	Yes
Panel C: Process				
Perceived skill	12.111***	5.982	53.779***	16.066***
	(3.987)	(5.486)	(4.679)	(3.478)
Size × η	38.938***	22.032*	22.307***	18.546*
	(10.908)	(12.623)	(6.817)	(9.641)
Fees	−12.487	−4.499	−35.460***	4.461
	(10.739)	(6.697)	(10.570)	(10.022)
N	698	650	2830	2626
Adj. R^2	0.08	0.46	0.21	0.53
Controls and FEs	No	Yes	No	Yes

The table shows regressions of pillar scores on fund and manager characteristics for cross-sections of funds in December 2020. Specifications (1) and (2) use U.S.-domiciled funds with an Analyst Rating. Specifications (3) and (4) use U.S.-domiciled funds with an Analyst Rating or a Quantitative Rating. Specifications (2) and (4) use the same controls and fixed effects as in specifications (2) and (4) in Table 6 in the main paper, with the exception that fund family fixed effects are not used in the regressions with parent pillar scores. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

E Regressions using panel data and fund fixed effects

E.1 Regressions of actual returns on size using fund fixed effects

Panels A and B of Table E1 show regressions of realized before-fee alphas on lagged fund size using the OLS estimator, the fund fixed effects estimator, the recursive demeaning estimator of Pástor et al. (2015) (RD1), and the recursive demeaning estimator of Zhu (2018) (RD2). All specifications using the preferred RD2 estimator show a significantly negative impact of fund size on fund returns.

When we restrict the sample to U.S.-domiciled funds and to the 1995–2014 period, the sample period used by Zhu (2018), we obtain estimates very close to hers despite calculating alphas slightly differently.[9] For example, in untabulated results of the regressions using monthly data, the coefficient estimates become −0.14 for the fund fixed effects estimator and −0.22 for the RD2 estimator, compared with −0.16 and −0.26, respectively, of Zhu (2018).

The benchmark with which to compute alphas in Panels A and B is dictated by Morningstar's methodology. Adams, Hayunga, and Mansi (2022) revisit the results of Pástor et al. (2015), albeit focusing on their industry-level as opposed to fund-level results, and find conflicting results regarding industry- and fund-level decreasing returns to scale. Adams et al. (2022) state that a major concern is the incorrect use of Morningstar's current performance benchmarks to measure historical return performance. Note that we use Morningstar's historical category assignments as opposed to the most recent ones, so our analysis should not be subject to this concern. Nonetheless, in Panels C and D of Table E1 we redo the analysis of Panels A and B using a combination of Vanguard index funds as benchmarks, as done by Berk and van Binsbergen (2015). Again and similar to the reply to the above paper by Pástor, Stambaugh, Taylor, and Zhu (2022), we find a significantly negative impact of size on returns in all specifications using the RD2 estimator.

[9] Both Zhu (2018) and Pástor et al. (2015) calculate alphas as the simple difference between the fund return and the benchmark return, without adjusting for different exposures to the benchmark.

Table E1: Decreasing returns to scale

	U.S. sample				All fund sample			
	(1) OLS	(2) FE	(3) RD1	(4) RD2	(5) OLS	(6) FE	(7) RD1	(8) RD2
Panel A: Monthly data, Morningstar benchmark								
Size (× 100)	−0.007***	−0.105***	−0.375***	−0.179***	0.002	−0.108***	−0.365**	−0.084***
	(0.002)	(0.006)	(0.131)	(0.021)	(0.002)	(0.004)	(0.175)	(0.030)
N	633540	633540	633540	633540	3020581	3020581	3020581	3020581
Panel B: Annual data, Morningstar benchmark								
Size (× 100)	−0.254***	−1.384***	−2.113**	−1.856***	−0.093***	−1.687***	1.484	−0.998***
	(0.040)	(0.097)	(1.077)	(0.278)	(0.032)	(0.066)	(1.305)	(0.349)
N	41643	41643	41643	41643	174924	174924	174924	174924
Panel C: Monthly data, Vanguard benchmark								
Size (× 100)	−0.004**	−0.088***	−0.175*	−0.133***	−0.003	−0.129***	−0.243	−0.146***
	(0.002)	(0.005)	(0.102)	(0.017)	(0.003)	(0.006)	(0.217)	(0.046)
N	633540	633540	633540	633540	3020581	3020581	3020581	3020581
Panel D: Annual data, Vanguard benchmark								
Size (× 100)	−0.218***	−1.204***	−0.455	−1.409***	0.200***	−1.905***	0.438	−2.739***
	(0.040)	(0.104)	(1.253)	(0.257)	(0.048)	(0.096)	(1.772)	(0.557)
N	41643	41643	41643	41643	174924	174924	174924	174924

The table shows coefficient estimates on lagged fund size in regressions of gross abnormal fund returns on lagged fund size in an unbalanced panel from 1979 to 2020. FE refers to the estimator that includes fund fixed effects. RD1 refers to the recursive demeaning estimator of Pástor et al. (2015), which recursively forward-demeans all variables and instruments for forward-demeaned fund size using backward-demeaned fund size while imposing a zero intercept in the first stage. RD2 refers to the recursive demeaning estimator of Zhu (2018), which instead instruments for forward-demeaned fund size using total fund size and includes an intercept in the first-stage regression. The U.S.-domiciled sample of funds includes funds from the following nine Morningstar Categories: U.S. Fund Large Growth, U.S. Fund Large Blend, U.S. Fund Large Value, U.S. Fund Small Growth, U.S. Fund Small Blend, U.S. Fund Small Value, U.S. Fund Mid-Cap Growth, U.S. Fund Mid-Cap Blend, and U.S. Fund Mid-Cap Value. Size is the logarithm of the fund's total assets under management (AUM) at the end of the previous period expressed in millions of December 2020 USD. Standard errors are shown in parentheses and clustered by Morningstar Category times year–month or Morningstar Category times year. Abnormal returns are relative to each fund's Morningstar Category benchmark or relative to a combination of Vanguard index funds. Standard errors are additionally clustered by fund in the RD specifications. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

