Are Subjective Expectations Formed as in Rational Expectations Models of Active Management?

Magnus Dahlquist, Markus Ibert, and Felix Wilke*

January 2024

Abstract

We recover forward-looking expected net-of-fee abnormal returns (alphas) for active equity mutual funds from analyst ratings. In contrast to the typical equilibrium implication of zero alphas, analyst alphas are negative for most funds, but positive for the largest funds. We compare analysts' subjective expectations with expectations from a rational expectations learning model. The model's rational learner believes that an increase in fund size leads to a decrease in returns, but we find no evidence that analysts believe so. Overall, analysts' expectations and the capital that follows analysts' recommendations are difficult to reconcile with existing rational expectations models of active management.

JEL: G11, G12, G14, G23.

Keywords: Alpha, expectations formation, mutual funds.

*Dahlquist: Stockholm School of Economics and CEPR; e-mail: magnus.dahlquist@hhs.se. Ibert: Copenhagen Business School and Danish Finance Institute; e-mail: mi.fi@cbs.dk. Wilke: Nova School of Business and Economics; e-mail: felix.wilke@novasbe.pt. We thank Laurent Barras, Sylvain Benoit (discussant), Richard B. Evans, Katrin Gödker (discussant), Campbell Harvey, Jose Vicente Martinez, Lasse H. Pedersen, Alberto G. Rossi, Andrea Rossi (discussant), Ran Xing (discussant), and seminar and conference participants at Copenhagen Business School, ESSEC Business School, the Stockholm School of Economics, the Federal Reserve Board of Governors, the Triangle Macro-Finance workshop, the Australasian Finance & Banking Conference 2020, the Annual Hedge Fund Research Conference 2023, the HEC-McGill Winter Finance Workshop 2023, and the American Finance Association Meeting 2024 for their comments and suggestions. Support from the Danish Finance Institute (DFI) is gratefully acknowledged. The paper was previously circulated under the title "Expectations of Active Mutual Fund Performance."

1 Introduction

Recent years have seen a surge of research on subjective expectations of investors and professional analysts to examine the predictions of rational expectations models in all areas of finance and economics (see, e.g., Coibion and Gorodnichenko, 2012, 2015; Greenwood and Shleifer, 2014; Bordalo, Gennaioli, La Porta, and Shleifer, 2019; Bordalo, Gennaioli, Ma, and Shleifer, 2020). We take this idea to the literature on actively managed mutual funds.

Rational expectations models in this literature make precise predictions about how return expectations are formed (see, e.g., Berk and Green, 2004) and existing research has reached opposing conclusions regarding the predictions of these models based on the revealed preferences of investors. Using data on fund flows, some researchers conclude that mutual fund investors are sophisticated Bayesian learners (see, e.g., Berk and van Binsbergen, 2015; Franzoni and Schmalz, 2017; van Binsbergen, Kim, and Kim, 2021; Barras, Gagliardini, and Scaillet, 2022; Kim, 2022), whereas others conclude that investors have limited financial sophistication (see, e.g., Song, 2020; Ben-David, Li, Rossi, and Song, 2022).

We recover forward-looking expected net-of-fee abnormal returns (henceforth, “alphas”) as perceived by mutual fund analysts for virtually all active equity mutual funds worldwide from analyst ratings provided by Morningstar. We compare analysts’ expectations with expectations implied by rational expectations models of active management. They differ. First, in contrast to the typical equilibrium implication of zero alphas, not all analyst alphas are zero. Second, we do not find any evidence that analysts’ expectations decrease as a fund’s size increases. If anything, analysts’ expectations *increase* as a fund’s size increases. In contrast and consistent with a large literature on decreasing returns to scale, we do find that *realized* fund returns decrease as a fund’s size increases (see, e.g., Chen, Hong, Huang, and Kubik, 2004; Pástor, Stambaugh, and Taylor, 2015; Zhu, 2018; Roussanov, Ruan, and Wei, 2021). We conclude that analysts seem to misjudge returns to scale in active management.

Analysts' expectations versus investors' expectations

Given the apparent ambiguity of investor fund flows to examine the predictions of rational expectations models in this literature, we believe that our novel focus on subjective expectations is valuable. However, it does not come without a cost. While professional analysts could be akin to sophisticated investors, there is little reason to believe that analysts' expectations are those of the marginal investor. That said, such a concern applies to all professional analysts—not just mutual fund analysts. In fact, it also applies to all surveys of actual investors, as researchers can never be sure to have surveyed the marginal investor (see, e.g., Choi and Robertson, 2020). One advantage of the mutual fund setting relative to professional forecasts in other settings is that we can actually test whether analysts' recommendations matter to some investors. They do: the effect of analysts' ratings on fund flows is up to 82% of the effect of the popular Star Ratings. Conceptually, our results are similar to those of Greenwood and Shleifer (2014), who show that subjective stock market return expectations are diametrically opposed to expectations implied by rational expectations asset pricing models. Similar to their survey evidence, mutual fund analysts' expectations can hardly be the expectations of a representative investor. This is because expectations that increase with size imply that unlimited amounts of capital should flow into all funds. Ultimately, misunderstandings of returns to scale in active management could help explain why many funds grow too large and underperform.

Predictions of rational expectations models

In the typical rational expectations model of active management, investors are uncertain about some parameters of the economy (e.g., managerial skill) and update their expectations from observed fund returns, which decrease as a fund's size increases. This latter concept of decreasing returns to scale is central to understanding the typical model. The decreasing returns to scale in realized returns that we and the literature document do not necessarily imply that a large fund will likely perform worse than a smaller fund. Decreasing returns to scale imply that—all else being equal—an increase in size leads to a decrease in returns relative to the passive benchmark, summarizing the notion that good investment ideas are not arbitrarily scalable. Finally, in equilibrium investors allocate capital to funds competitively such that alphas are zero (see, e.g., Berk and Green, 2004) or close to zero (see, e.g., Pástor and Stambaugh, 2012).

Data on subjective expectations

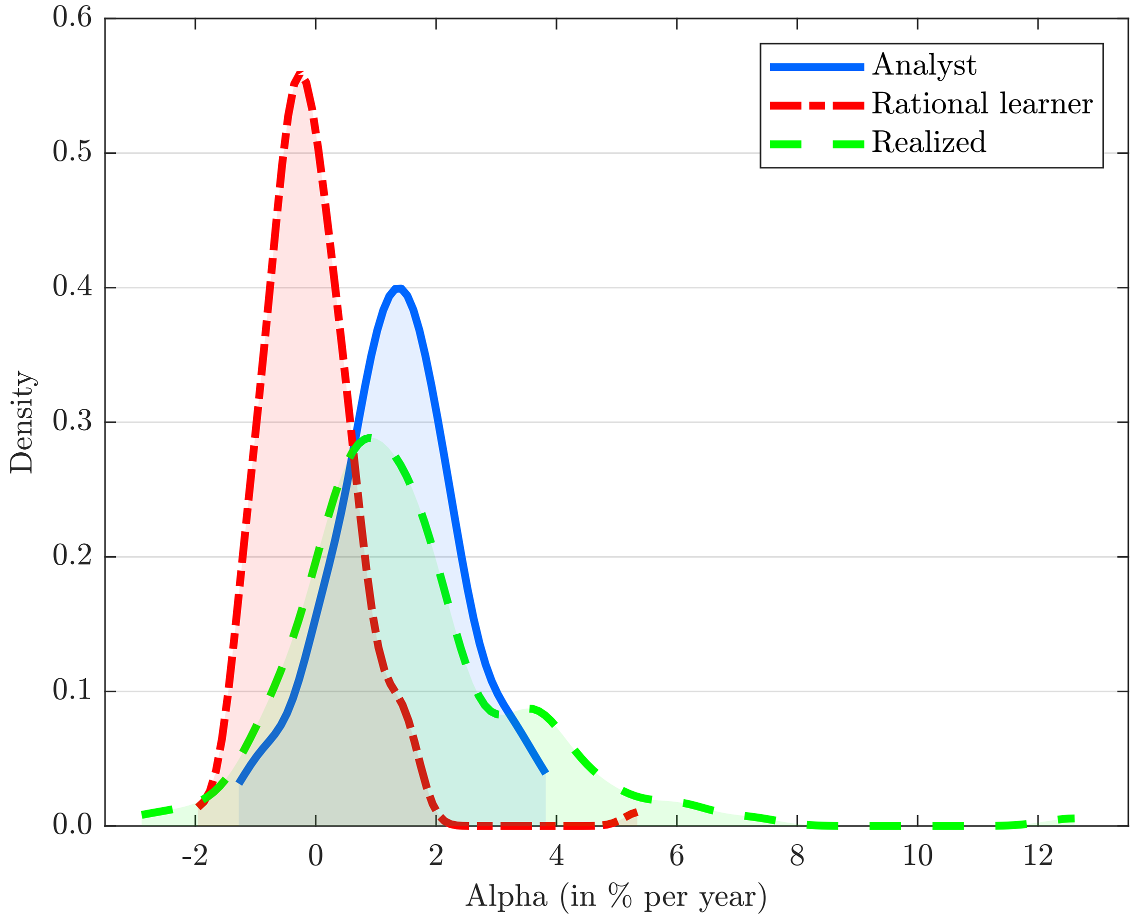
We recover expectations from analyst ratings provided by Morningstar, a leading financial services firm in the USD 13 trillion active equity mutual fund industry. As Morningstar overhauled the methodology for its forward-looking ratings in October 2019 and then provided a detailed description of how the overhauled ratings are constructed, we can recover detailed measures of expectations since then. Analysts assign the ratings according to a five-tier scale with three positive ratings of Gold, Silver, and Bronze, as well as a Neutral rating and a Negative rating. Under the new methodology, Morningstar constructs a distribution of alphas and then groups the alphas (which are not reported in the database) to arrive at the final Morningstar Analyst Ratings (which are reported in the database). We replicate Morningstar's new methodology to recover the alphas that the analysts use. When we translate our alphas into ratings, we can replicate 93% of the ratings.

Analysts' expectations versus model-implied expectations

Figure 1 illustrates our main results. The figure shows analyst alphas (in blue), alphas implied by a rational expectations learning model that we estimate and introduce below (in red), and backward-looking historically realized alphas (in green), all for the cross-section of the largest ten percent of analyst-rated funds in December 2020.

First, the Berk and Green (2004) equilibrium implication of a zero alpha for each and every fund is trivially counterfactual when compared with analyst alphas: not all analyst

Figure 1: 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 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.

alphas are zero. In fact, analyst alphas are only positive for the largest funds (shown in Figure 1), but negative for most other funds (not shown in Figure 1). That said, it is well known that not all realized alphas are zero (see, e.g., Kosowski, Timmermann, Wermers, and

White, 2006; Fama and French, 2010; Harvey and Liu, 2022) and previous research indeed interprets rational expectations models of active management in an approximate sense (see, e.g., Berk and Tonks, 2007). Thus, in what follows we relax the equilibrium implication of zero alphas.

Second, once the zero-alpha equilibrium implication is relaxed, the key prediction of rational expectations models of active management concerns decreasing returns to scale.[1] Consider now the distributions of analyst alphas and historically realized alphas. Figure 1 restricts the sample to the largest funds as of December 2020 because, if anywhere, the effect of decreasing returns to scale should be visible for these funds. On average, they have grown their assets under management (AUM) from USD 1 billion to USD 30 billion over the last 30 years. These increases in AUM are among the greatest in both absolute and relative terms. However, despite the growth in AUM, the figure shows that analysts extrapolate from past returns: they expect these funds to at least sustain the returns that they have earned in the past (in blue and green).[2] Such extrapolation for the funds that have seen the greatest increases in size is inherently difficult to reconcile with a belief in decreasing returns to scale.

Third, what do typical rational expectations models imply about expected returns going forward for the funds that have grown to be the largest? The distribution shown in red in Figure 1 is implied by a Berk and Green (2004)-type model once the equilibrium implication of zero alphas is relaxed. Without the equilibrium implication, their model is a filtering problem: a rational learner who is uncertain about managerial skill updates beliefs from past fund returns to form expectations of future returns. Any such Bayesian learning model

[1] Fama and French (2010) write: “For many readers, the important insight of Berk and Green (2004) is their assumption that there are diseconomies of scale in active management, not their detailed predictions about net fund returns (which are rejected in our tests).”

[2] In fact, for around 50% of the funds in Figure 1, analysts predict *larger* alphas going forward than these funds’ historically realized alphas, despite that these funds operate at record-high sizes. Similar to Linnainmaa (2013), in a simple learning model you would expect a fund’s alpha going forward to be bounded by a reasonable prior, say zero, and the historically realized alpha—unless you believe increases in size actually increase future returns.

can in principle "rationalize" analysts' expectations by imposing arbitrary priors (e.g., a prior belief in no decreasing returns to scale together with a high certainty around that prior). An important point of our paper is to take Bayesian models to the data rigorously and so we estimate the Berk and Green (2004) model using an empirical Bayes method as in Roussanov et al. (2021).[3] As in Roussanov et al. (2021), the estimation uncovers decreasing returns to scale in realized fund returns and so the distribution of alphas perceived by the rational learner shown in Figure 1 is notably shifted to the left for the funds that have grown to be the largest.

By imposing structure via the rational expectations learning model, we can also extend the results in Figure 1 to all funds (not just the largest ones). One advantage of the estimated rational expectations learning model is that its predictions can be tested using a simple cross-sectional regression of analyst alphas on the fund characteristics in the model: perceived managerial skill, size, and fees. Consistent with the rational expectations learning model, analyst alphas decrease with fees and increase with perceived skill. Inconsistent with the model, a 100% increase in AUM *increases* analysts' expectations by 9 basis points. Mirroring the results in the literature on decreasing returns to scale, the rational learner instead believes that a doubling of AUM leads to a 17-basis-points-decrease in alpha.

Potential concerns

You may be concerned that our results rest on the particular rational expectations learning model that we benchmark analyst alphas against, but they hardly do. The distributions of analyst alphas and historically realized alphas shown in Figure 1 do not rely on any particular rational expectations model. As long as there are decreasing returns to scale in realized fund

[3] In fact, estimating the model corresponds to the definition of "rational expectations" in this literature (see, e.g., p. 1274 in Berk and Green, 2004). The rational expectations paradigm has strict implications for the distribution of priors and other parameters in a Bayesian model: they cannot be arbitrary, but need to conform with the distribution of true parameters, which for any given model can be estimated from the data.

returns, you would expect a forecast incorporating decreasing returns to scale to be shifted to the left for this set of funds. This is in particular so given 30 years of potential learning to resolve any parameter uncertainties.

