



## Management Science

Publication details, including instructions for authors and subscription information:  
<http://pubsonline.informs.org>

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

Magnus Dahlquist, Markus Ibert, Felix Wilke

To cite this article:

Magnus Dahlquist, Markus Ibert, Felix Wilke (2025) Are Subjective Expectations Formed as in Rational Expectations Models of Active Management?. *Management Science*

Published online in *Articles in Advance* 21 Nov 2025

. <https://doi.org/10.1287/mnsc.2024.04419>

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. You are free to download this work and share with others, but cannot change in any way or use commercially without permission, and you must attribute this work as “*Management Science*. Copyright © 2025 The Author(s). <https://doi.org/10.1287/mnsc.2024.04419>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.”

Copyright © 2025 The Author(s)

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes. For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

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

Magnus Dahlquist,<sup>a,b,\*</sup> Markus Ibert,<sup>c,d</sup> Felix Wilke<sup>e</sup>

<sup>a</sup>Stockholm School of Economics, SE-113 83 Stockholm, Sweden; <sup>b</sup>Centre for Economic Policy Research (CEPR), London EC1R 5HL, United Kingdom; <sup>c</sup>Department of Finance, Copenhagen Business School, DK-2000 Frederiksberg, Denmark; <sup>d</sup>Danish Finance Institute (DFI), 2000 Frederiksberg, Denmark; <sup>e</sup>Nova School of Business and Economics, Universidade NOVA de Lisboa, 2775-405 Carcavelos, Portugal

\*Corresponding author

Contact: magnus.dahlquist@hhs.se,  <https://orcid.org/0000-0002-1846-3113> (MD); mi.fi@cbs.dk (MI); felix.wilke@novasbe.pt,  <https://orcid.org/0000-0002-6580-825X> (FW)

Received: January 9, 2024

Revised: December 23, 2024; April 19, 2025

Accepted: May 1, 2025

Published Online in Articles in Advance:  
November 21, 2025

<https://doi.org/10.1287/mnsc.2024.04419>

Copyright: © 2025 The Author(s)

**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. Consistently, counterfactual ratings based on the rational model tend to outperform analysts' ratings out of sample. Investor fund flows respond significantly to analyst ratings.

**History:** Accepted by Lukas Schmid, finance.



**Open Access Statement:** This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. You are free to download this work and share with others, but cannot change in any way or use commercially without permission, and you must attribute this work as "Management Science. Copyright © 2025 The Author(s). <https://doi.org/10.1287/mnsc.2024.04419>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by-nc-nd/4.0/>."

**Funding:** Support from the Center for Big Data in Finance [Grant DNR167], the Danish Finance Institute, and the Swedish House of Finance is gratefully acknowledged. This work was funded by Fundação para a Ciência e a Tecnologia (UIDB/00124/2025, UID/PRR/124/2025, Nova School of Business and Economics) and LISBOA2030 (DataLab2030 - LISBOA2030-FEDER-01314200).

**Supplemental Material:** The online appendices and data files are available at <https://doi.org/10.1287/mnsc.2024.04419>.

**Keywords:** alpha • expectations formation • mutual funds

## 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 (Coibion and Gorodnichenko 2012, 2015; Greenwood and Shleifer 2014; Bordalo et al. 2019, 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 (Berk and Green 2004). 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 (Berk and van Binsbergen 2015, Franzoni and Schmalz 2017, Barras et al. 2022, Kim 2022, van Binsbergen et al. 2023), whereas others conclude that investors have limited financial sophistication (Song 2020, Ben-David et al. 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 then 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 (Chen et al. 2004, Pástor et al. 2015, Zhu 2018, Roussanov et al. 2021). We conclude that analysts misjudge returns to scale in active management. Third, such misjudgments have implications for the predictive ability of analysts' forecasts: We find that counterfactual

ratings that incorporate decreasing returns to scale tend to outperform actual analyst ratings.