E.2 Replication of ratings in the time series

Since October 2019, we can infer alphas from the updated Morningstar Analyst and Quantitative Rating methodology. While all Quantitative Ratings are according to the new methodology as of October 2019, analyst-rated funds have been gradually updated over the following months until we only observe funds with ratings according to the new methodology in our cross-section as of December 2020. However, for the *replication* of ratings in every month since October 2019, we need to span the entire distribution of net-of-fee alphas—including those of funds not yet rated under the new methodology—and then bin the alphas into final ratings. We have all the required inputs to compute net-of-fee alphas for funds with an Analyst Rating under the old methodology except for the new individual pillar scores ("Parent," "People," and "Process"). Under the old methodology, individual pillar scores ranged from "Negative" via "Neutral" to "Positive." We assume that these three verbal expressions correspond to pillar scores of –1, 0, and +1, respectively. Then, for each of the three pillars, we translate the scoring scale into the new scoring scale ranging from –2 to +2:

1. First, we regress the new pillar ratings on a set of characteristics for the sample of updated funds:[10]

$$\text{PillarScore}_i = \gamma_0 + \gamma' X_i + \psi_i, \tag{C9}$$

   where the vector of characteristics, X_i, includes a fund's old pillar rating, its old Morningstar Analyst Rating, and its annual fee. The adjusted R^2 values of these regressions range from 61% to 75%.

2. Then, we use the coefficients obtained from the above regressions to predict the pillar score for a not yet updated fund:

$$\text{PillarScore}_j^{predicted} = \hat{\gamma}_0 + \hat{\gamma}' X_j. \tag{C10}$$

We also perform our replication exercise for the time period before the methodology change in October 2019. That is, we test to what extent constructing ratings according to the new methodology recovers the actual ratings in the database, even though, as far as we

[10] This is similar to the process that Morningstar recommends for predicting the new ratings of not yet updated funds: "For instance, if we run a fund through the updated methodology and that fund sits in the same peer group; has similar People, Process, and Parent Pillar ratings; and sports a similar expense ratio to a fund that hasn't gone through yet, then the peer fund's Analyst Rating can offer clues into how that fund will eventually be rated under the new methodology" (Ptak, 2019).

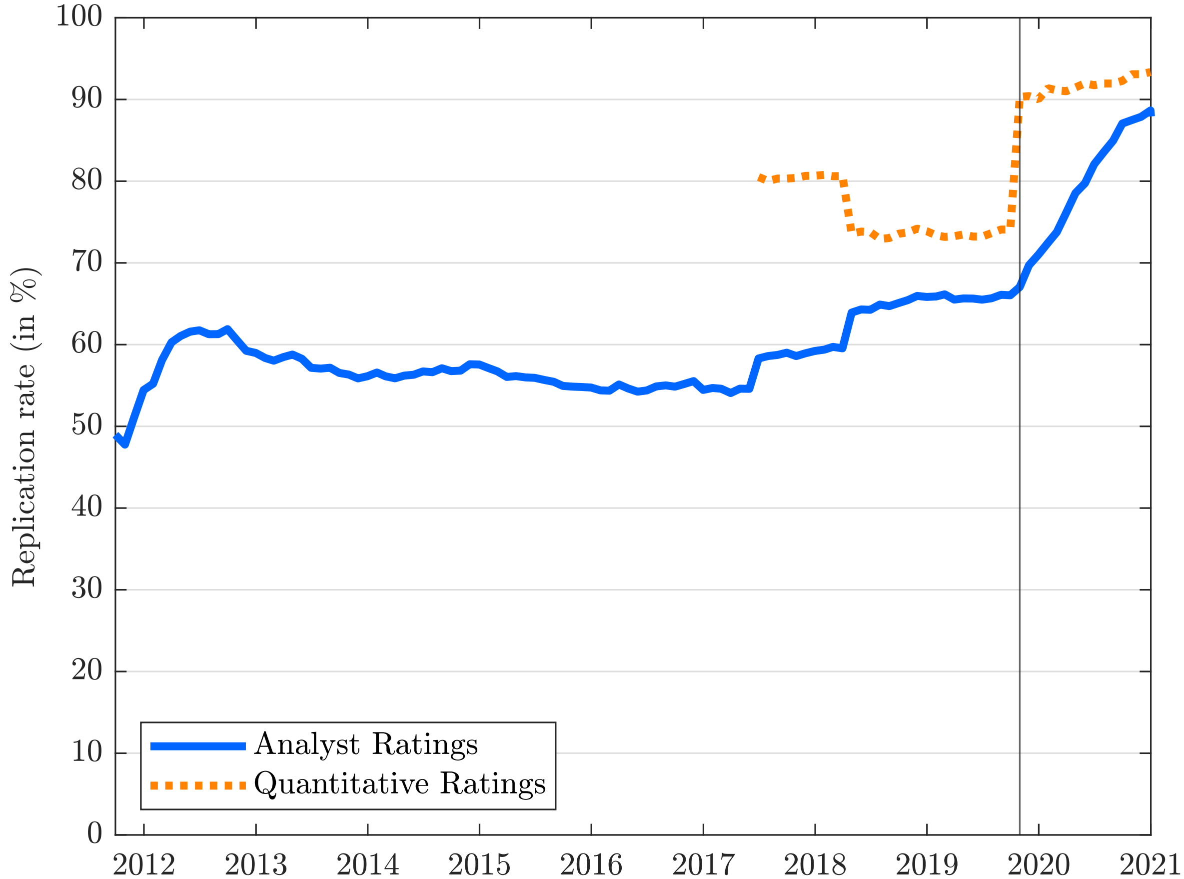
know, the ratings have not been awarded using the new methodology. For this purpose we predict pillar scores according to the new methodology using the coefficients from the above regressions for funds with an old Analyst Rating and employ a similar procedure for funds with a Quantitative Rating.