Apart from that, our results are robust to various extensions of the rational expectations learning model (including features from, among others, Pástor and Stambaugh, 2012; Pástor et al., 2015; Berk and van Binsbergen, 2015; Barras et al., 2022), they are robust when we control for additional manager and fund characteristics in reduced form, they are not confined to a particular cross-section of funds, and we also find decreasing returns to scale in realized fund returns using the estimator in Zhu (2018). Among the additional characteristics that matter for analysts' expectations are manager ownership (Khorana, Servaes, and Wedge, 2007; Evans, 2008; Ibert, 2023), manager tenure (Greenwood and Nagel, 2009), and fund family fixed effects.

Ultimately, attempts to reconcile analysts' expectations with rational expectations models of active management would need to generate measures of perceived managerial skill that, once controlled for, could flip the estimates on size from positive to negative in our regressions. With R^2 values above 60%, our specifications that control for additional manager and fund characteristics in reduced form leave little room for that. If such measures existed, they would be crucial for future model development and in turn highlight the importance of our key contribution: contrasting the rational expectations paradigm for mutual funds with subjective expectations.

Another potential concern is that analysts' forecasts may not represent their best attempts. Instead, their forecasts could also reflect incentive structures or career concerns. Morningstar is an independent research firm, has a substantial business reputation at stake, and previous research has used the Morningstar Analyst Rating as a *benchmark* of independent analysis (see, e.g, Cookson, Jenkinson, Jones, and Martinez, 2021). Moreover, our textual analysis of more than 20,000 reports and notes that accompany the Morningstar

Analyst Ratings suggests that analysts do discuss fund size, and even more so in the case of larger funds. We conclude that analysts' forecasts are their best attempts to forecast future returns, but that analysts seem to misjudge returns to scale in active management.

Related literature

Our paper relates to several strands of literature. First, a large literature examines the predictions of the Berk and Green (2004) model, including the key prediction whether an increase in size leads to a decrease in realized fund returns.[4] Contrasting the rational expectations paradigm for mutual funds with analysts' subjective expectations is novel. In their study of how leading financial theories describe individual investor behavior, Choi and Robertson (2020) also include a statement about decreasing returns to scale and report that only 18% of respondents believe in decreasing returns to scale (see also Bender, Choi, Dyson, and Robertson, 2022). Apart from their different focus, the usual caveats regarding survey data apply. It is unclear whether the surveyed investors are representative and whether they act on their expectations. Framing also matters: their statement does not allow for expectations to increase with size—something our results suggest. Overall, we view analysts' expectations as an important addition to survey-based expectations.

Second, our paper relates to a literature that examines expectations regarding fund performance (see, e.g., Jenkinson, Jones, and Martinez, 2016; Jones and Martinez, 2017; Armstrong, Genc, and Verbeek, 2019; Cookson et al., 2021). The analyst alphas we recover are an important improvement over previous work, as they can be confronted with model-implied alphas and, ultimately, can be used to compute forecast errors for virtually every fund in the universe of active equity mutual funds.[5]

[4]See Berk and van Binsbergen (2017) for a review of this literature. For recent studies that examine decreasing returns to scale in realized returns, in addition to the papers already cited, see, e.g., McLemore (2019), Pástor, Stambaugh, and Taylor (2020), Roussanov, Ruan, and Wei (2020), Dyakov, Jiang, and Verbeek (2020), Busse, Chordia, Jiang, and Tang (2021), Reuter and Zitzewitz (2021), Harvey, Liu, Tan, and Zhu (2021), and Pástor, Stambaugh, Taylor, and Zhu (2022).

[5]Armstrong et al. (2019) examine the ability of Analyst Ratings to predict fund performance from 2011

Third, our paper relates to a literature on models of active management. The rational expectations model and its perturbations, as presented here, are most closely related to the model of Berk and Green (2004). Compared to their model, the models of Dangl, Wu, and Zechner (2008), Glode and Green (2011), and Pástor and Stambaugh (2012) share similar features, that is, learning about some parameters, returns that decrease with size, and the competitive provision of capital. For a different modeling approach see, e.g., Gârleanu and Pedersen (2018). The models of Gennaioli, Shleifer, and Vishny (2015) and Spiegler (2020) allow for deviations from rational expectations.

2 Forward-looking Morningstar Ratings

2.1 Old and new ratings

Morningstar has provided monthly Analyst Ratings for a selected number of funds since 2011. Unlike the backward-looking Morningstar Rating (often referred to as the "Star Rating"), the Analyst Rating is the summary expression of Morningstar's forward-looking long-term analysis of a fund. Morningstar analysts assign the Analyst Ratings on a five-tier scale with three positive ratings of Gold, Silver, and Bronze, as well as a Neutral rating and a Negative rating. The Internet Appendix presents an example of how the Analyst Rating is displayed on Morningstar's website.

Up to October 2019, an Analyst Rating was based on an analyst's conviction of a fund's ability to outperform its peer group *and/or* relevant benchmark on a risk-adjusted basis over the long term. In October 2019, Morningstar overhauled its Analyst Rating system. The most important changes were a greater emphasis on fees and a share-class-specific rating in

to 2015 and find some evidence for it. It is impossible to recover analyst alphas before October 2019. That professional analysts' recommendations have some predictive power is not inconsistent with our results, but expected: Carhart (1997) shows that even a simple measure such as past returns has some predictive power for future returns, at least for the worst-performing funds.

contrast to a fund-level rating. Different share classes of the same fund generally earn the same return before fees, but fees differ across share classes. Under the new rating system, a fund is expected to beat both its peer group *and* a relevant benchmark on a risk-adjusted basis to earn a medalist rating (i.e., a Bronze, Silver, or Gold rating). The new rating system is therefore informative about alpha, as alpha measures the performance relative to a passive benchmark. In contrast, the old rating system is not necessarily informative about alpha, as a fund may have received a medalist rating if it was expected to outperform its peers, but not a passive benchmark.

In addition, in an effort to increase transparency, Morningstar for the first time also published a document detailing how the Analyst Ratings are constructed under the new methodology. Under the new methodology, Morningstar constructs alphas by combining a strategy's overall potential with pillar ratings for a fund's "Parent," "People," and "Process." Morningstar then groups the resulting alphas (which are not published in their database) into the aforementioned ratings (which are published in their database).

The number of funds that receive an Analyst Rating is limited by the size of the Morningstar analyst team. There are currently 72 unique analysts. To expand the number of funds covered, since 2017 Morningstar has also provided forward-looking Quantitative Ratings. These are similar to Analyst Ratings, but are based on a machine-learning algorithm that attempts to mimic a human analyst's decision-making process. Morningstar assigns Quantitative Ratings to funds not covered by human analysts. Each fund can receive either an Analyst Rating or a Quantitative Rating, but in general not both. We also include funds with a Quantitative Rating in most of our analyses. Table 1 provides a summary of the different Morningstar ratings.

2.2 Analyst and Quantitative Ratings methodology

This section details how Morningstar constructs its ratings under the new methodology and how we recover analyst alphas. The Internet Appendix contains additional details about our replication and the data.

Under the new rating system, Morningstar's exact methodology for constructing the ratings follows a three-step process. First, for each fund, Morningstar estimates performance-evaluation regressions on a rolling window starting in January 2000:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i(R_{b,i,t} - R_{f,t}) + \zeta_{i,t}, \tag{1}$$

where t runs over a rolling 36-month window, $R_{i,t}$ is the gross (i.e., before-fee) return of fund i, $R_{f,t}$ is a risk-free rate proxy, and $R_{b,i,t}$ is a fund-specific benchmark return. The performance-evaluation regressions are estimated on the fund level, not the share-class level. The estimated intercepts are grouped by fund strategy (e.g., U.S. equity large-cap blend) to form a distribution of realized alphas. Morningstar then calculates the semi-interquartile range (SIQR) of the distribution (i.e., the 75th percentile minus the 25th percentile divided by 2). The SIQR measures the historically realized alpha dispersion and summarizes Morningstar's assessment of the potential of a given strategy.

Second, Morningstar analysts score a fund based on the three individual pillars "People," "Parent," and "Process." Under the new methodology, the scores range from –2 to +2. The labels of the scores –2, –1, 0, +1, and +2 are "Low," "Below Average," "Average," "Above Average," and "High," respectively, and written as such in Morningstar products. The Analyst Rating pillar scores are assigned based on an in-depth analysis, must be approved by a ratings committee, and are explained in detail in a written report for each rated fund. The Internet Appendix includes an anonymized example of such a report. The Quantitative Rating pillar scores are assigned using the aforementioned machine-learning algorithm. The

SIQR and the pillar scores are then combined to give an estimate of the expected gross abnormal return of a fund:

$$\mathrm{E}_t^s[r_{i,t+1} + f_{i,t+1}] = \mathrm{SIQR}_{k,t} \times \left(0.10 \times \mathrm{Parent}_{i,t} + 0.45 \times \mathrm{People}_{i,t} + 0.45 \times \mathrm{Process}_{i,t}\right), \quad (2)$$

where E_t^s is the analyst's subjective expectation and $r_{i,t+1} + f_{i,t+1}$ is the fund's gross-of-fee abnormal return. The SIQR depends on the type of strategy, k, and acts as a scaling factor. The pillar ratings determine whether a fund receives a positive or negative gross analyst alpha.

Third, Morningstar subtracts the share-class-specific fee to arrive at a net-of-fee alpha for each share class, j, of fund i, that is, $\mathrm{E}_t^s[r_{i,j,t+1}]$. Conditional on a positive net alpha within a particular Morningstar Category, the top 15% of share classes receive a Gold rating, the next 35% receive a Silver rating, and the bottom 50% receive a Bronze rating. Conditional on a negative or zero net alpha within a particular category, the top 70% of share classes receive a Neutral rating and the bottom 30% receive a Negative rating.

The SIQR is not reported in the Morningstar database, so we need to recover Morningstar's SIQR estimate. Morningstar groups funds from around the world in closely related Morningstar Categories to estimate the SIQR, but is not explicit about the grouping. We group funds according to their Global Category (a Morningstar variable that groups closely related Morningstar Categories from different fund domiciles), use a fund's Morningstar Category Index as the benchmark, and use the three-month Treasury bill rate as the risk-free rate. In contrast to the SIQR, the pillar scores and fees are reported in the database, so we have all the inputs needed in order to recover the alphas before they are binned into the final ratings.

2.3 Replication

The predictions of the rational expectations model introduced below can be tested using a simple cross-sectional regression. We can recover analyst alphas since October 2019, but use the cross-section of analyst alphas in December 2020 for our main analysis. Funds with an Analyst Rating have been gradually updated since October 2019 using the new methodology, and this process was completed by December 2020. All funds with a Quantitative Rating are rated under the new methodology as of October 2019. We discuss the use of panel data in the robustness section and the Internet Appendix.

Table 2 shows that we can replicate the vast majority of Morningstar's Analyst and Quantitative Ratings, suggesting that we indeed recovered the alphas that Morningstar uses to construct the ratings. Panel A shows that for the 8697 share classes with an Analyst Rating under the new methodology, Morningstar assigns a Neutral rating to 3218 share classes. In this case, we assign a Neutral rating in 3035 cases, yielding a replication rate of 94%. Our overall replication rate for the Analyst Ratings is 89%. Panel B shows our replication of the Morningstar Quantitative Ratings. Our overall replication rate for Quantitative Ratings is 93%. In total, we can replicate 92.7% of all ratings (the average of 89% and 93% weighted by the number of share classes that have an Analyst or Quantitative Rating, respectively).

While we believe that we can replicate Morningstar's methodology reasonably well to recover analyst alphas, there is measurement error in the dependent variable. Under standard assumptions, measurement error in the dependent variable does not bias coefficient estimates, but inflates standard errors. This works against finding significant results, as our standard errors are larger than they would be without measurement error.

3 Data

We obtain gross returns, AUM, ratings, and fees for active open-end equity mutual funds from Morningstar Direct. We include all funds in the database to correctly replicate Morningstar's methodology. The sample contains both U.S.-domiciled and non-U.S.-domiciled funds. Morningstar only uses data as of January 2000 to construct the Analyst Ratings, so we use the same data in our replication of the ratings. In addition, we use the full time series available in Morningstar to estimate the rational expectations model of fund performance. The monthly sample for the estimation starts in January 1979, the first month for which Morningstar provides benchmark returns, and ends in December 2020. We convert all returns and assets to USD. As is common in the literature, we aggregate share-class-level variables (e.g., fees, returns, and analyst alphas) to the fund level by taking an AUM-weighted average.

Figure 2 plots the AUM of funds with an Analyst Rating, a Quantitative Rating, or no rating over time. As is evident from the figure, Morningstar assigns ratings to the vast majority of funds in the 13 USD trillion active equity fund industry. Table 3 presents summary statistics for the cross-section of funds in December 2020. The number of funds with a Quantitative Rating is large but the assets of these funds are much smaller on average. Moreover, the table shows that funds with Analyst Ratings have much larger analyst alphas and larger perceived skill (a measure of past performance adjusted for decreasing returns to scale, which is introduced below). Put differently, Morningstar assigns Analyst Ratings as opposed to Quantitative Ratings to funds that are larger and have performed better in the past, and to funds that Morningstar expects to perform well in the future.

We report our main results for both the sample of "all funds" (i.e., the sample of funds with an Analyst Rating or a Quantitative Rating) and the sample of funds with only an Analyst Rating. In the former case, the sample contains virtually all global equity mutual funds. Concerns about sample selection and the representativeness of funds in our sample should therefore be small. In the latter case, a narrower interpretation of our results is that they "only" apply to the USD 7 trillion managed by the funds with an Analyst Rating.

4 Baseline rational expectations model

In this section, we outline the baseline rational expectations model with which to compare analyst alphas. Similar to Berk and Green (2004), we model the abnormal return of fund i in year $t+1$ as

$$r_{i,t+1} + f_{i,t+1} = a_{i,t} - c(\mathrm{AUM}_{i,t}) + \epsilon_{i,t+1}, \tag{3}$$

where $\epsilon_{i,t+1} \sim N(0, \sigma_\epsilon^2)$, $r_{i,t+1}$ is the fund's net abnormal return, $f_{i,t+1}$ is fees, $a_{i,t}$ is unobservable managerial skill, and the function $c(\mathrm{AUM}_{i,t})$ captures decreasing returns to scale. We refer to $\mathrm{E}_t[r_{i,t+1}]$ as the alpha implied by the rational expectations model.