### 1.1. Analysts' Expectations vs. 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. Although professional analysts could be akin to sophisticated investors, it is unclear how representative analysts' expectations are more generally.

That said, one advantage of the mutual fund setting relative to professional forecasts in other settings is that we can evaluate 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. A natural conjecture is then that investors who follow analysts' ratings would have been better off had they followed forecasts that incorporate decreasing returns to scale instead. Indeed, we provide evidence that they would have been.

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 the marginal investor. This is because expectations that increase with size imply that unlimited amounts of capital should flow into all funds.

### 1.2. 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 (Berk and Green 2004) or close to zero (Pástor and Stambaugh 2012).

### 1.3. 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

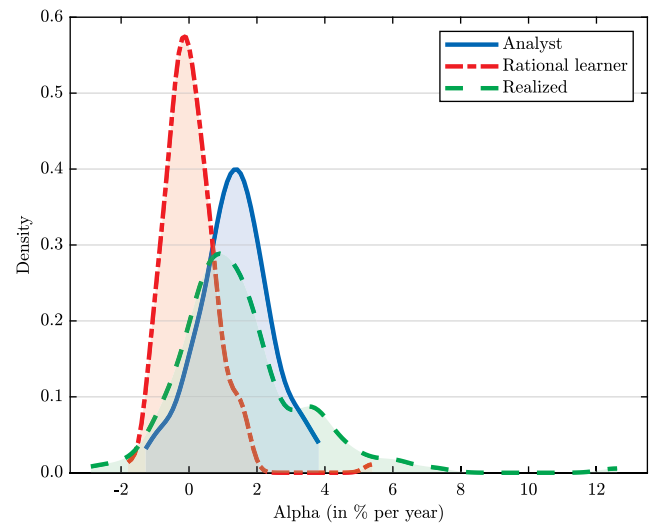
provided a detailed description of how the overhauled ratings are constructed, we can recover detailed measures of expectations since then. Analysts assign the ratings according to a five-tier scale with three positive ratings of Gold, Silver, and Bronze, as well as a Neutral rating and a Negative rating. 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.

### 1.4. Analysts' Expectations vs. Model-Implied Expectations

Figure 1 illustrates our main results. The figure shows analyst alphas, alphas implied by a rational expectations learning model that we estimate and introduce below, and backward-looking historically realized alphas, all for the cross section of the largest 10% 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 compared with analyst alphas: Not all analyst alphas are zero. In fact, analyst alphas are only positive for the largest funds (shown in Figure 1) but

**Figure 1.** (Color online) Alphas of the 10% Largest Analyst-Rated Funds



*Notes.* Cross-sectional distributions of analyst alphas and alphas as implied by a rational expectations learning model, as well as backward-looking historically realized alphas, all as of December 2020. Realized alphas are computed over the lifetime of a fund. The sample is restricted to the 10% 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.

negative for most other funds (not shown in Figure 1). That said, it is well documented that not all realized alphas are zero (Kosowski et al. 2006, Fama and French 2010, Harvey and Liu 2022), and previous research indeed interprets rational expectations models of active management in an approximate sense (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.<sup>1</sup> Consider now the distributions of analyst alphas and historically realized alphas. Figure 1 restricts the sample to the largest funds 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.<sup>2</sup> 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 for the “rational learner” 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 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).<sup>3</sup> 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 10 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 15-basis-point decrease in alpha.

We also test whether counterfactual ratings based on the rational learner’s expectations outperform actual analyst ratings, using an older batch of Morningstar’s analyst ratings that has been available since 2011. Consistent with Armstrong et al. (2019), we find that analyst ratings have some predictive power for future performance. However, counterfactual ratings based on our rational model have even more predictive power. For instance, the difference in alpha between funds that the rational model ranks “Gold” and the ones that Morningstar analysts rank “Gold” is 114 basis points per year.