Figure E1 shows the percentage of Analyst and Quantitative Ratings that we can replicate by applying the new methodology retroactively since 2011. The replication rate for Quantitative Ratings immediately jumps to about 90% in October 2019 when the new methodology was first introduced. Instead, the replication rate for Analyst Ratings steeply but gradually increases to this level over the following months as the number of funds for which we need to predict pillar scores decreases. The prediction procedure affects the replication of ratings for funds still rated under the old methodology by introducing an additional source of estimation error in net-of-fee alphas. However, this also indirectly affects the replication of ratings for funds already rated under the new methodology because the entire distribution of all alphas serves as the basis for binning into final ratings. Therefore, the replication rate for Quantitative Ratings further increases after October 2019 even though the methodology change was fully implemented for the entire universe of Quantitative Ratings in October 2019. While the cross-section of ratings as of December 2020, which we consider for our main analyses in the paper, only contains ratings according to the new methodology, the replication rate for analyst-rated funds is still slightly lower than the replication rate for funds with a Quantitative Rating. This is because Analyst Ratings are generally updated once a year, so the Analyst Ratings that we observe as of December 2020 were partly awarded in preceding months, in which our prediction procedure still affected the distribution of net-of-fee alphas.[11] While this complicates recovering the actual rating label, importantly, it does *not* affect our ability to accurately infer alphas.

The replication of ratings from before the methodology change in October 2019 is less successful, as indicated by a substantially lower replication rate (see Figure E1). Note that we do not necessarily expect to recover any of the old ratings by using the new methodology

[11] The Morningstar Direct database contains a snapshot variable indicating the date the last Analyst Rating and accompanying analyst report were published. This allows us to check whether we can recover the Analyst Rating for the month in which the rating was actually published. While this is the more accurate comparison in terms of timing—in contrast to checking whether we can recover the rating as of December 2020, knowing that the rating was based on the distribution of net-of-fee alphas in a preceding month—it entails the risk that the imputation procedure might affect the distribution of net-of-fee alphas in that earlier month, as not necessarily all ratings are updated to the new methodology. Furthermore, since only the date on which the last Analyst Rating was published is available in the database, we cannot base the comparison of replicated and actual ratings on the month a rating was published for Analyst Ratings other than the one most recently published.

Figure E1: Replication of ratings in the time series



The figure shows the percentage of Analyst and Quantitative Ratings that we can replicate by applying the new methodology to funds predominantly rated under the old methodology from September 2011 to December 2020. The vertical line indicates October 2019, the month the new methodology was first introduced.

due to notable differences from the old methodology (e.g., Analyst Ratings were awarded on the fund level as opposed to the share class level before October 2019; see Table 1 in the main paper for further differences). Therefore, the lower replication rate before October 2019 merely serves as an indication of a significant change in the methodology.

E.3 Regressions of expectations on size using fund fixed effects

Panel A of Table E2 shows regressions of monthly analyst alphas on lagged fund size with fund fixed effects. The sample is restricted to funds with ratings according to Morningstar's new methodology since October 2019. For comparability to the main text, we use net-of-fee analyst alphas as dependent variables, but the results are similar when we use gross-of-fee analyst alphas (unreported). Of course, the results must be interpreted with caution as the time-series dimension is short. Panel B of Table E2 shows ordered logistic regressions of ordinal scale ratings (i.e., Gold, Silver, Bronze, Negative, and Neutral) on lagged fund size with fund fixed effects using the time series since 2011 when the ratings were first introduced. The ratings before October 2019 do not necessarily measure alphas, but they are still a measure of expected future performance. Since ratings are fairly persistent over time for a given fund, it is important to cluster standard errors by fund. All specifications show a significantly positive impact of fund size on expectations of fund performance.

Note that the small-sample (downward) bias of the fixed effects estimator is likely less severe with expectations as the dependent variable since there is no mechanical relationship between expectations and size. If it is severe, it will work against us and our reported coefficient estimates will be smaller than they would be without the bias.

However, as we have emphasized in the main text, regressions with fund fixed effects are less powerful in our context. Clearly, they are evidence against full-information rational expectations, but the prevailing hypothesis in the literature on mutual funds is noisy-information rational expectations. With noisy information, agents are uncertain about some of the parameters of the economy (e.g., managerial skill).

In fact, without controlling for measures and proxies of time-varying perceived skill as we do in our main analysis, we *expect* analysts' expectations to increase as a given fund grows larger: presumably both analysts and investors, who ultimately determine fund size, update their beliefs in the same direction in response to positive news (e.g., positive fund returns).

For the same reason, the mere predictability of forecast errors would be evidence against full-information rational expectations, but not necessarily against rational expectations, as predictable forecast errors may simply indicate departures from the full-information assumption (see, e.g., Coibion and Gorodnichenko, 2015). For instance, Pástor and Stambaugh (2012) describe investors who have rational expectations but are uncertain about both managerial skill and the decreasing returns to scale parameter. Investors in their model continue to expect positive returns from active management even though active management repeatedly underperforms, so forecast errors are predictable. Similarly, forecast errors in Berk and Green (2004) and in the rational expectations learning model are predictable by the difference between true skill and skill as perceived by investors.

Table E2: Panel regressions of expectations on size with fund fixed effects

	Analyst Ratings		Analyst and Quantitative Ratings	
	U.S. sample	All fund sample	U.S. sample	All fund sample
Panel A: Alphas				
Size (× 100)	0.074^{**}	0.116^{***}	0.518^{***}	0.180^{***}
	(0.034)	(0.033)	(0.088)	(0.029)
N	3915	12664	22628	197565
Panel B: Ratings				
Size	3.178^{***}	2.248^{***}	1.444^{***}	0.895^{***}
	(0.427)	(0.157)	(0.133)	(0.043)
N	40302	161502	96019	587724