Following Roussanov et al. (2021), we generalize Berk and Green (2004) to allow for time-varying skill:

$$a_{i,t+1} = (1-\rho)a_0 + \rho a_{i,t} + \sqrt{1-\rho^2} \cdot \nu_{i,t+1}, \tag{4}$$

where $\rho \in [0,1]$, the shock is distributed as $v_{i,t+1} \sim N(0, \sigma_{a,0}^2)$, and skill when a fund is born is distributed as $N(a_0, \sigma_{a,0}^2)$. A rational learner updates her beliefs about managerial skill, i.e., $a_{i,t+1}$ (the only parameter she is uncertain about), from past returns. Allowing for time-varying skill allows the learner to rationally place a greater weight on more recent past performance. A Kalman filter argument implies that beliefs at each point in time are given by:

$$\widehat{a}_{i,t+1} = \rho \left(\widehat{a}_{i,t} + \frac{\widehat{\sigma}^2_{a,t}}{\widehat{\sigma}^2_{a,t} + \sigma^2_{\epsilon}} (r_{i,t+1} - \widehat{a}_{i,t} + c(\text{AUM}_{i,t}) + f_{i,t+1}) \right) + (1-\rho) a_0, \tag{5}$$

$$\widehat{\sigma}^2_{a,t+1} = \rho^2 \widehat{\sigma}^2_{a,t} \left(1 - \frac{\widehat{\sigma}^2_{a,t}}{\widehat{\sigma}^2_{a,t} + \sigma^2_{\epsilon}} \right) + (1-\rho^2) \sigma^2_{a,0}, \tag{6}$$

where $\widehat{\sigma}^2_{a,t+1}$ describes the uncertainty concerning the perceived skill, $\widehat{a}_{i,t+1}$, given initial conditions a_0 and $\sigma^2_{a,0}$. We assume a logarithmic specification for the decreasing returns to scale; that is, $c(\text{AUM}) = \eta \log(\text{AUM})$, where η is a parameter capturing the sensitivity of fund returns to an increase in AUM. We examine a more flexible functional form in Section 7.1 and in the Internet Appendix. The results in the Internet Appendix suggest that the logarithmic specification fits the data well.[6]

We use maximum likelihood to estimate the model on the fund level (using gross fund returns and fund size).[7] We run a performance-evaluation regression as in Equation (1), but over the entire life of a fund using the same benchmark that analysts use, and then form $r_{i,t+1} + f_{i,t+1} = \widehat{\alpha}_i + \zeta_{i,t+1}$, where $\widehat{\alpha}_i$ is the sample average of realized gross abnormal returns.[8] We then annualize the monthly abnormal returns to form the annual abnormal returns. The AUM is measured at the end of the previous year in millions of 2020 USD. Together with the log specification for the decreasing returns to scale, this implies that $a_{i,t}$ is the return on the first USD 1 million invested in the fund.

[6] In the most general version of our model with indexing in Section C.2 of the Internet Appendix, if $\gamma = 1$ (the parameter controlling the shape of decreasing returns to scale) and $\rho = 1$ (constant managerial skill), our model collapses to the model and parameterization in Berk and Green (2004) (see their Equation [11] and their parameterization in their Section IV).

[7] The model assumes that the residuals are uncorrelated across observations. This assumption is more likely to hold for fund returns than share class returns, as the share class returns of a given fund are highly correlated.

[8] One concern is that this procedure could create a bias towards finding decreasing returns to scale similar to the bias that troubles finite-sample fixed effects regressions (see, e.g., Pástor et al., 2015, and note that $\widehat{\alpha}_i$ is a fund fixed effect that is computed using information over the entire life of a fund). In the Internet Appendix, we alternatively estimate $\widehat{\alpha}_i$ using three-year rolling window averages, which eliminates this potential bias. The results are similar.

Table 4 presents the parameter estimates and their standard errors. Our parameter estimates are similar to those of Roussanov et al. (2021). Note that their sample differs from ours, as they focus on U.S.-domiciled funds, whereas we also include funds from other domiciles to be consistent with Morningstar's methodology. The estimated prior mean of managerial skill is 2.30% per year, the prior standard deviation is 2.09%, the residual volatility is 8.11%, and the persistence parameter is 0.95. With a standard deviation of log(AUM) of 1.90, the decreasing returns to scale parameter estimate of 0.25% implies that a one-standard-deviation increase in log(AUM) leads to a 0.48-percentage-point decrease in returns. Alternatively, a doubling of AUM, corresponding to a log increase of 0.69, leads to a 0.17-percentage-point decrease in returns.

The model laid out so far is a filtering problem, independent of the equilibrium argument of Berk and Green (2004). Their equilibrium implication is that alphas are zero at any point in time. Otherwise, the money of risk-neutral investors would flow into and out of funds, affecting alphas through decreasing returns to scale and ultimately competing away any alphas. In contrast, a rational learner who is agnostic to the equilibrium concept expects the abnormal return net of fees to be

$$\mathrm{E}_t[r_{i,t+1}] = \widehat{a}_{i,t} - \eta \log(\mathrm{AUM}_{i,t}) - f_{i,t+1}, \tag{7}$$

which may or may not equal zero. If the rational learner also has rational expectations, she uses the true parameter values of a_0, $\sigma_{a,0}$, η, σ_ϵ, and ρ, which are approximated by our estimates, to form her expectations. We assume rational expectations to form the alphas in December 2020, for every fund according to Equation (7).

5 Main empirical results

5.1 Descriptive statistics

Table 3 shows that analyst alphas are dispersed and obviously inconsistent with the equilibrium implication of a zero alpha for every fund. In fact, analysts actually expect most funds to underperform their benchmarks. The median analyst alpha for the sample of funds with an Analyst or a Quantitative Rating is −124 basis points per year.

Initial evidence that analysts' expectations are tilted towards larger funds comes from the equal- and value-weighted means in Table 3. For the sample of funds with an Analyst or a Quantitative Rating, the equal-weighted mean of analyst alphas is −139 basis points, whereas the value-weighted mean is 51 basis points. This implies that analysts expect the largest funds to outperform significantly.

5.2 Analyst alphas and perceived skill, size, and fees

According to the rational expectations model, three variables determine alphas: perceived skill, fund size, and fees. We start by investigating the univariate relationship between alphas and size. We sort funds into deciles according to their size in December 2020 and then compute average alphas across deciles for both analysts and the rational learner.

Panel (a) of Figure 3 shows the results for the sample of funds with an Analyst Rating and Panel (b) shows the results for the sample of funds with an Analyst or a Quantitative Rating. Analysts' expectations increase with size, whereas the rational learner's expectations are unrelated to size. In general, analysts are more optimistic about funds with an Analyst Rating than about funds with a Quantitative Rating. Since funds with a Quantitative Rating constitute most of the sample in Panel (b), the average analyst alphas are significantly lower in Panel (b) than in Panel (a). The figure also shows that, while analysts are optimistic about the largest funds, they are excessively pessimistic about the smallest funds. This again foreshadows our main conclusion that analysts' expectations are difficult to square with a belief in decreasing returns to scale. However, a belief that larger funds perform better does not necessarily imply a belief in increasing returns to scale: analysts may simply expect larger funds to be able to hire better managers and so perceived managerial skill is an omitted variable. In a similar vein, larger funds may simply charge lower fees.

Therefore, we formally evaluate the rational expectations model in multivariate regressions. One advantage of the model's predictions is that they can be tested using a simple cross-sectional regression. Equation (7), together with the assumption of rational expectations, makes clear predictions for a regression of analyst alphas on size (measured as the logarithm of AUM), perceived skill, and fees: the coefficient estimates should be $-\eta$, 1, and –1, respectively.[9] Table 5 presents two cross-sectional regressions: specification (1) uses the sample of funds with an Analyst Rating; specification (2) uses the sample of funds with an Analyst or a Quantitative Rating. In brackets, we report p-values for the null hypothesis that the coefficients equal the values predicted by the rational expectations model.

Fund size. The estimate on size is statistically positive in both columns and has the opposite sign to that of the model's prediction, which leads us to reject the rational expectations model. For instance, in specification (2) the coefficient estimate on size is 0.13% as opposed to -0.25%.

Perceived skill. As the rational expectations model predicts, greater perceived skill is associated with a larger analyst alpha. However, the coefficient estimate on perceived skill is smaller than and statistically different from one in both specifications.

[9] Moreover, in theory the constant should be zero and the R^2 should be 100%; similarly, in theory the error terms are homoscedastic. In our empirical analysis, we allow for more conservative standard errors clustered by fund family. For our main results, we also focus on net-of-fee alphas; the main result is similar when we take fees out of the equation and impose the restriction that the coefficient on fees is equal to -1.

Fees. As the rational expectations model predicts, an increase in fees is associated with a decrease in analyst alpha. The coefficient estimate on fees is not statistically different from minus one in specification (1), but is statistically different from minus one in specification (2).

You may be concerned that our regressions omit other variables, correlated with both analysts' unobserved perceptions of managerial skill and size, that bias the coefficient estimate on size. This is a valid concern—exogenous variation in size is difficult to obtain.

However, Figure 1 shows that we do not even need to identify the effect of size on analysts' expectations to argue that analysts' expectations are tilted too much towards larger funds. That said, the figure is consistent with two interpretations. Under a first interpretation, analyst alphas for the funds that have grown to be the largest are too large because analysts perceive these funds to be me much more skilled than they actually are (a too large $\widehat{a}_{i,t}$)—while still believing that an increase in size deteriorates future returns. Under a second interpretation, analyst alphas for the largest funds are too large because analysts do not believe that an increase in size actually deteriorates future returns (a wrong η). By imposing structure and modeling alphas as a linear function of perceived managerial skill and size, the results of this subsection support the latter interpretation.

Finally, we add additional variables to our empirical specifications in the next subsection and extend the model in various ways in the robustness section, but the estimates on size remain positive.

5.3 Additional determinants of expectations

Morningstar's methodology suggests that the rational expectations model omits variables relevant to analysts' expectation formation. We are guided by Morningstar's methodology and previous research in choosing additional variables to explain analysts' expectations. We group variables corresponding to the three pillars "People," "Process," and "Parent." Most of our variables can be obtained directly from Morningstar Direct, which ensures that they are available to analysts. We then simply include these variables in reduced form in our cross-sectional regressions.[10]

For "People," we include manager tenure (the longest tenure, in months, of the managers of a fund), manager ownership (the average dollar amount managers of a fund personally invest in the fund), managerial multitasking (the average number of additional funds that the managers of a fund manage), and a dummy for whether a fund is team managed. Manager ownership has been shown to predict fund performance in the U.S. and Sweden (see, e.g., Khorana et al., 2007; Ibert, 2023). However, since ownership information is only publicly available for U.S.-domiciled funds, our sample is restricted.[11]

For "Process," we include a fund's top 10 assets (the percentage of AUM in the ten largest positions), a fund's tracking error (the standard deviation of returns in excess of the benchmark over the life of the fund), fund turnover (as reported to the SEC), a dummy for whether a fund is primarily held by retail investors, and a dummy for whether a fund is primarily sold through a broker.[12] Top 10 assets and tracking error serve as measures of diversification and activeness, respectively. There is evidence that more active funds outperform (see, e.g., Cremers and Petajisto, 2009). In contrast, broker-sold funds and funds held primarily by retail investors have underperformed on average (Bergstresser, Chalmers,

[10]An alternative approach would be to include additional variables in our structural estimation via the measurement equation, Equation (3). One caveat to this approach is that, for many of our additional variables, time series are not readily available from Morningstar Direct.

[11]As of 2005, the SEC requires that mutual fund managers publicly report personal investments in their own funds. Managers must report whether their dollar ownership in their funds falls into one of the following ranges: USD 0, USD 1–10,000, USD 10,001–50,000, USD 50,001–100,000, USD 100,001–500,000, USD 500,001–1,000,000, or above USD 1,000,000. As done by Khorana et al. (2007), we use midpoints of the disclosed ownership ranges to calculate manager ownership, except for the maximum range, "USD 1,000,001 and above," for which we use the bottom of the range.

[12]We winsorize fund turnover at the 1st and 99th percentiles as done by Pástor, Stambaugh, and Taylor (2017) and do the same with the top 10 assets. The retail dummy takes the value of one if more than 2/3 of a fund's assets come from share classes open to retail investors. The broker-sold dummy takes the value of one if more than 2/3 of a fund's assets come from share classes that charge front-end or back-end loads or a 12b-1 fee of more than 0.25%.

and Tufano, 2009; Del Guercio and Reuter, 2014).

For "Parent," we include fund family fixed effects. The literature on the role of the fund family has highlighted the fund family's impact on individual fund performance (see, e.g., Massa, 2003; Gaspar, Massa, and Matos, 2006; Ferreira, Matos, and Pires, 2018).

Since our measure summarizing past fund performance—perceived skill—requires a belief in decreasing returns to scale to compute it, we also control for alternative measures of past performance that analysts may consider. Morningstar Star Ratings are a prominent alternative measure of past performance, so we include Morningstar Star Rating fixed effects. We also include Morningstar Category and Sustainability Rating fixed effects. Overall, our set of controls is extensive. The effect of size on analyst alphas is identified from variation across funds within the same fund family, within the same category, with the same Star and Sustainability ratings, and with the same levels of the various observables we consider.

Table 6 shows four specifications. The first two are for the sample of U.S.-domiciled funds with an Analyst Rating and the last two are for the sample of all rated U.S.-domiciled funds. Specifications (1) and (3) replicate the specifications in Table 5 for the restricted sample of U.S.-domiciled funds and show similar results. Specifications (2) and (4) include "People" and "Process" variables as well as various fixed effects. We standardize "People" and "Process" variables to mean zero and unit standard deviation, but leave perceived skill, size, and fees unstandardized for comparison to previous tables.

As expected, other characteristics besides perceived skill, size, and fees are important to analysts' expectation formation. In both specifications (2) and (4), manager tenure, manager ownership, and managerial multitasking are positively related to analysts' expectations. In specification (4), one-standard-deviation increases in tenure and ownership increase analyst alphas by 0.25- and 0.19-percentage-points, respectively. In contrast, funds predominantly held by retail investors are expected to perform worse, consistent with earlier evidence on the realized performance of such funds.

The point estimates on fund size become smaller both economically and statistically, suggesting that some of the additional characteristics are correlated with both size and expected returns. Nonetheless, the point estimates on size remain positive in all columns. Most importantly, the point estimates are still far from the -0.25 point estimate implied by the rational expectations model. The p-value for the null hypothesis that the coefficient equals -0.25 is 0.00.

Another piece of evidence comes from the coefficient estimates on fees. The impact of fund size on fund returns is perhaps hard to grasp given the sophistication required to detect decreasing returns to scale in realized fund returns and some mixed empirical evidence in previous studies. However, common sense suggests that, all else being equal, a one-percentage-point increase in fees should decrease expected returns by one percentage point. The estimates on fees in (2) and (4) are close to minus one and not statistically different from minus one, suggesting that these specifications satisfy this basic principle of common sense. These specifications give us confidence that we have not overlooked other important characteristics that could, once included, lead to a negative coefficient estimate on size.

In fact, R^2 values of above 60% suggest that specifications (2) and (4) capture analyst alphas reasonably well. The increases in R^2 values are driven by the inclusion of fund family fixed effects. We hypothesize that governance and incentives could play a large role. For instance, fund manager compensation practices are likely important and have been shown to differ systematically across fund families (Ibert, Kaniel, Van Nieuwerburgh, and Vestman, 2018; Ma, Tang, and Gómez, 2019).