### 1.5. Robustness and Potential Concerns

Our results are robust to various extensions of the rational expectations learning model (including features from, among others, Pástor and Stambaugh (2012), Berk and van Binsbergen (2015), Pástor et al. (2015), and Barras et al. (2022)), they are robust when we control for additional manager and fund characteristics in reduced form, and 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 et al. 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 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 and has a substantial business reputation at stake, and previous research has used the Morningstar analyst rating as a *benchmark* of independent analysis (Cookson et al. 2021). We believe that analysts’ forecasts are their best attempts to forecast future returns, but ultimately it is

difficult to rule out that analysts' forecasts may be driven by incentive structures or career concerns.

## 1.6. 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.<sup>4</sup> Our contrasting of 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 (Bender et al. 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, a large literature examines the predictions of the Berk and Green (2004) model studying investors' capital allocations and fund returns (Berk and van Binsbergen 2015, Franzoni and Schmalz 2017, Song 2020, Barras et al. 2022, Ben-David et al. 2022, Kim 2022, van Binsbergen et al. 2023). By studying revealed preferences, this literature does not have to worry about whether subjective expectations are "representative." That said, as mentioned above, this literature has reached opposing conclusions. This is not surprising. After all, rational and behavioral theories are observationally equivalent using prices and quantities alone (Cochrane 2017). One way to break this observational equivalence is by studying subjective expectations.

For instance, Berk and Tonks (2007) report that persistence in performance for the worst-performing funds results from investors' unwillingness to respond to bad performance and to withdraw capital. This is evidence against the literal version of Berk and Green (2004), but it is unclear whether Berk and Green (2004) does not hold because investors in the worst-performing funds are irrationally overoptimistic or because they face other frictions that are not part of the model. For instance, investors could have rational expectations but face high information acquisition costs (Mankiw and Reis 2002). Relative to the literature that tests rational expectations models by studying capital allocations and fund returns, the novelty of our paper is to contrast rational expectations models with their key ingredients: actual expectations.

Third, our paper relates to a literature that examines expectations regarding fund performance (Jenkinson et al. 2016, Jones and Martinez 2017, Armstrong et al. 2019, Cookson et al. 2021). The analyst alphas we

recover are an important improvement over previous work, because 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.<sup>5</sup>

Fourth, 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 with their model, the models of Dangl et al. (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 Gârleanu and Pedersen (2018). The models of Gennaioli et al. (2015) and Spiegelger (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. Online Appendix A presents an example of how the analyst rating is displayed on Morningstar's website.

Until 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 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 because alpha measures the performance relative to a passive benchmark. In contrast, the old rating system is not necessarily informative about alpha, because 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 (Morningstar 2021). 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 72 unique analysts covering equity funds in 2020. 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. Online Appendix B contains additional details about our replication and the data.

Under the new rating system, Morningstar’s 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:

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

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 two). 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. Online

Appendix A 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:

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

where  $E_i^s$  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,  $E_i^s[r_{i,j,t+1}]$ . 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 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 Online 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. The sample consists of worldwide active equity mutual funds.<sup>6</sup>

**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	<i>New:</i> Three-pillar ratings (people, process, and parent), SIQR (dispersion of CAPM alphas of fund strategy), and share-class fees <i>Old:</i> Five-pillar ratings (people, process, parent, performance, and price)	<i>New:</i> 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 <i>Old:</i> 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	Five globes Four globes Three globes Two globes One globe
Rating level	Share class	<i>New:</i> Share class <i>Old:</i> 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
Rating peer group	Morningstar category	Morningstar category	Morningstar category	Morningstar global category
Medalist ranking (Gold, Silver, and Bronze) requirement		<i>New:</i> Beat benchmark index <i>and</i> peer group average <i>Old:</i> Beat benchmark index <i>and/or</i> peer group average	<i>New:</i> Beat benchmark index <i>and</i> peer group average <i>Old:</i> Beat benchmark index <i>and/or</i> 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