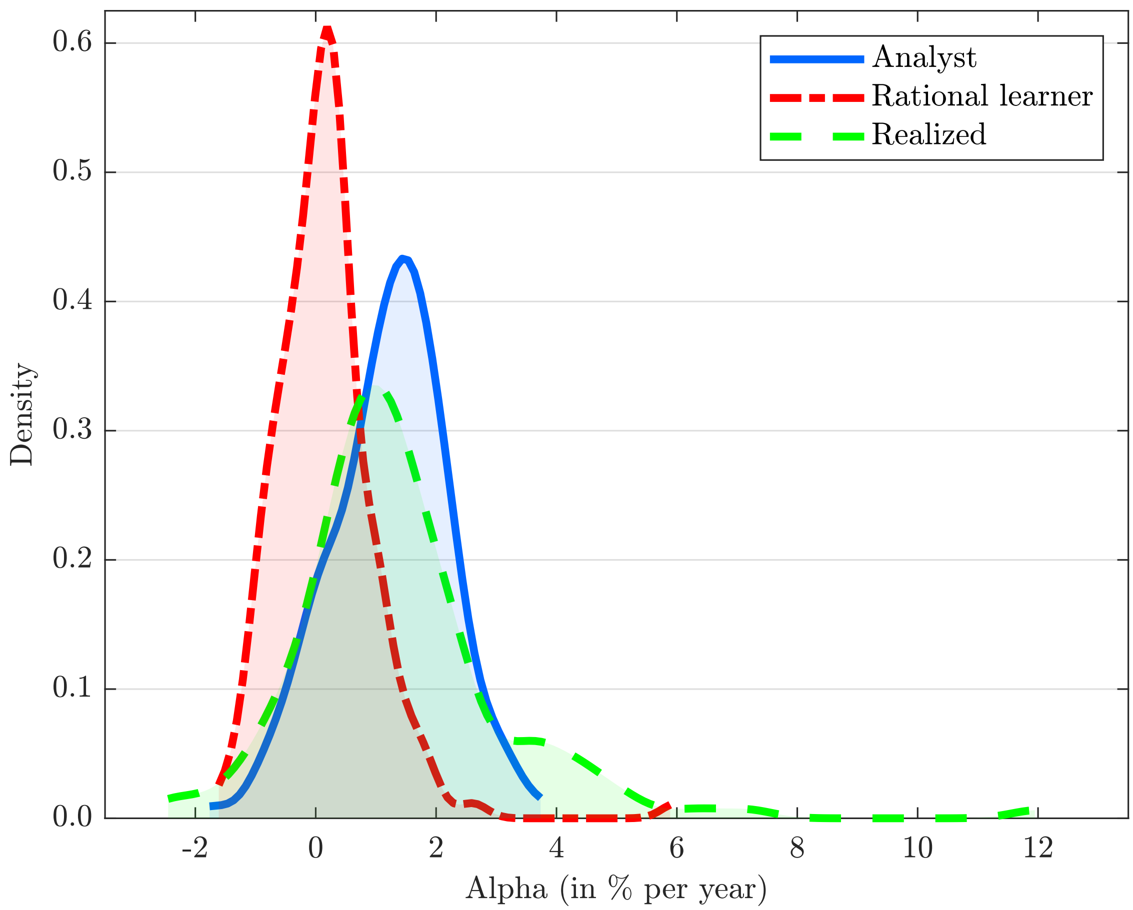
The table shows coefficient estimates on lagged fund size in regressions of expected fund performance on lagged fund size and fund fixed effects using monthly data. Specifications (1) and (2) use funds with an Analyst Rating. Specifications (3) and (4) use funds with an Analyst Rating or a Quantitative Rating. In Panel A, the dependent variables are analyst expected net-of-fee abnormal returns (alphas). The sample is restricted to funds with ratings according to Morningstar's updated methodology since October 2019. Panel B estimates ordered logistic regressions, in which the dependent variables are ratings on an ordinal scale (Gold = 5, Silver = 4, Bronze = 3, Neutral = 2, and Negative = 1). Specifications (1) and (2) use all funds with Analyst Ratings since 2011. The samples in specifications (3) and (4) start in 2017, which is when the Quantitative Ratings were first introduced. Quantitative Ratings as observed in the data are lagged by one month because Morningstar publishes each monthly batch of Quantitative Ratings near the end of the following month. The U.S.-domiciled sample of funds includes funds from the following nine Morningstar Categories: U.S. Fund Large Growth, U.S. Fund Large Blend, U.S. Fund Large Value, U.S. Fund Small Growth, U.S. Fund Small Blend, U.S. Fund Small Value, U.S. Fund Mid-Cap Growth, U.S. Fund Mid-Cap Blend, and U.S. Fund Mid-Cap Value. Size is the logarithm of the fund's total assets under management (AUM) at the end of the previous period expressed in millions of December 2020 USD. Standard errors are shown in parentheses and clustered by fund and year–month in Panel A and by fund in Panel B. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

E.4 Representativeness of cross-section in December 2020

The analysis in this Internet Appendix and the main paper uses data from our download on 9 February 2021. We also downloaded the same data again on 28 January 2022 in order to re-estimate our main tables using the cross-sections of analyst alphas in December 2020 as well as December 2021.

Figure E2 replicates Figure 1 in the paper using the cross-section of funds in December 2021. Tables E3 and E4 show panel regressions using both the cross-section in December 2020 and the one in December 2021. The results are similar to those discussed in the main text.

Figure E2: Alphas of the ten percent largest analyst-rated funds



The figure shows the cross-sectional distributions of analyst alphas (in blue) and alphas as implied by a rational expectations learning model (in red), as well as backward-looking historically realized alphas (in green), all as of December 2021. Realized alphas are computed over the lifetime of a fund. The sample is restricted to the ten percent largest funds with an Analyst Rating as of December 2021. Alphas are relative to each fund's Morningstar Category benchmark.

Table E3: Cross-sectional regressions of alphas on fund characteristics—December 2020 and December 2021

	Analyst Ratings	Analyst and Quantitative Ratings
	(1)	(2)
Perceived skill	0.384***	0.702***
	(0.069)	(0.039)
	[0.000]	[0.000]
Size × η	0.233	0.535***
	(0.142)	(0.109)
	[0.000]	[0.000]
Fees	−1.028***	−1.525***
	(0.141)	(0.055)
	[0.844]	[0.000]
Constant (× 100)	0.239	−1.377***
	(0.259)	(0.176)
	[0.358]	[0.000]
N	2893	28516
Adj. R^2	0.16	0.32

The table shows regressions of Morningstar analyst alphas on skill as perceived by a rational learner, fund size (logarithm of assets under management in millions of USD) × η, and fees for two cross-sections of funds in December 2020 and December 2021. Specification (1) uses funds with an Analyst Rating. Specification (2) uses funds with an Analyst Rating or a Quantitative Rating. Alphas are relative to each fund's Morningstar Category benchmark. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient. In brackets are p-values for the null hypothesis that the coefficients of skill, size × η, fees, and the constant equal the model-predicted parameters of +1, −1, −1, and 0, respectively.