6 Analysts' expectations and investors' expectations

We study *analysts'* subjective expectations. Analysts could be akin to sophisticated investors, but in general our paper says little about *investors'* subjective expectations. A valid approach for learning about investors' subjective expectations is to directly survey investors. However, surveys entail well-known drawbacks, as is explained in detail in Choi and Robertson (2020). For instance, it is unclear whether survey respondents act on their expectations and, thus, whether their expectations are reflected in their capital allocations.

While we do not observe investors' subjective expectations, one advantage of working with mutual fund data is that we can test whether better ratings lead to larger investor fund flows. The Internet Appendix shows that they do, using the ordinal ratings that are available for a longer time series. That flows follow ratings shows that analysts' expectations matter to some investors, regardless of whether these investors have the same expectations of future performance, have different expectations, or have even formed their expectations.

Figure 4 summarizes the results regarding flows shown in the Internet Appendix. The figure shows coefficient estimates on Star Rating dummies, Analyst Rating dummies, and Quantitative Rating dummies in a regression of monthly fund flows on the dummies, a battery of control variables, and fund, year–month, as well as category fixed effects (see also Armstrong et al., 2019). The effect of the Analyst Rating on flows can be close to the effect of the popular Star Rating. For instance, when a fund with no Star Rating is assigned a five-star rating, monthly flows increase by 1.39 percentage points (i.e., $= 1.56 - 0.17$). Similarly, when a fund with no Analyst Rating is assigned a Gold Analyst Rating, monthly flows increase by 1.14 percentage points. In contrast, while statistically significant, the effect of Quantitative Ratings on flows is considerably smaller.

7 Additional issues

7.1 Robustness

This section summarizes some robustness tests. The Internet Appendix discusses these and other robustness tests in more detail.

The results are robust to controlling for value added, a generic measure of skill that does not rely on any model's particular assumptions to derive perceived managerial skill (Berk and van Binsbergen, 2015). For our main results, we have assumed a logarithmic functional form for the decreasing returns to scale technology. We re-estimate the baseline model using a more flexible functional form (see also Roussanov et al., 2020). The results suggest that the logarithmic assumption fits the data well. We also consider specifications that allow the impact of size on returns to vary across funds based on common characteristics. Consistent with Pástor et al. (2015), we do find that funds with higher turnover, funds that invest in small-cap stocks, and funds that are more active face steeper decreasing returns to scale in realized fund returns. However, none of these patterns are mirrored in analysts' expectations. We also extend the baseline model to account for uncertainty in the decreasing returns to scale parameter and industry size (Pástor and Stambaugh, 2012). In the former case, the effect of size on returns varies fund-by-fund (as in Barras et al., 2022), just like managerial skill. While we cannot estimate cross-sectional regressions in this case, our conclusion is robust: as in our main analysis, analysts' expectations are too large for the largest funds and too small for the vast majority of other funds. To allow for a structural break in the relationship between returns, skill, size, and fees in our model, we also estimate the baseline model using only funds incepted since 2000, which is the first year of data that enters Morningstar's methodology through the SIQR computation. Again, the results are robust.

Our regressions of expectations on fund characteristics identify the coefficient estimate on size using cross-sectional variation. The Internet Appendix shows that our results are robust to estimating an ordered logit model with fund fixed effects using the ordinal ratings that are available since 2011. These regressions are analogous to the regressions that researchers have estimated using realized before-fee fund returns and identify the coefficient estimate using time-series variation (see, e.g., Pástor et al., 2015). However, fund fixed effects are less powerful in our context. Intuitively, fund fixed effects control for analysts' perceptions of skill that remain constant over time, but such perceptions most likely vary over time as analysts update their beliefs about true skill.[13] We also document robust evidence of decreasing returns to scale in realized returns using the fund fixed effects recursive demeaning estimator of Zhu (2018). After the initial writing of this paper, we have also updated the data up to December 2021 and conducted an out-of-sample test of our main results. Again, the Internet Appendix shows that the results are robust.

7.2 Conflicts of interest

A general concern when studying analysts' expectations is that biases in expectations may not necessarily reflect cognitive misunderstandings. For instance, if Morningstar or its analysts had misguided incentives to assign better ratings to larger funds, analysts' expectations would not necessarily reflect a genuine cognitive misunderstanding of returns to scale.

We believe that such conflicts of interest are limited. Morningstar claims that its research activities are independent of its commercial activities. Moreover, as a leading financial services firm in the mutual fund industry, Morningstar has a substantial business reputation

[13] For a similar reason, the predictability of forecast errors would not be powerful evidence against rational expectations models of active management. Too see this, consider the investors in Pástor and Stambaugh (2012). Investors in their model continue to expect positive returns from active management even though active management repeatedly underperforms. Thus, forecast errors are predictable, even though the investors in Pástor and Stambaugh (2012) clearly have rational expectations. Similarly, forecast errors in Berk and Green (2004) and in the models of our paper are predictable. The reason for the predictability of forecast errors in all these cases is the wedge between true skill and perceived skill that is induced by parameter uncertainty and learning.

at stake. In contrast to credit-rating issuers, Morningstar does not receive a fee from fund issuers for its fund analysis. Finally, Morningstar's primary business model does not entail acting as a seller of mutual funds, so it is likely not subject to the conflicts of interest that have been shown to affect broker-sold funds (see, e.g., Bergstresser et al., 2009). In line with these arguments, Cookson et al. (2021) use the Morningstar Analyst Rating as a benchmark of independent analysis when studying investment platforms' mutual fund recommendations.

7.3 Textual analysis of written reports

In a similar vein, one may wonder whether the Analyst Ratings do not account for the effect of fund size on return expectations by design. There is no pillar rating for the effect of fund size on fund returns, so it is not immediately clear at which point in Morningstar's methodology such an effect should enter. On one hand, if true, such a design flaw would of course trivially support our conclusion: realized returns decrease with size, but expected returns do not (by design). On the other hand, this conclusion would perhaps be less interesting, as expectations would not truly reflect a cognitive misunderstanding of returns to scale in active management by analysts, but rather a design flaw on Morningstar's part.

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 (2023), who collects an exhaustive 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 5 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 into their assessments. In the context of Figure 1, analysts' extrapolation of past returns also does not happen mechanically: nothing restricts analysts from assigning lower pillar scores to the funds that have grown to be the largest to bring down expectations of future returns.

8 Conclusion

We introduce data on subjective expectations to the mutual fund literature. We find that there is little evidence that analysts form their expectations as in a workhorse model and so a discussion seems warranted about whether we—researchers in this area—can build more realistic models of active management.

Given no evidence of decreasing returns to scale in analysts' expectations even after decades of potential learning, building rational expectations models to match analysts' expectations might be challenging. Future research could also depart from the rational expectations assumption in developing models to match analysts' expectations. Such development would be similar to the development of asset pricing models to match extrapolative subjective stock market return expectations (see, e.g., Barberis, Greenwood, Jin, and Shleifer, 2015; Adam, Marcet, and Beutel, 2017; Nagel and Xu, 2022). The models of active management of Gennaioli et al. (2015) and Spiegler (2020) allow for deviations from rational expectations and therefore could constitute starting points.

As mentioned in the introduction, such models could hardly be representative agent models: expectations that increase with size imply that all funds should receive unlimited amounts of capital. Expectations that merely do not decrease with size—as opposed to increase with size—imply that all funds with a negative alpha should manage no capital and all funds with a positive alpha should manage unlimited amounts of capital. It follows that misunderstandings of returns to scale in active management could help explain the enormous size and poor performance of the active fund industry. An investor who believes that returns increase with size allocates more and more capital to funds in the hope that the additional capital aids funds to earn better future returns. However, this additional capital actually deteriorates future returns due to decreasing returns to scale in realized returns.

References

Adam, Klaus, Albert Marcet, and Johannes Beutel, 2017, Stock Price Booms and Expected Capital Gains, *American Economic Review* 107, 2352–2408.

Armstrong, Will J., Egemen Genc, and Marno Verbeek, 2019, Going for Gold: An Analysis of Morningstar Analyst Ratings, *Management Science* 65, 2310–2327.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2015, X-CAPM: An Extrapolative Capital Asset Pricing Model, *Journal of Financial Economics* 115, 1–24.

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

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2022, What Do Mutual Fund Investors Really Care About?, *Review of Financial Studies* 35, 1723–1774.

Bender, Svetlana, James J. Choi, Danielle Dyson, and Adriana Z. Robertson, 2022, Millionaires Speak: What Drives Their Personal Investment Decisions?, *Journal of Financial Economics* 146, 305–330.

Bergstresser, Daniel, John M.R. Chalmers, and Peter Tufano, 2009, Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry, *Review of Financial Studies* 22, 4129–4156.

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 Ian Tonks, 2007, Return Persistence and Fund Flows in the Worst Performing Mutual Funds, *Working Paper*.

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

Berk, Jonathan B., and Jules H. van Binsbergen, 2017, Mutual Funds in Equilibrium, *Annual Review of Financial Economics* 9, 147–167.

van Binsbergen, Jules H., Jeong H. Kim, and Soohun Kim, 2021, Capital Allocation and the Market for Mutual Funds: Inspecting the Mechanism, *Working Paper*.

Blake, Christopher R., and Matthew R. Morey, 2000, Morningstar Ratings and Mutual Fund Performance, *Journal of Financial and Quantitative Analysis* 35, 451–483.

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

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer, 2020, Overreaction in Macroeconomic Expectations, *American Economic Review* 110, 2748–2782.

Busse, Jeffrey A., Tarun Chordia, Lei Jiang, and Yuehua Tang, 2021, Transaction Costs, Portfolio Characteristics, and Mutual Fund Performance, *Management Science* 67, 1227–1248.

Carhart, Mark M., 1997, On Persistence in Mutual Fund Performance, *Journal of Finance* 52, 57–82.

Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey D. Kubik, 2004, Does Fund Size Erode Mutual Fund Performance? The Role of Liquidity and Organization, *American Economic Review* 94, 1276–1302.

Choi, James J., and Adriana Z. Robertson, 2020, What Matters to Individual Investors? Evidence from the Horse's Mouth, *Journal of Finance* 75, 1965–2020.

Coibion, Olivier, and Yuriy Gorodnichenko, 2012, What Can Survey Forecasts Tell Us about Information Rigidities?, *Journal of Political Economy* 120, 116–159.

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.

Cookson, Gordon, Tim Jenkinson, Howard Jones, and Jose Vicente Martinez, 2021, Best Buys and Own Brands: Investment Platforms' Recommendations of Mutual Funds, *Review of Financial Studies* 34, 227–263.

Cremers, Martijn K.J., and Antti Petajisto, 2009, How Active Is Your Fund Manager? A New Measure that Predicts Performance, *Review of Financial Studies* 22, 3329–3365.

Dangl, Thomas, Youchang Wu, and Josef Zechner, 2008, Market Discipline and Internal Governance in the Mutual Fund Industry, *Review of Financial Studies* 21, 2307–2343.

Del Guercio, Diane, and Jonathan Reuter, 2014, Mutual Fund Performance and the Incentive to Generate Alpha, *Journal of Finance* 69, 1673–1704.

Del Guercio, Diane, and Paula A. Tkac, 2008, Star Power: The Effect of Morningstar Ratings on Mutual Fund Flow, *Journal of Financial and Quantitative Analysis* 43, 907–936.

Dyakov, Teodor, Hao Jiang, and Marno Verbeek, 2020, Trade Less and Exit Overcrowded Markets: Lessons from International Mutual Funds, *Review of Finance* 24, 677–731.

Evans, Allison L., 2008, Portfolio Manager Ownership and Mutual Fund Performance, *Financial Management* 37, 513–534.

Evans, Richard B., and Yang Sun, 2021, Models or Stars: The Role of Asset Pricing Models and Heuristics in Investor Risk Adjustment, *Review of Financial Studies* 34, 67–107.

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.

Ferreira, Miguel A., Pedro Matos, and Pedro Pires, 2018, Asset Management within Commercial Banking Groups: International Evidence, *Journal of Finance* 73, 2181–2227.

Franzoni, Francesco, and Martin C. Schmalz, 2017, Fund Flows and Market States, *Review of Financial Studies* 30, 2621–2673.

Gârleanu, Nicolae, and Lasse Heje Pedersen, 2018, Efficiently Inefficient Markets for Assets and Asset Management, *Journal of Finance* 73, 1663–1712.

Gaspar, José Miguel, Massimo Massa, and Pedro Matos, 2006, Favoritism in Mutual Fund Families? Evidence On Strategic Cross-Fund Subsidization, *Journal of Finance* 61, 73–104.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny, 2015, Money Doctors, *Journal of Finance* 70, 91–114.

Glode, Vincent, and Richard C. Green, 2011, Information Spillovers and Performance Persistence for Hedge Funds, *Journal of Financial Economics* 101, 1–17.

Greenwood, Robin, and Stefan Nagel, 2009, Inexperienced Investors and Bubbles, *Journal of Financial Economics* 93, 239–258.

Greenwood, Robin, and Andrei Shleifer, 2014, Expectations of Returns and Expected Returns, *Review of Financial Studies* 27, 714–746.

Hartzmark, Samuel M., and Abigail B. Sussman, 2019, Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows, *Journal of Finance* 74, 2789–2837.

Harvey, Campbell R., and Yan Liu, 2022, Luck versus Skill in the Cross-Section of Mutual Fund Returns: Reexamining the Evidence, *Journal of Finance* 77, 1921–1966.

Harvey, Campbell R., Yan Liu, Eric K. M. Tan, and Min Zhu, 2021, Crowding: Evidence from Fund Managerial Structure, *Working Paper.*

Ibert, Markus, 2023, What Do Mutual Fund Managers' Private Portfolios Tell Us About Their Skills?, *Journal of Financial Intermediation* 53, 100999.

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.

Jenkinson, Tim, Howard Jones, and Jose Vicente Martinez, 2016, Picking Winners? Investment Consultants' Recommendations of Fund Managers, *Journal of Finance* 71, 2333–2370.

Jones, Howard, and Jose Vicente Martinez, 2017, Institutional Investor Expectations, Manager Performance, and Fund Flows, *Journal of Financial and Quantitative Analysis* 52, 2755–2777.

Khorana, Ajay, and Edward Nelling, 1998, The Determinants and Predictive Ability of Mutual Fund Ratings, *Journal of Investing* 7, 61–66.

Khorana, Ajay, Henri Servaes, and Lei Wedge, 2007, Portfolio Manager Ownership and Fund Performance, *Journal of Financial Economics* 85, 179–204.

Kim, Jeong Ho, 2022, Investor Learning and the Aggregate Allocation of Capital to Active Management, *Working Paper.*

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

Linnainmaa, Juhani T., 2013, Reverse Survivorship Bias, *Journal of Finance* 68, 789–813.

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.