Table E4: Cross-sectional regressions of alphas on additional fund characteristics—December 2020 and December 2021

	Analyst Ratings		Analyst and Quantitative Ratings	
	(1)	(2)	(3)	(4)
Rational learner				
Perceived skill	0.305***	0.093	0.782***	0.269***
	(0.071)	(0.058)	(0.073)	(0.043)
Size × η	0.606***	0.353**	0.463***	0.159
	(0.206)	(0.175)	(0.112)	(0.107)
Fees	−1.448***	−0.980***	−1.888***	−1.158***
	(0.164)	(0.108)	(0.164)	(0.168)
People				
Manager tenure		0.123***		0.231***
		(0.038)		(0.028)
Manager ownership		0.133***		0.185***
		(0.050)		(0.034)
Managerial multitasking		0.608***		0.303***
		(0.169)		(0.103)
Management team		0.071		0.423***
		(0.104)		(0.094)
Process				
Top 10 assets (%)		0.102		−0.025
		(0.090)		(0.067)
Tracking error		−0.015		−0.072
		(0.068)		(0.068)
Turnover ratio		−0.351***		−0.087
		(0.109)		(0.066)
Retail		−0.255***		−0.176***
		(0.076)		(0.068)
Broker-sold		−0.283***		−0.041
		(0.086)		(0.087)
N	1347	1335	5524	5460
Adj. R^2	0.24	0.66	0.28	0.66
Sustainability FE	No	Yes	No	Yes
Star FE	No	Yes	No	Yes
Morningstar Category FE	No	Yes	No	Yes
Fund Family FE	No	Yes	No	Yes

The table shows regressions of Morningstar analyst alphas on fund and manager characteristics for two cross-sections of funds in December 2020 and December 2021. Specifications (1) and (2) use U.S.-domiciled funds with an Analyst Rating. Specifications (3) and (4) use U.S.-domiciled funds with an Analyst Rating or a Quantitative Rating. Alphas are relative to each fund's Morningstar Category benchmark. Manager tenure is the maximum tenure (in months) taken over all managers, manager ownership is the average amount managers of a fund personally invest in the fund, managerial multitasking is the average number of additional funds that managers of a particular fund manage, and management team is a dummy for team-managed funds. Top 10 assets is the percentage of AUM in the ten largest positions, tracking error is the standard deviation of returns in excess of the benchmark over the life of the fund, turnover is a fund's trading activity as reported to the SEC, retail is a dummy for whether a fund is primarily held by retail investors, and broker-sold is a dummy for whether a fund is primarily sold through brokers. "People" and "Process" variables are standardized to zero mean and unit standard deviation (except for the dummy variables), and the coefficient estimates are multiplied by 100. Standard errors are clustered by fund family and shown in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively, for the null hypothesis of a zero coefficient.

F Fund flows and ratings

Table F1 shows regressions of monthly fund flows,

$$\text{Flow}(\%) = \frac{\text{AUM}_{i,t} - \text{AUM}_{i,t-1} \times (1 + R_{i,t})}{\text{AUM}_{i,t-1} \times (1 + R_{i,t})} \times 100, \tag{C11}$$

on Morningstar Analyst Ratings, Star Ratings, various control variables, as well as fund, category (also known as "style"), and year–month (time) fixed effects. Fund fixed effects are included to rule out that any unobserved fund heterogeneity that is constant over time drives the effect of ratings on flows. In specification (1), Gold-rated funds receive 0.775-percentage-point larger monthly flows than do Neutral-rated funds; equivalently, all else being equal, Gold-rated funds receive more than 9-percentage-point larger yearly flows than do Neutral-rated funds. Specification (2) shows that the effect of Analyst Ratings on flows weakens once the Star Rating is included. Nevertheless, funds recommended by analysts—Gold-, Silver-, and Bronze-rated funds—still attract significantly more flows than do funds with a Neutral Analyst Rating. Gold-rated funds receive 8-percentage-point larger flows per year than do Neutral-rated funds with the same Star Rating.

Specifications (3) and (4) repeat specifications (1) and (2) but include separate indicator variables for funds with a Quantitative Rating, which in (1) and (2) enter the "Unrated" group, since these funds do not have an Analyst Rating. Morningstar publishes each monthly batch of Quantitative Ratings near the end of the following month, so we lag the Quantitative Ratings that we observe in the data by one month to avoid look-ahead bias. The sample starts in 2017, which is when the Quantitative Ratings were first introduced. Controlling for the Star Rating in specification (4), funds with an Analyst Rating of Gold receive 12-percentage-point larger flows per year than do Neutral-rated funds. For example, if a fund with AUM of USD 4760 million (average fund size of analyst-rated funds in December 2020) is assigned a Gold as opposed to a Neutral rating, its flows increase by about USD 552 million per year. The impact of Quantitative Ratings on flows is weaker, but Gold-, Silver-, and Bronze-rated funds also attract significantly more flows than do funds with a Neutral Quantitative Rating. When a Quantitative Rating of Gold is assigned to the average-sized quantitative-rated fund, the fund receives USD 16 million ($0.335\% \times 12 \times$ USD 409 million) larger inflows per year than does a Neutral-rated fund with the same Star Rating.

Table F2 shows further robustness and reports the coefficient estimates on Analyst and Quantitative Ratings in regressions of flows on ratings without any controls or fixed effects and when explicitly controlling for perceived managerial skill from our baseline rational expectations learning model.

Table F1: Fund flows on Analyst Ratings

	Analyst Ratings 2011–2020		Analyst and Quantitative Ratings 2017–2020	
	(1)	(2)	(3)	(4)
Analyst Ratings				
Gold	0.775***	0.666***	1.091***	0.966***
	(0.116)	(0.100)	(0.224)	(0.222)
Silver	0.525***	0.438***	0.895***	0.756***
	(0.080)	(0.070)	(0.205)	(0.199)
Bronze	0.324***	0.257***	0.552***	0.448***
	(0.046)	(0.042)	(0.102)	(0.100)
Neutral				
Negative	−0.247*	−0.131	−0.254	−0.223
	(0.137)	(0.134)	(0.253)	(0.260)
Unrated	0.062	0.019	−0.163	−0.178
	(0.046)	(0.044)	(0.121)	(0.113)
Quantitative Ratings				
$Gold^Q$			0.484***	0.335***
			(0.059)	(0.060)
$Silver^Q$			0.245***	0.146***
			(0.051)	(0.055)
$Bronze^Q$			0.168***	0.102***
			(0.034)	(0.034)
$Neutral^Q$				
$Negative^Q$			−0.109**	−0.040
			(0.055)	(0.053)
$Unrated^Q$			−0.079	−0.069
			(0.068)	(0.066)
Star Ratings				
Five-star		1.703***		1.556***
		(0.045)		(0.070)
Four-star		0.592***		0.496***
		(0.021)		(0.038)
Three-star				
Two-star		−0.404***		−0.347***
		(0.030)		(0.036)
One-star		−0.760***		−0.831***
		(0.050)		(0.084)
No-star		0.262***		0.165
		(0.072)		(0.139)
N	1370489	1370489	529998	529998
Adj. R^2	0.12	0.12	0.15	0.15
Controls	Yes	Yes	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Morningstar Category FE	Yes	Yes	Yes	Yes