Ma, Linlin, Yuehua Tang, and Juan Pedro Gómez, 2019, Portfolio Manager Compensation in the U.S. Mutual Fund Industry, *Journal of Finance* 74, 587–638.

Massa, Massimo, 2003, How Do Family Strategies Affect Fund Performance? When Performance-Maximization Is Not the Only Game in Town, *Journal of Financial Economics* 67, 249–304.

McLemore, Ping, 2019, Do Mutual Funds Have Decreasing Returns to Scale? Evidence from Fund Mergers, *Journal of Financial and Quantitative Analysis* 54, 1683–1711.

Nagel, Stefan, and Zhengyang Xu, 2022, Asset Pricing with Fading Memory, *Review of Financial Studies* 35, 2190–2245.

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, and Lucian A. Taylor, 2017, Do Funds Make More When They Trade More?, *Journal of Finance* 72, 1483–1528.

Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor, 2020, Fund Tradeoffs, *Journal of Financial Economics* 138, 614–634.

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.

Reuter, Jonathan, and Eric Zitzewitz, 2021, How Much Does Size Erode Mutual Fund Performance? A Regression Discontinuity Approach, *Review of Finance* 25, 1395–1432.

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.

Sharpe, William F., 1998, Morningstar's Risk-Adjusted Ratings, *Financial Analysts Journal* 54, 21–33.

Song, Yang, 2020, The Mismatch Between Mutual Fund Scale and Skill, *Journal of Finance* 75, 2555–2589.

Spiegler, Ran, 2020, A Simple Model of a Money-Management Market with Rational and Extrapolative Investors, *European Economic Review* 127, 1–17.

Wilke, Felix, 2023, Mutual Fund Analysts as Information Intermediaries, *Working Paper*.

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

Table 1: Overview of Morningstar's fund ratings

	Star Rating	Analyst Rating	Quantitative Rating	Sustainability Rating
Introduction	1985	2011	2017	2016
Key inputs	Historical fund returns	*New*: Three-pillar ratings (People, Process, and Parent), SIQR (dispersion of CAPM alphas of fund strategy), and share-class fees *Old*: Five-pillar ratings (People, Process, Parent, Performance, and Price)	*New*: Three-pillar ratings (People, Process, and Parent) estimated using a machine-learning algorithm, SIQR (dispersion of CAPM alphas of fund strategy), and share-class fees *Old*: Five-pillar ratings (People, Process, Parent, Performance, and Price) estimated using a machine-learning algorithm	Sustainalytics' company-level ESG Risk Rating
Backward- or forward-looking	Backward-looking	Forward-looking	Forward-looking	Forward-looking
Rating scale	***** **** *** ** *	Gold Silver Bronze Neutral Negative	Gold Silver Bronze Neutral Negative	5 globes 4 globes 3 globes 2 globes 1 globe
Rating level	Share class	*New*: Share class *Old*: Fund	Share class	Fund
Ranking metric to award ratings	Morningstar Risk-Adjusted Return	Share-class alphas from Analyst and Quantitative Rating methodology	Share-class alphas from Analyst and Quantitative Rating methodology	Morningstar Historical Portfolio Sustainability Score

Continued on next page

Table 1 continued from previous page

	Star Rating	Analyst Rating	Quantitative Rating	Sustainability Rating
Rating peer group	Morningstar Category	Morningstar Category	Morningstar Category	Morningstar Global Category
Medalist ranking (Gold, Silver, and Bronze) requirement		*New*: Beat benchmark index *and* peer group average *Old*: Beat benchmark index *and/or* peer group average	*New*: Beat benchmark index *and* peer group average *Old*: Beat benchmark index *and/or* peer group average	
Major updates	06/2002: Ratings assigned within Morningstar Categories (before broad asset classes, e.g., equity)	10/2019: Ratings assigned at share-class level based on expected net-of-fee alphas, reduction to three pillars, and higher bar for medalist ranking	10/2019: Ratings assigned at share-class level based on expected net-of-fee alphas, reduction to three pillars, and higher bar for medalist ranking	10/2019: Replacement of Sustainalytics' company ESG Rating with its ESG Risk Rating
Selected academic sources and sample periods for the analysis	Ben-David et al. (2022), 1991–2011; Blake and Morey (2000), 1992–1997; Del Guercio and Tkac (2008), 1996–1999; Evans and Sun (2021), 1999–2005; Khorana and Nelling (1998), 1992–1995; Sharpe (1998)	Armstrong et al. (2019), 2011–2015		Hartzmark and Sussman (2019), 2016–2017

The table compares key features of Morningstar fund ratings. The Morningstar Rating (commonly referred to as the Star Rating) is a purely quantitative, backward-looking measure of a fund's past performance. The Morningstar Analyst Rating is forward looking and conveys an analyst's conviction of a fund's investment merits. The Morningstar Quantitative Rating is derived from a machine-learning model and attempts to replicate the Analyst Rating a human Morningstar analyst might assign to a fund. The Morningstar Sustainability Rating assesses the risk exposure of an investment portfolio to environmental, social, and governance (ESG) factors.

Table 2: **Replication of Morningstar Analyst and Quantitative Ratings**

Panel A: Morningstar Analyst Ratings

	Replicated rating						
Actual rating	Negative	Neutral	Bronze	Silver	Gold	Total	Rate
Negative	80	15	0	0	0	95	84%
Neutral	60	3035	121	2	0	3218	94%
Bronze	2	167	2293	201	10	2673	86%
Silver	0	1	213	1731	107	2052	84%
Gold	0	0	0	88	571	659	87%
Total	142	3218	2627	2022	688	8697	89%

Panel B: Morningstar Quantitative Ratings

	Replicated rating						
Actual rating	Negative	Neutral	Bronze	Silver	Gold	Total	Rate
Negative	12557	503	0	0	0	13060	96%
Neutral	416	26150	396	1	0	26963	97%
Bronze	2	906	6378	376	12	7674	83%
Silver	0	12	559	4252	179	5002	85%
Gold	0	3	1	312	2328	2644	88%
Total	12975	27574	7334	4941	2519	55343	93%

The table shows how well Morningstar Analyst and Quantitative Ratings on the share class level under the new ratings methodology are replicated for the cross-section of funds in December 2020. The actual Morningstar Analyst Ratings are tabulated in rows, whereas the replicated ratings are tabulated in columns. The column *Rate* indicates the percentage of ratings that we can replicate (e.g., we assign a Neutral rating to 3035 out of 3218 analyst-rated share classes receiving a Morningstar Analyst Rating of Neutral, yielding a replication rate of 94%).

Table 3: Summary statistics

	N	Mean (V.W.)	Mean (E.W.)	S.D.	10%	25%	50%	75%	90%
Panel A: Assets under management									
Analyst Rating	1454		4760	14036	154	406	1248	3882	10098
Quantitative Rating	12480		409	1257	10	30	100	336	931
All ratings	13934		863	4871	12	34	126	464	1477
No rating	4512		155	1251	6	13	37	112	291
All	18446		690	4290	9	25	89	341	1144
Panel B: Fees									
Analyst Rating	1454	0.79	1.06	0.39	0.64	0.84	1.00	1.24	1.59
Quantitative Rating	12480	1.11	1.44	0.72	0.65	0.96	1.36	1.82	2.27
All ratings	13934	0.92	1.40	0.70	0.64	0.94	1.29	1.77	2.23
No rating	4512	1.28	1.65	0.93	0.78	1.06	1.63	1.97	2.44
All	18446	0.94	1.46	0.77	0.67	0.95	1.37	1.84	2.27
Panel C: Perceived skill									
Analyst Rating	1454	3.21	2.90	0.92	1.84	2.28	2.78	3.39	4.10
Quantitative Rating	12480	2.69	2.29	0.94	1.27	1.75	2.26	2.74	3.39
All ratings	13934	2.99	2.36	0.95	1.32	1.79	2.30	2.83	3.50
No rating	4512	2.96	2.41	1.10	1.38	1.96	2.30	2.63	3.58
All	18446	2.99	2.37	0.99	1.34	1.83	2.30	2.78	3.51
Panel D: Analyst alphas									
Analyst Rating	1454	1.29	0.60	1.35	–1.09	–0.25	0.69	1.42	2.24
Quantitative Rating	12480	–0.55	–1.62	2.48	–4.82	–3.21	–1.57	0.04	1.50
All ratings	13934	0.51	–1.39	2.49	–4.67	–2.99	–1.24	0.34	1.66

The table shows value-weighted (V.W., by assets under management, AUM) and equal-weighted (E.W.) means, standard deviations, and various percentiles of AUM, fees, skill, and analyst alphas for global active equity mutual funds in December 2020. AUM is the fund size in millions of USD. Perceived skill is managerial skill estimated from a rational model of fund performance. Alphas are relative to each fund's Morningstar Category benchmark. Fees, perceived skill, and analyst alphas are expressed in % per year.

Table 4: Parameter estimates of the rational fund performance model

Parameter	Description	Estimate
η	Decreasing returns to scale (%)	0.251***
		(0.013)
a_0	Prior mean (%)	2.296***
		(0.063)
$\sigma_{a,0}$	Prior standard deviation (%)	2.095***
		(0.042)
σ_ϵ	Residual standard deviation (%)	8.111***
		(0.015)
ρ	Skill persistence	0.948***
		(0.006)

The table shows the parameter estimates of the rational fund performance model 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 5: Cross-sectional regressions of alphas on fund characteristics

	Analyst Ratings	Analyst and Quantitative Ratings
	(1)	(2)
Perceived skill	0.382***	0.706***
	(0.066)	(0.041)
	[0.000]	[0.000]
Size (× 100)	0.066**	0.133***
	(0.032)	(0.025)
	[0.000]	[0.000]
Fees	−0.959***	−1.536***
	(0.150)	(0.060)
	[0.787]	[0.000]
Constant (× 100)	0.042	−1.553***
	(0.276)	(0.174)
	[0.878]	[0.000]
N	1454	13934
Adj. R^2	0.15	0.32

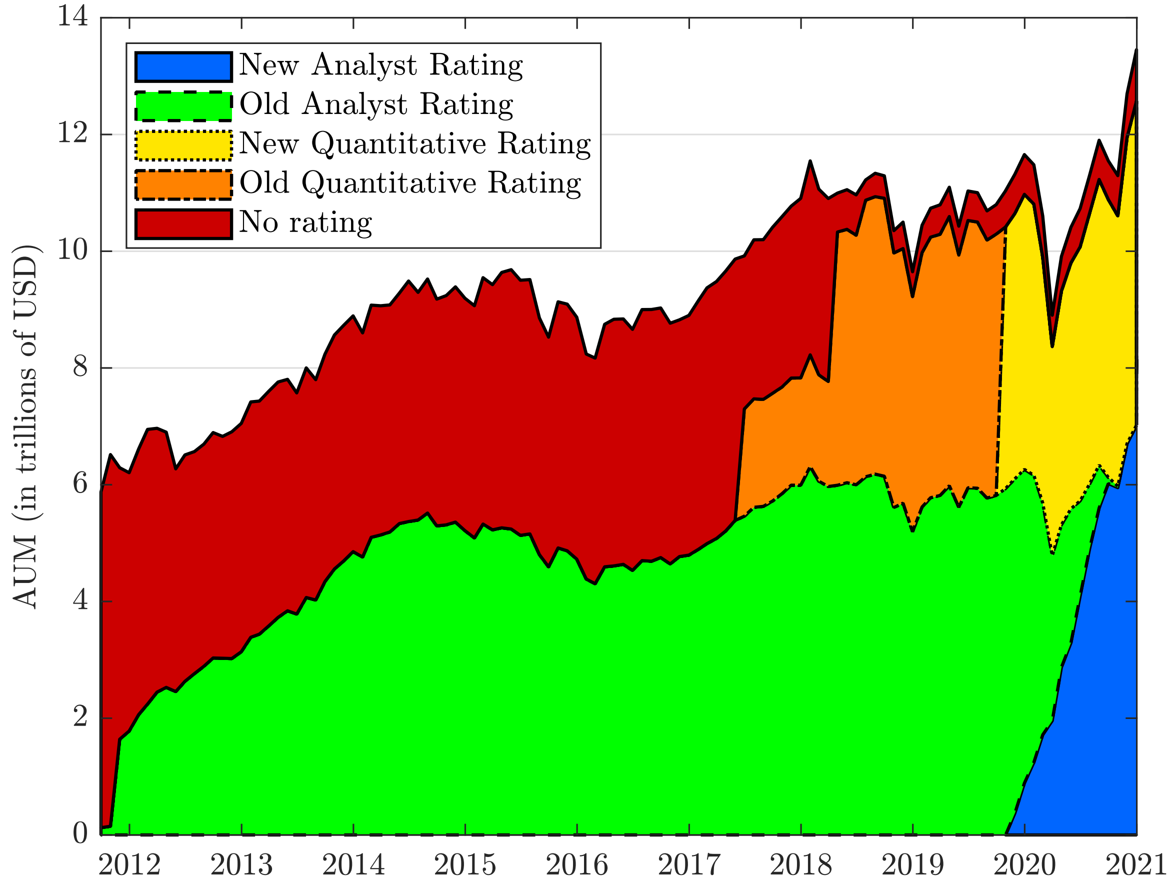
The table shows regressions of Morningstar analyst alphas on skill as perceived by a rational learner, fund size (logarithm of assets under management in millions of USD), and fees for cross-sections of funds in December 2020. Specification (1) uses funds with an Analyst Rating. Specification (2) uses 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, fees, and the constant equal the model-predicted parameters of +1, −0.251 (the estimate of η in Table 4), −1, and 0, respectively.

Table 6: Cross-sectional regressions of alphas on additional fund characteristics

	Analyst Ratings		Analyst and Quantitative Ratings	
	(1)	(2)	(3)	(4)
Rational learner				
Perceived skill	0.268***	0.106	0.860***	0.342***
	(0.066)	(0.071)	(0.080)	(0.058)
Size (× 100)	0.164***	0.078**	0.111***	0.052*
	(0.045)	(0.038)	(0.028)	(0.029)
Fees	−1.350***	−0.947***	−1.768***	−0.960***
	(0.146)	(0.115)	(0.195)	(0.211)
People				
Manager tenure		0.111***		0.247***
		(0.040)		(0.033)
Manager ownership		0.115**		0.192***
		(0.054)		(0.041)
Managerial multitasking		0.645***		0.576***
		(0.205)		(0.193)
Management team		0.094		0.496***
		(0.109)		(0.114)
Process				
Top 10 assets (%)		0.128		−0.028
		(0.131)		(0.091)
Tracking error		−0.010		−0.154*
		(0.064)		(0.093)
Turnover ratio		−0.486***		−0.108
		(0.156)		(0.081)
Retail		−0.290***		−0.157*
		(0.092)		(0.088)
Broker-sold		−0.267**		−0.068
		(0.116)		(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

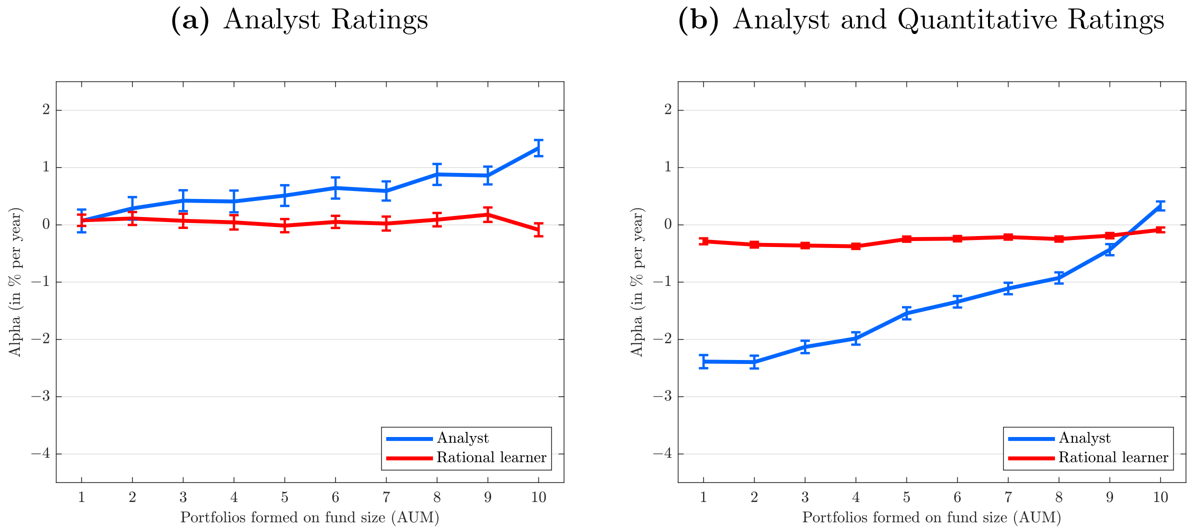
The table shows regressions of Morningstar analyst alphas on fund and manager characteristics for cross-sections of funds in December 2020. Specifications (1) and (2) use U.S.-domiciled funds with an Analyst Rating. Specifications (3) and (4) use U.S.-domiciled funds with an Analyst Rating or a Quantitative 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.

Figure 2: Size of active equity mutual fund industry



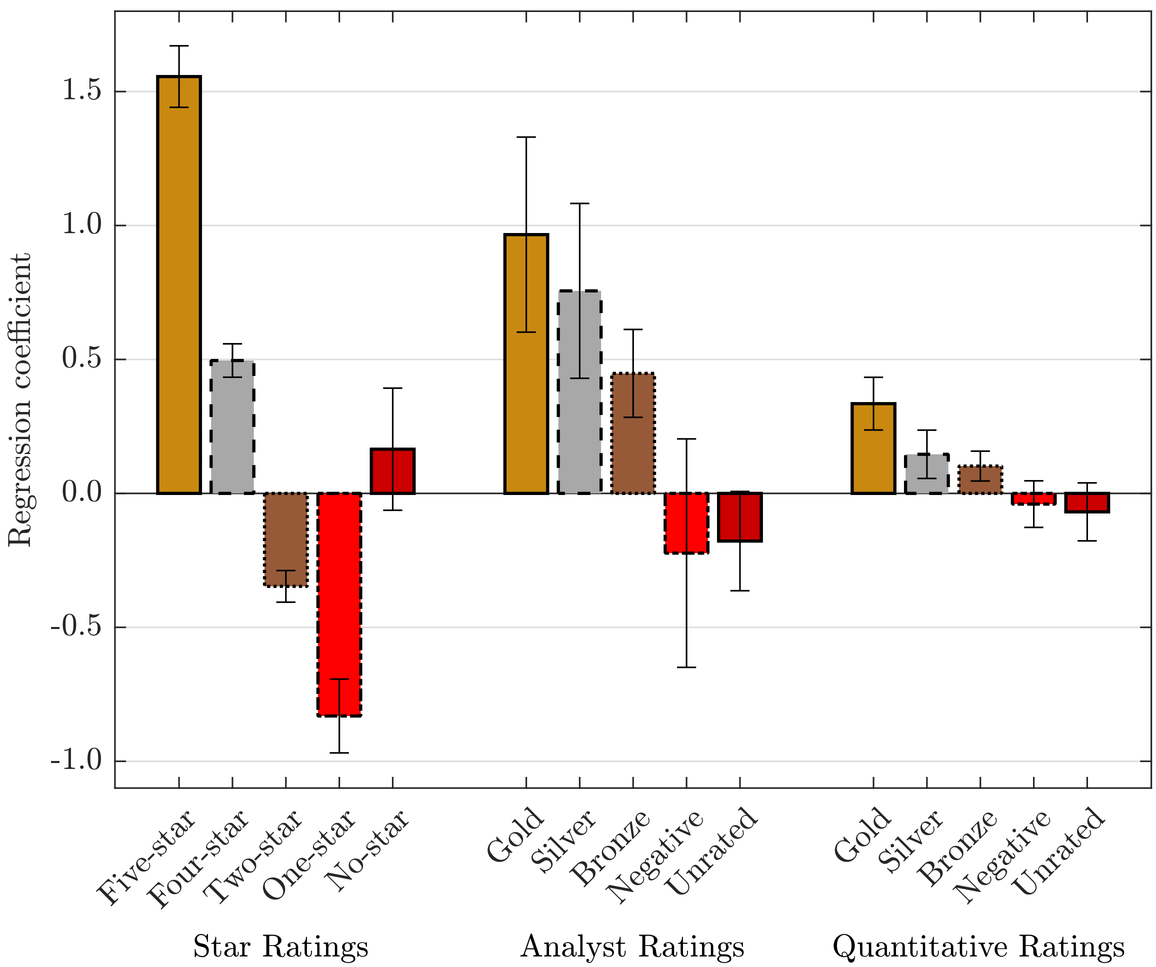
The figure shows the assets under management (AUM) of actively managed equity mutual funds up to December 2020. New Analyst Rating indicates funds with a Morningstar Analyst Rating according to the new methodology. Old Analyst Rating indicates funds with a Morningstar Analyst Rating under the old methodology. Similarly, Old Quantitative Rating and New Quantitative Rating indicate funds with a Morningstar Quantitative Rating under the old and new methodologies, respectively.

Figure 3: Alphas against fund size



The figure shows 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.

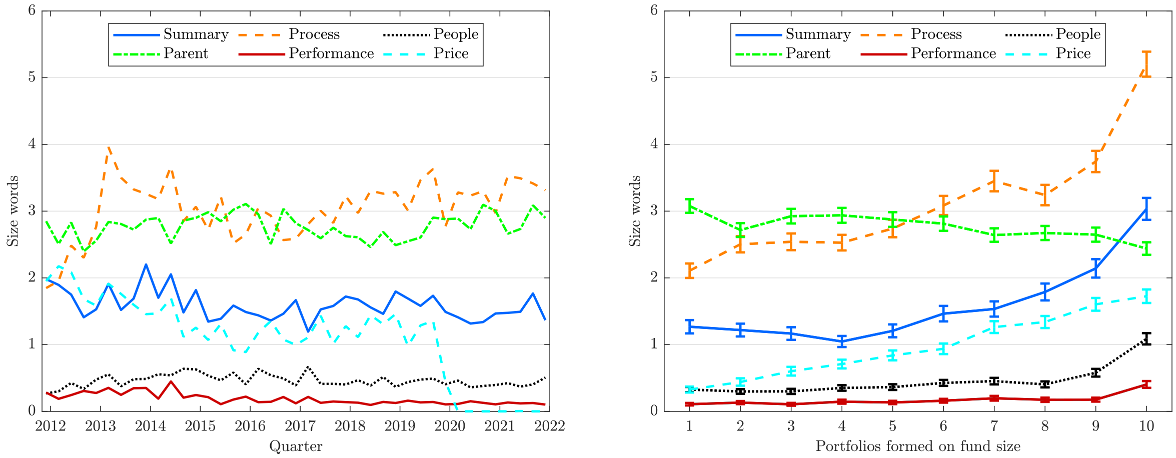
Figure 4: Fund flows and ratings



The figure shows coefficient estimates on Morningstar Star Rating, Analyst Rating, and Quantitative Rating dummy variables in a regression of monthly percentage equity mutual fund flows on the dummy variables, various observables, and fund, year–month, as well as category fixed effects. The coefficient estimates are from specification (4) of Table F1 in the Internet Appendix. The regression omits the three-star, the neutral-analyst, and the neutral Quantitative Rating dummy variables. The bars indicate 90% confidence bands.

Figure 5: 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 (2023) 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.

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

Magnus Dahlquist, Markus Ibert, and Felix Wilke

January 2024

A Data appendix

A.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]

A.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.

A.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.

A.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

A.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.

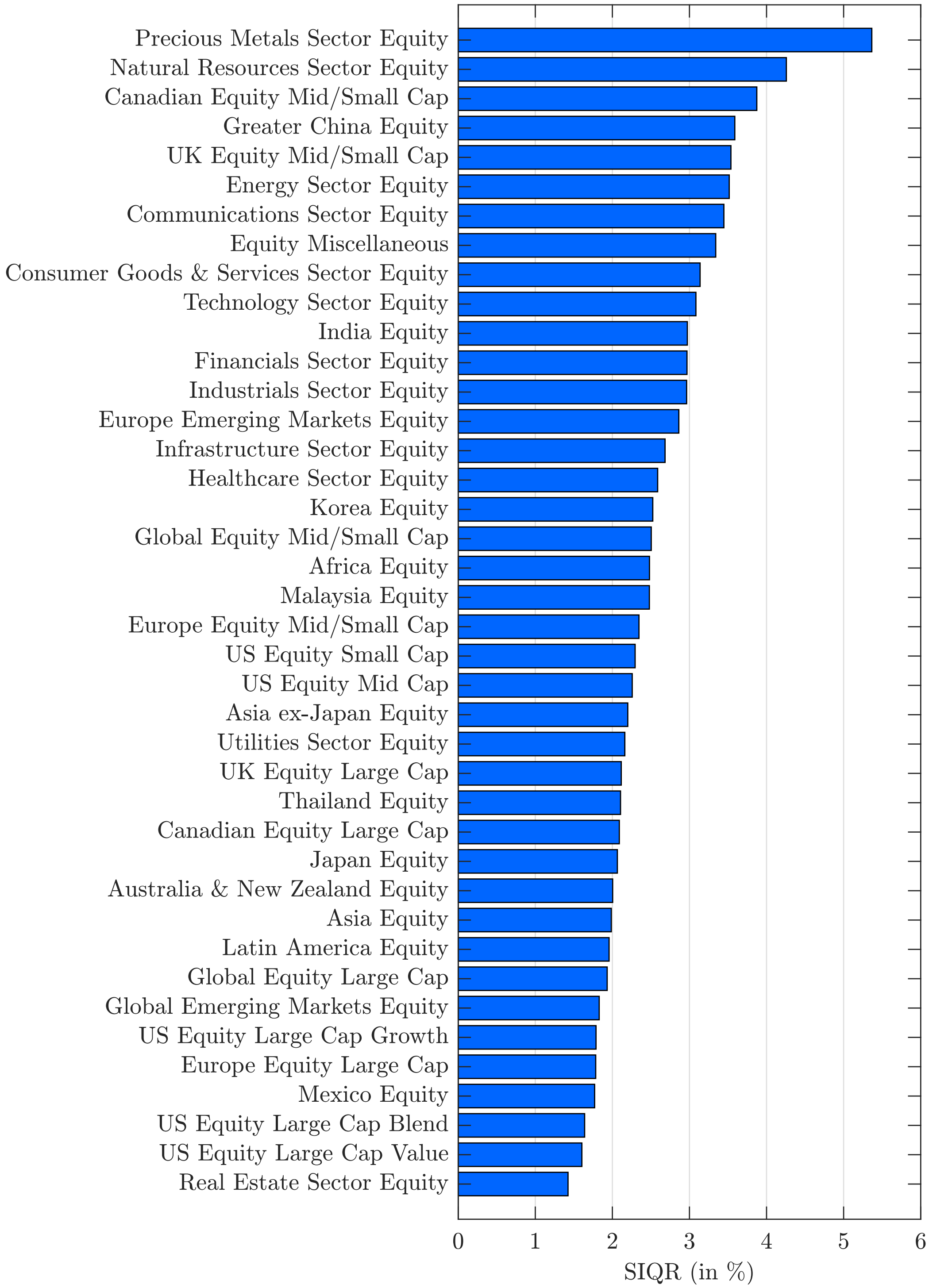
We group Morningstar categories based on the *Global Category* to calculate the SIQR and assign an SIQR to every fund based on its *Morningstar Category* in December 2020. First, we identify the most common *Global Category* among all funds within each *Morningstar Category*. Most funds within a *Morningstar 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 A1 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 A1: 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.

A.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.

A.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.

A.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.

A.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.

A.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,

$$\%\text{AUM}_t = \frac{\text{AUM}_t - \text{AUM}_{t-1}}{\text{AUM}_{t-1}}. \tag{A1}$$

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

$$\text{Reversal}_t = \frac{\text{AUM}_{t+1} - \text{AUM}_t}{\text{AUM}_t - \text{AUM}_{t-1}}. \tag{A2}$$

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

$$\text{abs}(\%\text{AUM}_t) >= 0.5, -0.75 > \text{Reversal}_t > -1.25, \text{ and } \text{AUM}_{t-1} >= 10 \text{ million}, \tag{A3}$$

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:

$$\text{AUM}_k = F \times \text{AUM}_{k-1}(1 + r_k), \text{ for } k \in [t+1, t+n-1], \tag{A4}$$

$$F = \left(\frac{1}{\prod_{k=t+1}^{t+n}(1 + r_k)} \frac{\text{AUM}_{t+n}}{\text{AUM}_t} \right)^{\frac{1}{n}}, \tag{A5}$$

   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.

A.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.

A.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.

A.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.

A.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%).

A.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.

A.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 A.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.

A.5 Rational learner alpha

Using the data from Section A.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 A.4) and the data for the model estimation (Section A.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.