The table shows regressions of monthly equity mutual fund flows on Morningstar Analyst, Quantitative, and Star Ratings up to December 2020. Specifications (1) and (2) include Analyst Rating dummy variables (Gold, Silver, Bronze, Negative, and Unrated; Neutral is the omitted category), whereas (3) and (4) additionally include Quantitative Rating dummy variables (indicated by a Q superscript). The controls include the logarithm of assets under management (AUM) and fund family AUM (in millions of USD), fund age (logarithm of number of months since fund inception), fees, past 12-month fund returns, past 12-month volatility of fund returns, past 12-month average fund flows, and maximum manager tenure. Quantitative Ratings as observed in the data are lagged by one month because Morningstar publishes each monthly batch of Quantitative Ratings near the end of the following month. Standard errors are calculated using the spatial estimator of Driscoll and Kraay (1998), which allows for both cross-sectional and serial correlation up to four lags in the errors as well as for heteroskedasticity in the errors. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

Table F2: Fund flows on Analyst Ratings

	Analyst Ratings 2011–2020			Analyst and Quantitative Ratings 2017–2020		
	(1)	(2)	(3)	(4)	(6)	(6)
Perceived skill		0.505***	0.197***		0.733***	0.450***
		(0.043)	(0.037)		(0.087)	(0.091)
Analyst Ratings						
Gold	0.829***	0.680***	0.648***	0.943***	1.029***	0.961***
	(0.051)	(0.114)	(0.101)	(0.085)	(0.244)	(0.241)
Silver	0.625***	0.437***	0.399***	0.633***	0.829***	0.735***
	(0.065)	(0.079)	(0.071)	(0.057)	(0.205)	(0.199)
Bronze	0.553***	0.277***	0.240***	0.511***	0.538***	0.457***
	(0.058)	(0.048)	(0.045)	(0.050)	(0.105)	(0.103)
Neutral						
Negative	−0.410*	−0.254*	−0.183	−1.137***	−0.324	−0.304
	(0.213)	(0.135)	(0.133)	(0.247)	(0.245)	(0.262)
Unrated	0.684***	0.058	0.022	1.115***	−0.158	−0.170
	(0.045)	(0.045)	(0.044)	(0.110)	(0.128)	(0.120)
Quantitative Ratings						
Gold^Q				0.558***	0.441***	0.328***
				(0.116)	(0.059)	(0.057)
Silver^Q				0.300***	0.215***	0.141**
				(0.077)	(0.056)	(0.056)
Bronze^Q				0.226***	0.146***	0.097***
				(0.050)	(0.035)	(0.032)
Neutral^Q						
Negative^Q				−0.290***	−0.054	−0.008
				(0.044)	(0.054)	(0.051)
Unrated^Q				0.511***	−0.111*	−0.092
				(0.089)	(0.064)	(0.063)
Star Ratings						
Five-star			1.633***			1.500***
			(0.046)			(0.067)
Four-star			0.554***			0.461***
			(0.022)			(0.039)
Three-star						
Two-star			−0.346***			−0.305***
			(0.028)			(0.032)
One-star			−0.605***			−0.704***
			(0.046)			(0.076)
No-star			0.234***			0.129
			(0.071)			(0.143)
N	1324208	1324208	1324208	513124	513124	513124
Adj. R^2	0.00	0.12	0.12	0.00	0.14	0.14
Controls	No	Yes	Yes	No	Yes	Yes
Fund FE	No	Yes	Yes	No	Yes	Yes
Time FE	No	Yes	Yes	No	Yes	Yes
Morningstar Category FE	No	Yes	Yes	No	Yes	Yes

The table shows regressions of monthly equity mutual fund flows on skill as perceived by a rational learner, Morningstar Analyst, Quantitative, and Star Ratings up to December 2020. Specifications (1)–(3) include Analyst Rating dummy variables (Gold, Silver, Bronze, Negative, and Unrated; Neutral is the omitted category), whereas (4)–(6) additionally include Quantitative Rating dummy variables (indicated by a Q superscript). Perceived skill is standardized to have a mean of zero and a standard deviation of one to ease interpretation. The controls include the logarithm of assets under management (AUM) and fund family AUM (in millions of USD), fund age (logarithm of number of months since fund inception), fees, past 12-month fund returns, past 12-month volatility of fund returns, past 12-month average fund flows, and maximum manager tenure. Quantitative Ratings as observed in the data are lagged by one month because Morningstar publishes each monthly batch of Quantitative Ratings near the end of the following month. Standard errors are calculated using the spatial estimator of Driscoll and Kraay (1998), which allows for both cross-sectional and serial correlation up to four lags in the errors as well as for heteroskedasticity in the errors. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

G Hypothetical size pillar

To quantify the error analysts are making, we ask how big of a weight size "should" be assigned in a methodology that incorporates decreasing returns to scale. Our results suggest that this weight should be non-negligible and around 43%.

We estimate a possible weight on a hypothetical "size pillar" as follows. First, we construct a size pillar score similar to the other pillars with values –2, –1, 0, +1, and +2. We do not know exactly how Morningstar assigns the pillar scores, but Morningstar does give some guidance in its Quantitative Rating methodology document on how the other pillar scores are binned (Morningstar, 2021b). Specifically, they suggest a distribution of 10% (high), 22.5% (above average), 35% (average), 22.5% (below average), and 10% (low). This distribution actually also mirrors the distributions of the People, Parent, and Process pillar scores in the data.