B 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 B1 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

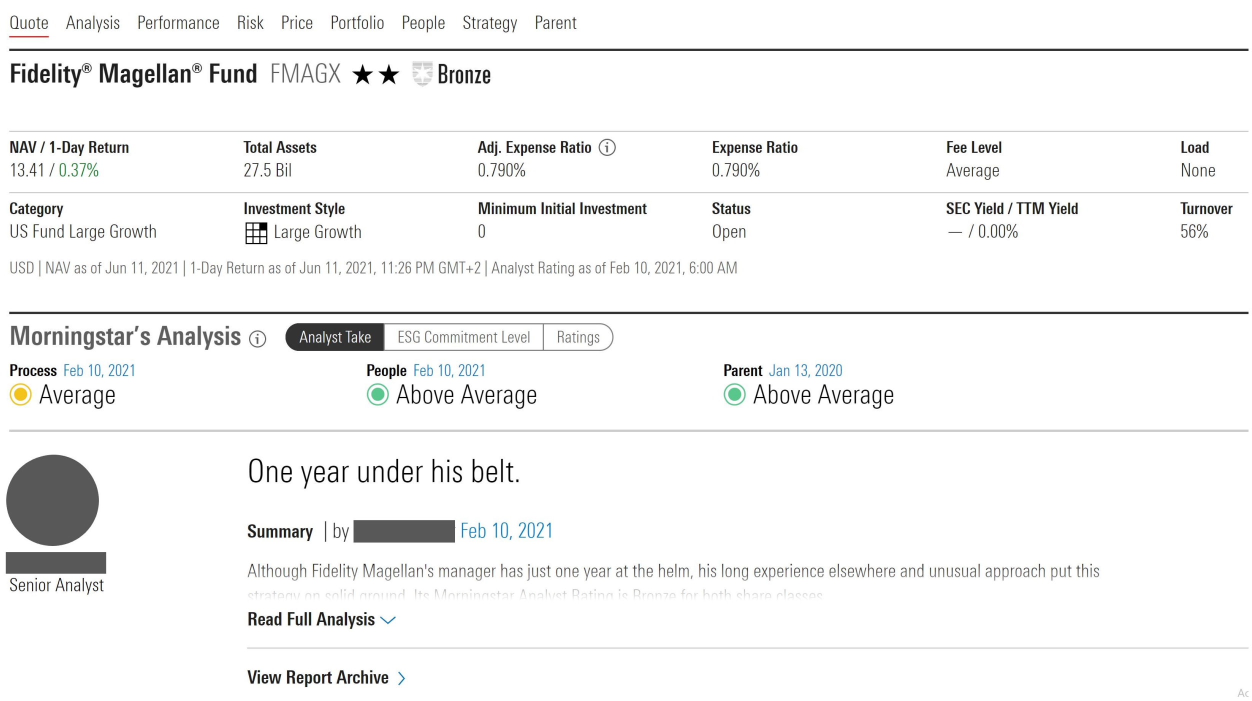
2009. *Manager A* and *manager B* became comanagers in early 2014, a few months before then-portfolio manager *manager D* left the firm. *Manager A* and *manager B* had 10- and eight-years' experience as analysts on the strategy, respectively, before their promotions. In 2014 *the lead manager* hired 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 B1: 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.

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 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 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.

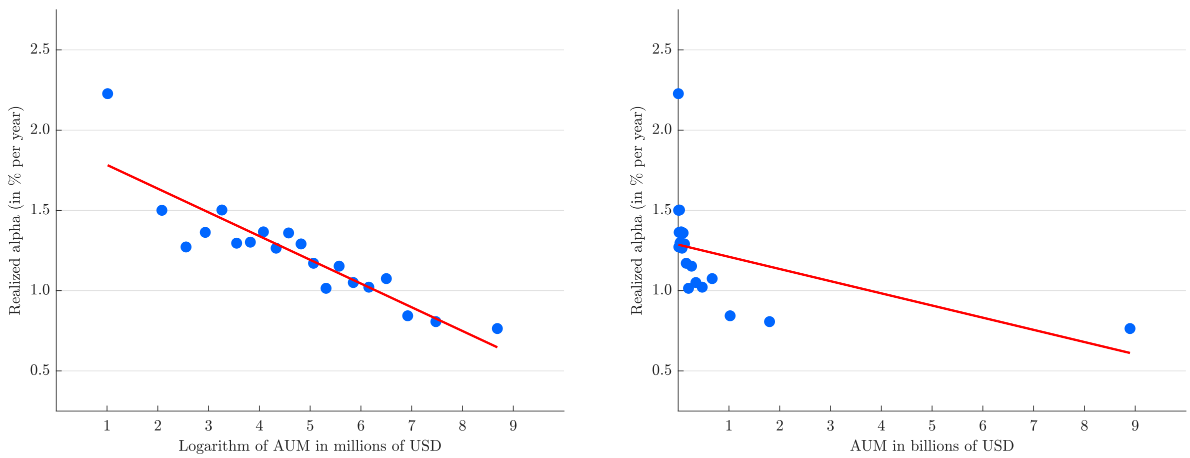
To formalize this argument, we allow for a more flexible impact of AUM on returns in the rational expectations learning model: $c(AUM) = \eta\frac{(AUM)^{\gamma}-1}{\gamma}$, with $\gamma \in (0, 1)$ and γ being an additional parameter to estimate (as in Roussanov, Ruan, and Wei, 2020). If $\gamma = 1$, 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 alphas against lagged fund size

(a) Logarithmic **(b)** Linear



The figure shows realized 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)	(0.025)
	[0.000]	[0.000]
Size (× 100)	0.027**	0.077***
	(0.013)	(0.012)
	[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, 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, fees, and the constant equal the model-predicted parameters of +1, −0.141 (the estimate of η in Table C1), −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 (× 100)	0.065***	0.031**	0.050***	0.020
	(0.017)	(0.015)	(0.013)	(0.015)
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

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

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

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, $-\eta$ on active fund size, and -1 on fees. As before, the coefficient estimates on scale, this time measured as actively managed size, are significantly positive and significantly different from the model-implied coefficient estimate of -0.12.

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 (× 100)	0.116***	0.221***
	(0.029)	(0.021)
	[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, 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, fees, and the constant equal the model-predicted parameters of +1, −0.115 (the estimate of η in Table C4), −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 ($\times$ 100)	0.235^{***}	0.120^{***}	0.239^{***}	0.107^{***}
	(0.032)	(0.033)	(0.025)	(0.025)
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 remains positive and far away from the −0.24 estimate implied by the model. The point estimate on the interaction between turnover and size is negative (−0.038), 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 $\times$ 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 (× 100)	0.117*	0.144***
	(0.067)	(0.031)
	[0.000]	[0.000]
Fees	−0.776***	−1.261***
	(0.241)	(0.123)
	[0.354]	[0.034]
Turnover (× 100)	−0.351	0.059
	(0.703)	(0.152)
	[0.135]	[0.000]
Turnover × Size (× 100)	−0.038	−0.055
	(0.100)	(0.034)
	[0.628]	[0.359]
Constant (× 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), 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, fees, turnover, turnover times size, and the constant equal the model-predicted parameters of +1, −0.240 (the estimate of η in Table C7), −1, 0.706, −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 × 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 (× 100)	0.057*	0.133***
	(0.034)	(0.025)
	[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), 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, fees, small-cap, small-cap times size, and the constant equal the model-predicted parameters of +1, −0.186 (the estimate of η in Table C9), −1, 3.818, −0.431, and 0, respectively.

C.3 Uncertainty about 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):

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

With a true decreasing returns to scale parameter that is constant over time (perceptions thereof will of course vary over time), the state transition equations are:

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

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 maximum likelihood assuming a log functional form for the cost function.

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 0.70. As in Barras et al. (2022), skill, a_0, and scale, η_0, are strongly positively correlated. The correlation is 0.99.

While we cannot estimate a cross-sectional regression when the effect of size on returns varies fund-by-fund, we can still test whether analysts form their expectations according to this model. First, we can test whether the difference between analyst alphas and model-implied alphas is different. They are statistically different (untabulated), but such a rejection is perhaps not too meaningful. Second and more importantly, we can return to our motivating evidence from the main text.

Similar to the main text, Panel (a) of Figure C2 shows that before-fee alphas derived from the model that allows for uncertainty about 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 using 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.

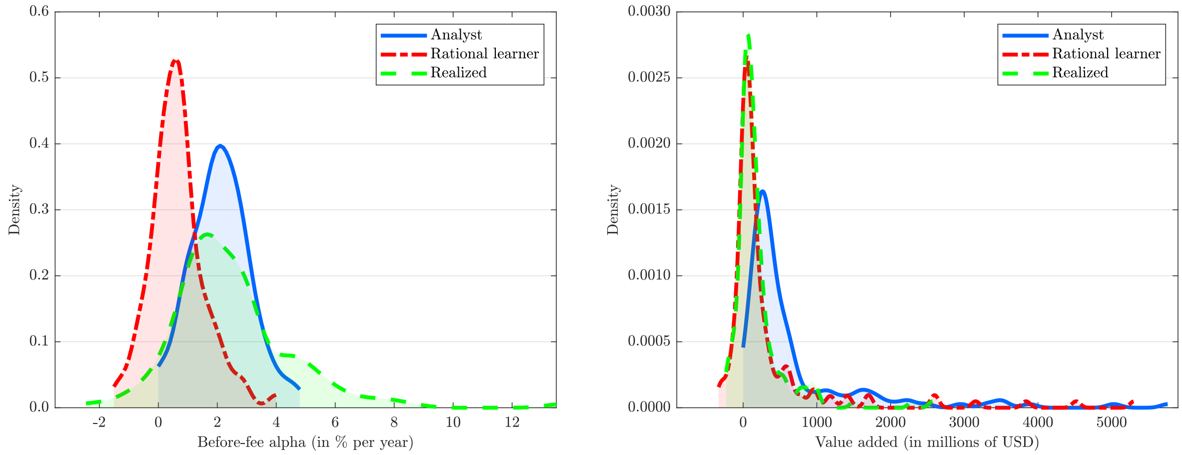
Table C11: Parameter estimates of the rational fund performance model with uncertainty about the decreasing returns to scale parameter

Parameter	Description	Estimate
η_0	Prior mean decreasing returns to scale (DRS) (%)	0.215***
		(0.022)
$\sigma_{\eta,0}$	Prior standard deviation of DRS (%)	0.701***
		(0.138)
a_0	Prior skill mean (%)	2.223***
		(0.076)
$\sigma_{a,0}$	Prior skill standard deviation (%)	5.170***
		(0.274)
σ_ϵ	Residual standard deviation (%)	8.085***
		(0.037)
ρ	Skill persistence	1.000***
		(0.070)
ρ_{η_0,a_0}	Correlation prior skill mean and prior mean DRS	0.991***
		(0.047)

The table shows the parameter estimates of the rational fund performance model with 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.

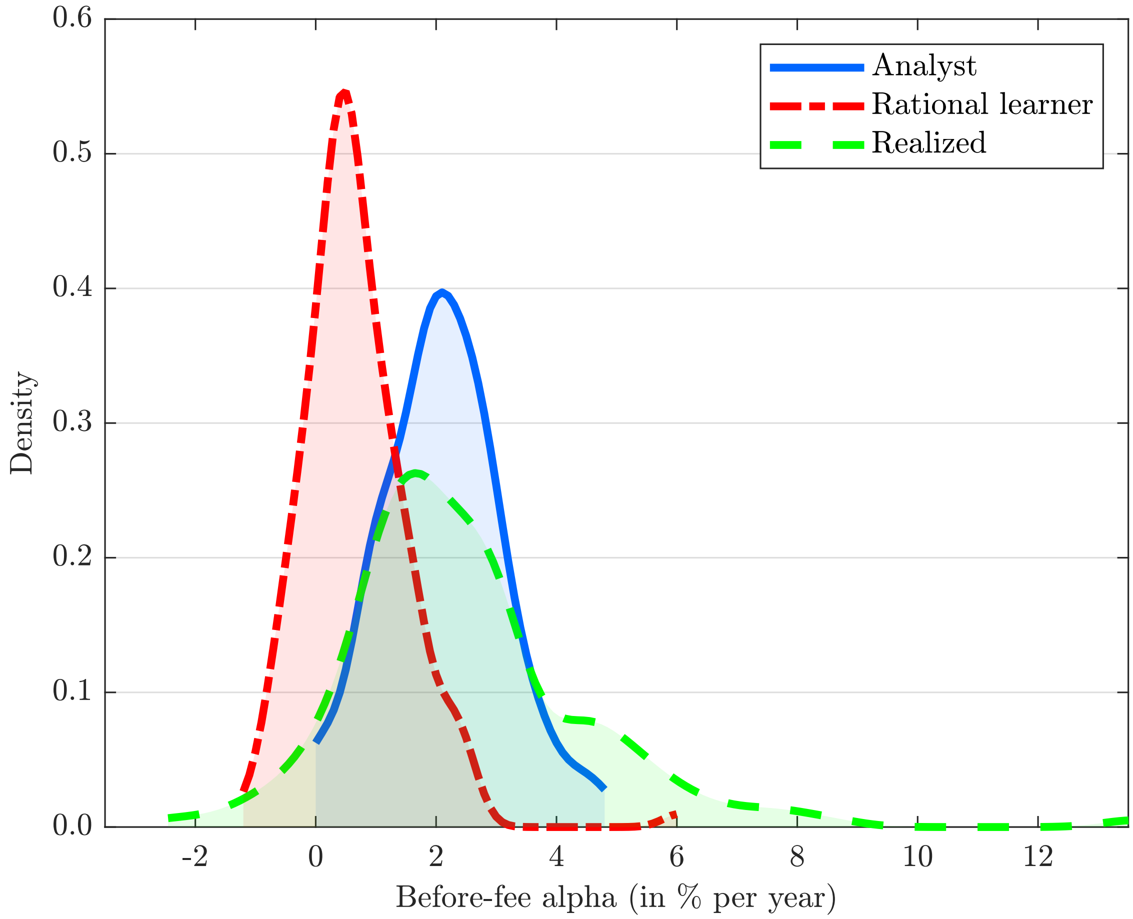
Figure C2: Before-fee alphas and value added of the ten percent largest analyst-rated funds with uncertainty about the decreasing returns to scale parameter

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



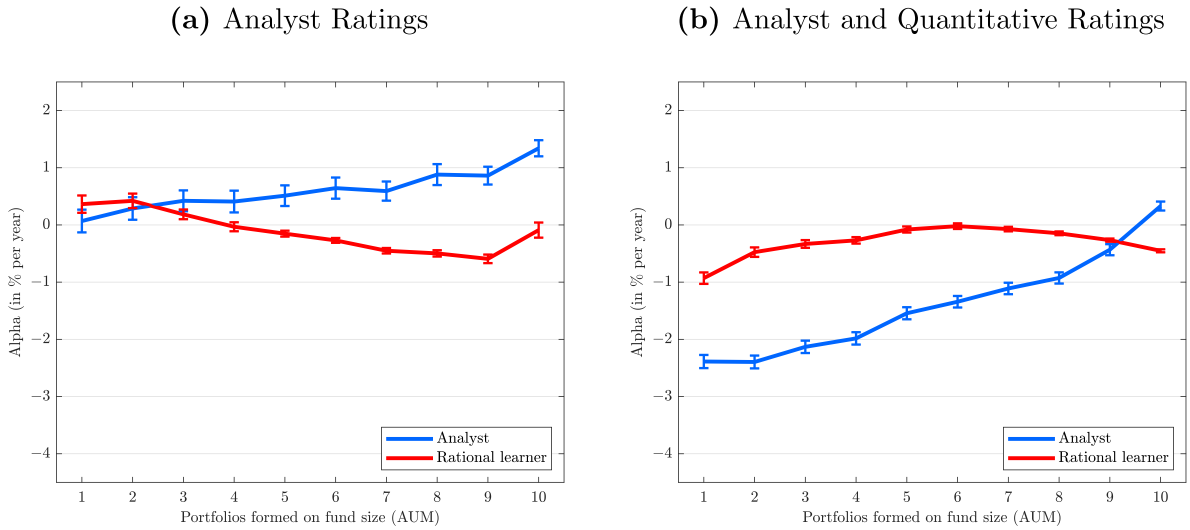
Panel (a) shows the cross-sectional distributions of analyst before-fee alphas (in blue) and before-fee alphas implied by a rational expectations learning model (in red), as well as backward-looking historically realized before-fee alphas (in green), all as of December 2020. Realized alphas are computed 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 uncertainty about 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



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 uncertainty about both managerial skill and the decreasing returns to scale parameter.

Table C12: 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 C12. 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 C13, 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 C13: 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 (× 100)	0.064**	0.134***
	(0.032)	(0.025)
	[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), 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, fees, and the constant equal the model-predicted parameters of +1, −0.248 (the estimate of η in Table C12), −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 C14 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 C14: 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 C15 and C16 show the results. The coefficient estimates on size are significantly positive in all specifications.

Table C15: 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	0.120***	0.183***
	(0.029)	(0.020)
	[0.000]	[0.000]
Size (× 100)	0.134***	0.238***
	(0.028)	(0.025)
	[0.000]	[0.000]
Fees	−0.905***	−1.493***
	(0.157)	(0.063)
	[0.545]	[0.000]
Constant (× 100)	0.288	−0.825***
	(0.270)	(0.176)
	[0.288]	[0.000]
N	1454	13934
Adj. R^2	0.13	0.28

The table shows regressions of Morningstar analyst alphas on skill as perceived by a rational learner, fund size, 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, fees, and the constant equal the model-predicted parameters of +1, the category-specific decreasing returns to scale parameter, −1, and 0, respectively.

Table C16: 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	0.090**	−0.002	0.248***	0.095**
	(0.035)	(0.055)	(0.026)	(0.037)
Size (× 100)	0.215***	0.099***	0.247***	0.095***
	(0.036)	(0.034)	(0.023)	(0.027)
Fees	−1.274***	−0.939***	−1.672***	−0.917***
	(0.144)	(0.113)	(0.199)	(0.211)
People				
Manager tenure		0.113***		0.256***
		(0.040)		(0.033)
Manager ownership		0.112**		0.190***
		(0.054)		(0.042)
Managerial multitasking		0.651***		0.573***
		(0.205)		(0.192)
Management team		0.078		0.487***
		(0.110)		(0.112)
Process				
Top 10 assets (%)		0.130		−0.034
		(0.129)		(0.094)
Tracking error		0.030		−0.077
		(0.057)		(0.101)
Turnover ratio		−0.495***		−0.110
		(0.155)		(0.080)
Retail		−0.287***		−0.156*
		(0.093)		(0.089)
Broker-sold		−0.282**		−0.102
		(0.119)		(0.107)
N	698	650	2830	2626
Adj. R^2	0.24	0.62	0.22	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 C17: Parameter estimates of the rational fund performance model restricted to funds incepted since 2000

Parameter	Description	Estimate
η	Decreasing returns to scale (%)	0.257^{***}
		(0.018)
a_0	Prior mean (%)	2.113^{***}
		(0.082)
$\sigma_{a,0}$	Prior standard deviation (%)	2.431^{***}
		(0.058)
σ_ϵ	Residual standard deviation (%)	8.158^{***}
		(0.021)
ρ	Skill persistence	0.954^{***}
		(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 C17 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 C18: 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 (%)	0.226^{***}
		(0.025)
a_0	Prior mean (%)	2.049^{***}
		(0.118)
$\sigma_{a,0}$	Prior standard deviation (%)	2.376^{***}
		(0.093)
σ_ϵ	Residual standard deviation (%)	8.018^{***}
		(0.034)
ρ	Skill persistence	0.880^{***}
		(0.020)

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 C18 reports parameter estimates of the rational model, which are similar to the ones in Table 4 in the main text.

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 +1, 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 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	2.871	-6.692	19.128***	9.870***
	(4.787)	(6.816)	(2.959)	(3.607)
Size	0.089**	0.117***	0.071***	0.070***
	(0.042)	(0.044)	(0.017)	(0.019)
Fees	-73.220**	-58.525*	-81.549***	-75.916***
	(30.575)	(30.989)	(12.452)	(13.946)
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	5.388	2.919	28.942***	10.849***
	(3.310)	(3.960)	(3.407)	(3.166)
Size	0.125***	0.048**	0.078***	0.039**
	(0.027)	(0.022)	(0.016)	(0.016)
Fees	-18.217***	13.454*	-30.174***	2.702
	(6.797)	(7.433)	(8.685)	(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	11.527***	5.345	52.255***	15.403***
	(3.902)	(5.345)	(4.513)	(3.392)
Size	0.083***	0.048*	0.043***	0.039*
	(0.024)	(0.027)	(0.015)	(0.021)
Fees	-12.210	-4.489	-35.347***	4.461
	(10.739)	(6.676)	(10.528)	(10.005)
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 ($\times$ 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 ($\times$ 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 ($\times$ 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 ($\times$ 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]

$$\text{PillarScore}_i = \gamma_0 + \gamma' X_i + \psi_i, \tag{C7}$$

   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:

$$\text{PillarScore}_j^{predicted} = \hat{\gamma}_0 + \hat{\gamma}' X_j. \tag{C8}$$

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).

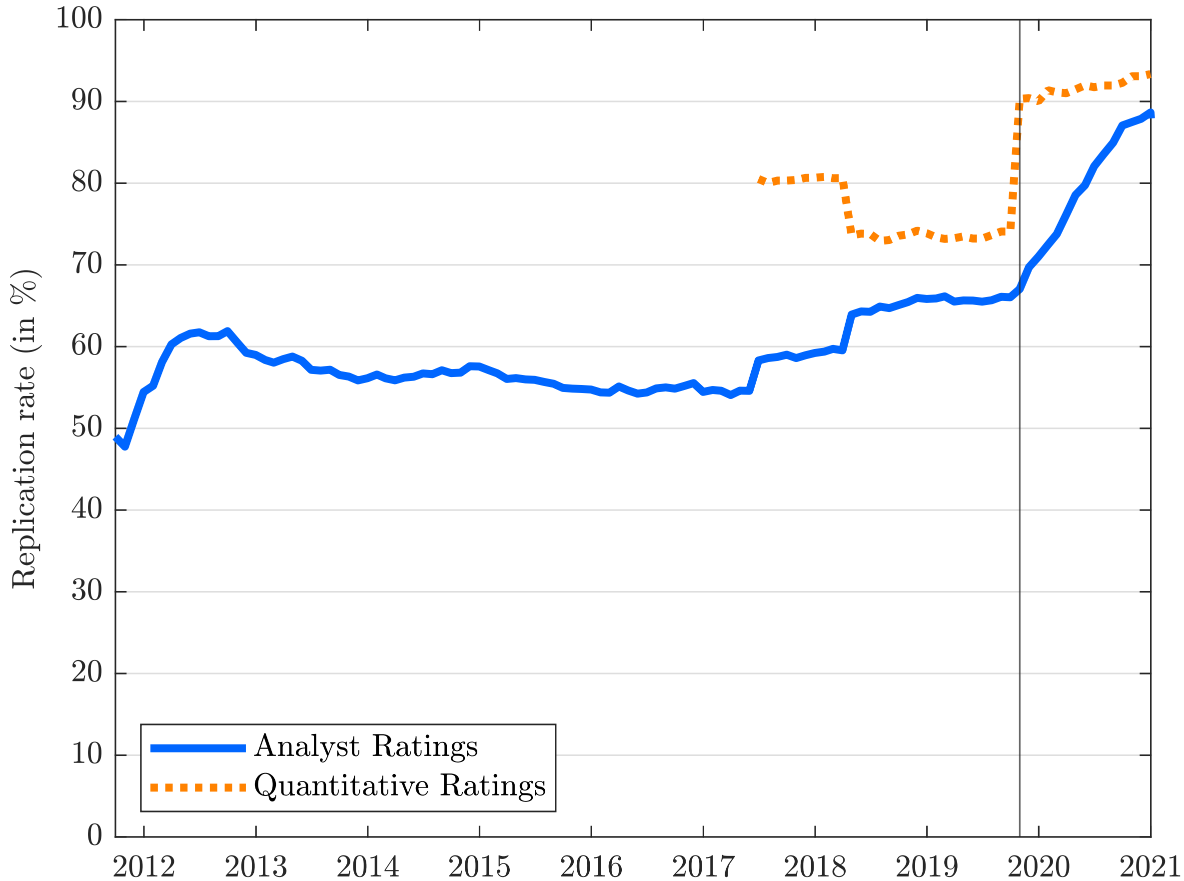
know, the ratings have not been awarded using the new methodology. For this purpose we predict pillar scores according to the new methodology using the coefficients from the above regressions for funds with an old Analyst Rating and employ a similar procedure for funds with a Quantitative Rating.

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 ($\times$ 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.

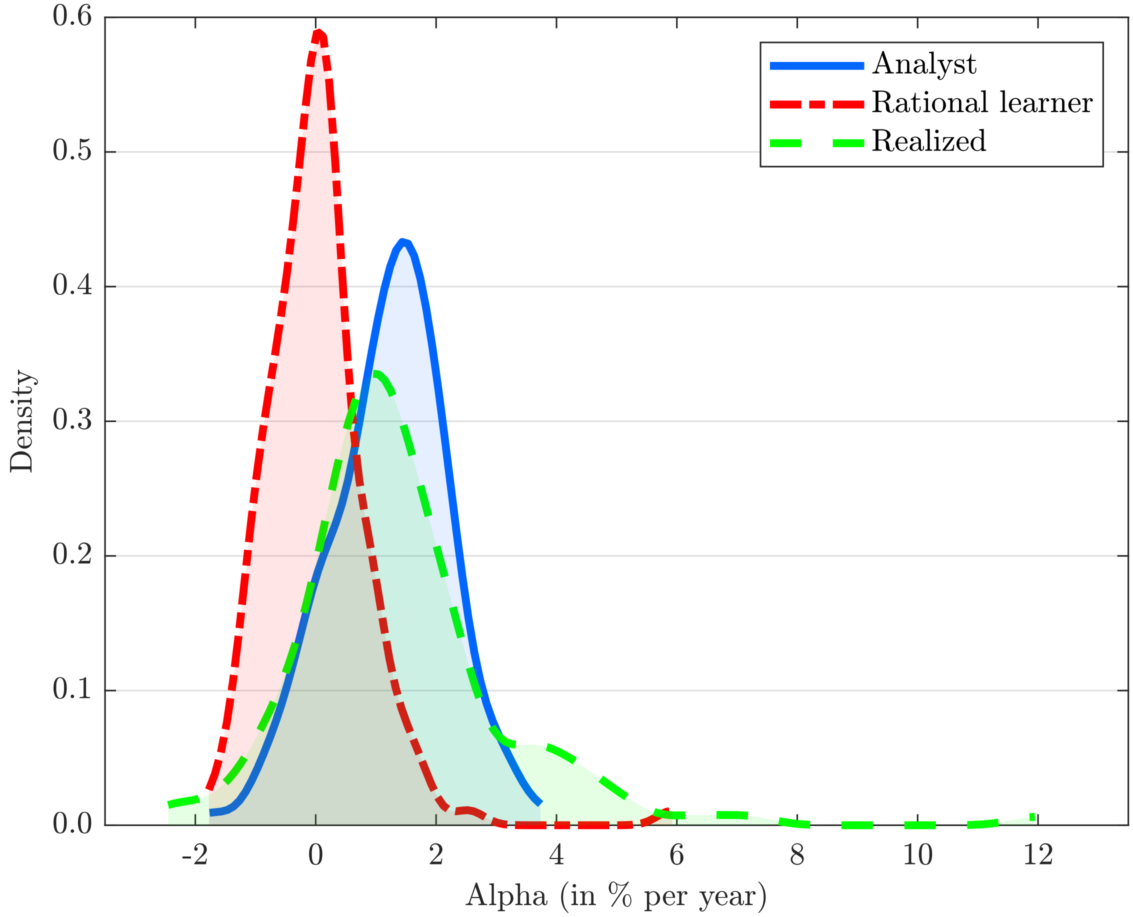
Figure E2 replicates Figure 1 in the paper using the cross-section of funds in December 2021. Tables E3 and E4 show panel regressions using both the cross-section in December 2020 and the one in December 2021. The results are similar to those discussed in the main text.

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	0.370***	0.681***
	(0.067)	(0.037)
Size (× 100)	0.054*	0.115***
	(0.032)	(0.024)
Fees	−1.018***	−1.523***
	(0.141)	(0.055)
Constant (× 100)	0.172	−1.440***
	(0.264)	(0.179)
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), 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.

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 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	0.294***	0.083	0.758***	0.256***
	(0.068)	(0.054)	(0.069)	(0.041)
Size ($\times$ 100)	0.141***	0.095**	0.101***	0.041*
	(0.046)	(0.040)	(0.025)	(0.024)
Fees	−1.423***	−0.970***	−1.872***	−1.150***
	(0.158)	(0.108)	(0.162)	(0.166)
People				
Manager tenure		0.120***		0.228***
		(0.037)		(0.028)
Manager ownership		0.131***		0.184***
		(0.050)		(0.034)
Managerial multitasking		0.617***		0.304***
		(0.169)		(0.103)
Management team		0.069		0.422***
		(0.104)		(0.093)
Process				
Top 10 assets (%)		0.105		−0.024
		(0.091)		(0.067)
Tracking error		−0.012		−0.070
		(0.067)		(0.068)
Turnover ratio		−0.344***		−0.085
		(0.109)		(0.067)
Retail		−0.258***		−0.178***
		(0.077)		(0.068)
Broker-sold		−0.283***		−0.043
		(0.088)		(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,

$$\text{Flow}(\%) = \frac{\text{AUM}_{i,t} - \text{AUM}_{i,t-1} \times (1 + R_{i,t})}{\text{AUM}_{i,t-1} \times (1 + R_{i,t})} \times 100, \tag{C9}$$

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 eight-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 571 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 ($0.355\% \times 12 \times$ USD 409 million) larger inflows per year than does a Neutral-rated fund with the same Star Rating.

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.

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.

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.

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.

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.

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