We use these percentile thresholds and rank all funds in December 2020 based on size separately within every Morningstar category. The largest size bracket gets a score of –2, the smallest size bracket gets a score of +2.

We then compute before-fee analyst alphas similar to Equation (2) in the main paper:

$$\begin{aligned} \mathrm{E}_t^s[r_{i,t+1} + f_{i,t+1}] = \Big[& (1 - w_{size}) \, (\, 0.10 \times \mathrm{Parent}_{i,t} + 0.45 \times \mathrm{People}_{i,t} \\ & + 0.45 \times \mathrm{Process}_{i,t} \,) + w_{size} \times \mathrm{Size}_{i,t} \Big] \times \mathrm{SIQR}_{k,t}. \end{aligned} \tag{G12}$$

That is, we keep the relative weights from Morningstar's methodology on the Parent, People, and Process pillars unchanged. However, the combined weight on the three pillar scores amounts to $1 - w_{size}$ and we add the size pillar score with weight w_{size} to the equation. We compute the squared deviation between these alternative analyst alphas and the rational learner alphas and then estimate the weight on the size pillar score, w_{size}, that minimizes the mean squared error. By construction, these counterfactual ratings incorporate decreasing returns to scale, because they are closer to the rational forecast, which incorporates decreasing returns to scale. Table G1 reports the estimated weight on the size pillar and the mean squared error. The estimated weight on the size pillar is a non-negligible 43%.

Next, we compare the effect of size in the counterfactual ratings to the effect of fees. A one-percentage-point increase in fees decreases analyst alphas by one percentage point. From Equation (G12), a one-unit decrease in the size pillar score decreases analyst alphas by $w_{size} \times SIQR$.

To compare the hypothetical effect of a size pillar to fees, we standardize the effect of fees by subtracting the mean of fees in the data from 1 (percentage point) and divide by the standard deviation of fees in the data. Expressed this way, a one-standard-deviation increase in fees decreases analyst alphas by -0.39% for the set of analyst-rated funds and by -0.70% for the set of analyst- and quant-rated funds, respectively.

It turns out that a one-standard-deviation increase in size on average decreases the size pillar score by exactly one unit.[12] The effect of a one-standard-deviation increase in size is, thus, $w_{size} \times -1 \times \text{SIQR}$. As w_{size}, we choose the weight of 43.26% from above and the median SIQR in the data is 1.93%, both for the set of analyst-rated and the set of analyst- and quant-rated funds. Thus, a one-standard-deviation increase in size decreases hypothetical analyst alphas that incorporate a size pillar by -0.84%.

Overall, the effect of a hypothetical size pillar is larger than the effect of fees on hypothetical analyst alphas. A one-standard-deviation increase in fees decreases analyst alphas by -0.39% and -0.70%, respectively, whereas a one-standard-deviation increase in size decreases the counterfactual analyst alphas that incorporate DRS by -0.84%.

[12]This is based on the change in the size pillar score, within every category separately, when increasing fund size from the category median to the category median plus one-standard-deviation. The median change in the pillar score over all categories is -1.

Table G1: Estimated weight on a hypothetical size pillar

	Size pillar score percentiles					Weight on size pillar	MSE (in %)
	+2	+1	0	–1	–2		
Morningstar's methodology (no size pillar)			0–100			0.0000	0.0619
Pillar score thresholds from Quantitative Rating methodology	0–10	10–32.5	32.5–67.5	67.5–90	90–100	0.4326	0.0351

The table reports results for adding a hypothetical size pillar to the computation of analyst alphas. We compare Morningstar's actual methodology, which does not include a size pillar, with adding a size pillar using thresholds from Morningstar's Quantitative Rating methodology. We compute the difference between analyst alphas and our counterfactual rational learner alphas and minimize the mean squared error to estimate the weight on the hypothetical size pillar score.

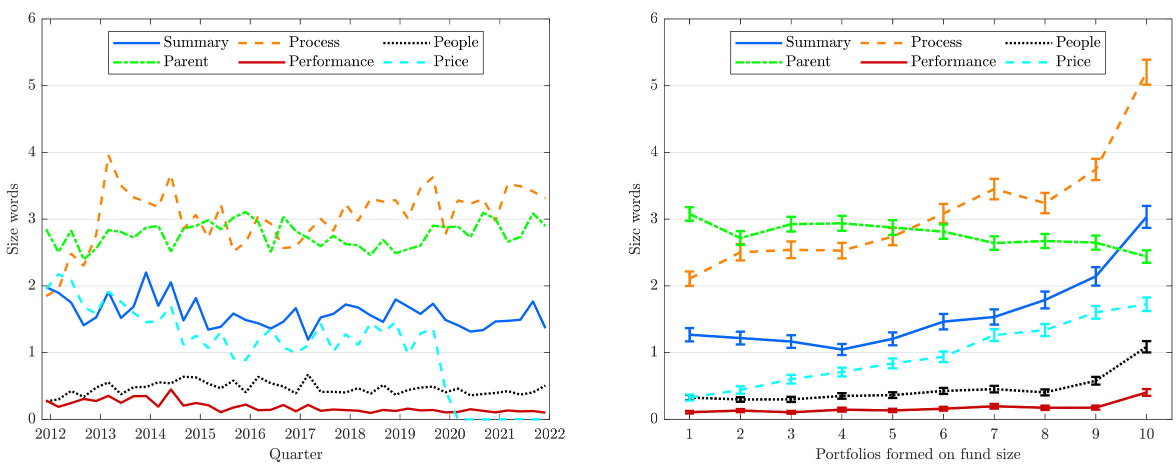
H Textual analysis

To provide evidence that analysts account for fund size in forming their expectations, even though size is not explicitly considered in the pillar ratings, we perform a textual analysis of more than 20,000 reports and notes that analysts wrote to accompany the ratings. The textual analysis follows the methodology outlined by Wilke (2025), who collects a list of size-related words in the spirit of a negative word list and other sentiment dictionaries (see, e.g., Loughran and McDonald, 2011). Candidate words are used in the context of discussing a fund's AUM. Importantly, these words are specific to this topic and rarely used otherwise, to avoid contextual misclassifications. Panel (a) of Figure H1 shows that analysts use words related to fund size in the "Process" and "Parent" pillars. Thus, even though there is no explicit pillar rating for fund size, analysts seem concerned with fund size. Panel (b) shows that analysts focus on fund size even more in the case of larger funds, corroborating the evidence of Panel (a).

Overall, while it is not possible to look inside analysts' minds, the textual analysis suggests that Morningstar's methodology does not restrict analysts from incorporating fund size when forming expectations of fund performance. Figure H1 shows that the Process pillar section contains the most size-related language of all sections in analyst reports. This is consistent with Morningstar's methodology document suggesting an evaluation "of the process with the resources backing the strategy and with the size of the asset base tied to the strategy" (Morningstar, 2021a). Moreover, fund size is not prominently discussed in the Performance pillar section. In fact, it is the section with the fewest size-related words in the entire report.

Figure H1: Size-related words in analyst reports

(a) Size-related word count **(b)** Size-related words against fund size



Panel (a) shows the number of size-related words mentioned in each part of the analyst report, averaged over all reports published per quarter from Q4 2011 to Q4 2021. Panel (b) shows the number of size-related words against fund size (AUM). Note that "Performance" and "Price" pillar commentary is still part of written analyst reports, even though "Performance" and "Price" pillar ratings ceased to exist under the new methodology in 2019. Only the remaining three pillar ratings (i.e., "People," "Parent," and "Process") feed into the calculation of the final Analyst Rating. The size-related words are from Wilke (2025) and are ASSET, ASSETS, AUD, AUM, BALLOON, BALLOONED, BALLOONING, BASE, BASES, BILLION, BILLIONS, BLOAT, BLOATED, CAD, CAPACITY, CHF, CLOSED, CLOSES, CLOSING, CLOSURE, CORPUS, EUR, FUM, GBP, GIRTH, INFLOW, INFLOWS, INR, JPY, MILLION, MILLIONS, NIMBLE, NIMBLENESS, NIMBLER, NOK, NZD, OUTFLOW, OUTFLOWS, RECLOSE, RECLOSED, REOPEN, REOPENED, REOPENING, SCALE, SGD, SIZE, SIZES, SURGING, SWELL, SWELLED, SWELLING, TRILLION and USD. The bars indicate 90% confidence bands.

References

Adams, John, Darren Hayunga, and Sattar Mansi, 2022, Scale and Performance in Active Management are Not Negatively Related, *Critical Finance Review* 11, 541–592.

Amihud, Yakov, and Ruslan Goyenko, 2013, Mutual Fund's R2 As Predictor of Performance, *Review of Financial Studies* 26, 667–694.

Barras, Laurent, Patrick Gagliardini, and Olivier Scaillet, 2022, Skill, Scale, and Value Creation in the Mutual Fund Industry, *Journal of Finance* 77, 601–638.

Berk, Jonathan B., and Richard C. Green, 2004, Mutual Fund Flows and Performance in Rational Markets, *Journal of Political Economy* 112, 1269–1295.

Berk, Jonathan B., and Jules H. van Binsbergen, 2015, Measuring Skill in the Mutual Fund Industry, *Journal of Financial Economics* 118, 1–20.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, 2019, Diagnostic Expectations and Stock Returns, *Journal of Finance* 74, 2839–2874.

Coibion, Olivier, and Yuriy Gorodnichenko, 2015, Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts, *American Economic Review* 105, 2644–2678.

Cooper, Michael, Michael Halling, and Wenhao Yang, 2020, The Persistence of Fee Dispersion among Mutual Funds, *Review of Finance* 25, 365–402.

Dempster, A. P., N. M. Laird, and D. B. Rubin, 1977, Maximum Likelihood from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society. Series B (Methodological)* 39, 1–38.

Driscoll, John, and Aart Kraay, 1998, Consistent Covariance Matrix Estimation With Spatially Dependent Panel Data, *Review of Economics and Statistics* 80, 549–560.

Fama, Eugene F., and Kenneth R. French, 2010, Luck versus Skill in the Cross-Section of Mutual Fund Returns, *Journal of Finance* 65, 1915–1947.

Gerakos, Joseph J., Juhani T. Linnainmaa, and Adair Morse, 2021, Asset Managers: Institutional Performance and Factor Exposures, *Journal of Finance* 76, 2035–2075.

Ibert, Markus, Ron Kaniel, Stijn Van Nieuwerburgh, and Roine Vestman, 2018, Are Mutual Fund Managers Paid for Investment Skill?, *Review of Financial Studies* 31, 715–772.

Loughran, Tim, and Bill McDonald, 2011, When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks, *Journal of Finance* 66, 35–65.

Morningstar, 2021a, Morningstar Analyst Rating for Funds: Methodology.

Morningstar, 2021b, Morningstar Quantitative Rating: Pillar Methodology.

Pástor, Ľuboš, and Robert F. Stambaugh, 2012, On the Size of the Active Management Industry, *Journal of Political Economy* 120, 740–781.

Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor, 2015, Scale and Skill in Active Management, *Journal of Financial Economics* 116, 23–45.

Pástor, Ľuboš, Robert F. Stambaugh, Lucian A. Taylor, and Min Zhu, 2022, Diseconomies of Scale in Active Management: Robust Evidence, *Critical Finance Review* 11, 593–611.

Ptak, Jeffrey, 2019, Introducing the Enhanced Morningstar Analyst Rating for Funds.

Roussanov, Nikolai, Hongxun Ruan, and Yanhao Wei, 2020, Mutual Fund Flows and Performance in (Imperfectly) Rational Markets?, *Working Paper*.

Roussanov, Nikolai L., Hongxun Ruan, and Yanhao Wei, 2021, Marketing Mutual Funds, *Review of Financial Studies* 34, 3045–3094.

Shumway, Robert H., and David S. Stoffer, 1982, An Approach to Time Series Smoothing and Forecasting Using the EM Algorithm, *Journal of Time Series Analysis* 3, 253–264.

Watson, Mark W., and Robert F. Engle, 1983, Alternative Algorithms for the Estimation of Dynamic Factor, Mimic and Varying Coefficient Regression Models, *Journal of Econometrics* 23, 385–400.

Wilke, Felix, 2025, Learning about Managerial Skill and Fund Scale from Mutual Fund Analysts, *Working Paper*.

Zhu, Min, 2018, Informative Fund Size, Managerial Skill, and Investor Rationality, *Journal of Financial Economics* 130, 114–134.