*Notes.* 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							
Actual rating	Replicated rating					Total	Rate
	Negative	Neutral	Bronze	Silver	Gold		
Negative	80	15	0	0	0	95	84%
Neutral	60	3,035	121	2	0	3,218	94%
Bronze	2	167	2,293	201	10	2,673	86%
Silver	0	1	213	1,731	107	2,052	84%
Gold	0	0	0	88	571	659	87%
Total	142	3,218	2,627	2,022	688	8,697	89%

Panel B: Morningstar quantitative Ratings							
Actual rating	Replicated rating					Total	Rate
	Negative	Neutral	Bronze	Silver	Gold		
Negative	12,557	503	0	0	0	13,060	96%
Neutral	416	26,150	396	1	0	26,963	97%
Bronze	2	906	6,378	376	12	7,674	83%
Silver	0	12	559	4,252	179	5,002	85%
Gold	0	3	1	312	2,328	2,644	88%
Total	12,975	27,574	7,334	4,941	2,519	55,343	93%

*Notes.* 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 3,035 of 3,218 analyst-rated share classes receiving a Morningstar analyst rating of Neutral, yielding a replication rate of 94%).

Panel A shows that for the 8,697 share classes with an analyst rating under the new methodology, Morningstar assigns a Neutral rating to 3,218 share classes. In this case, we assign a Neutral rating in 3,035 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).

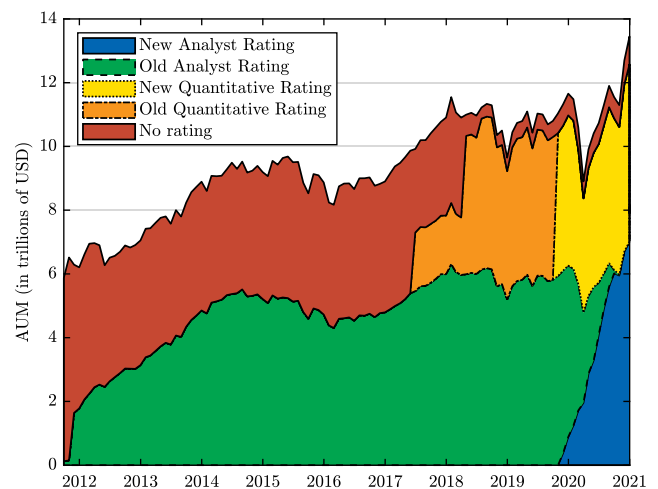
Although we believe that we can replicate Morningstar’s methodology reasonably well to recover analyst alphas, there is measurement error in the dependent variable. Under standard assumptions, measurement error in the dependent variable does not bias coefficient estimates, but inflates standard errors. 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.** (Color online) Size of Active Equity Mutual Fund Industry



*Notes.* 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 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 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

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

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  $c(\text{AUM}_{i,t})$  captures decreasing returns to scale. We refer to  $E_t[r_{i,t+1}]$  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:

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

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 beliefs about managerial skill, that is,  $a_{i,t+1}$  (the only parameter for which there is uncertainty), 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

**Table 3.** Summary Statistics

	N	Mean (V.W.)	Mean (E.W.)	Standard deviation	10%	25%	50%	75%	90%
Panel A: Assets under management									
Analyst rating	1,454		4,760	14,036	154	406	1,248	3,882	10,098
Quantitative rating	12,480		409	1,257	10	30	100	336	931
All ratings	13,934		863	4,871	12	34	126	464	1,477
No rating	4,512		155	1,251	6	13	37	112	291
All	18,446		690	4,290	9	25	89	341	1,144
Panel B: Fees									
Analyst rating	1,454	0.79	1.06	0.39	0.64	0.84	1.00	1.24	1.59
Quantitative rating	12,480	1.11	1.44	0.72	0.65	0.96	1.36	1.82	2.27
All ratings	13,934	0.92	1.40	0.70	0.64	0.94	1.29	1.77	2.23
No rating	4,512	1.28	1.65	0.93	0.78	1.06	1.63	1.97	2.44
All	18,446	0.94	1.46	0.77	0.67	0.95	1.37	1.84	2.27
Panel C: Perceived skill									
Analyst rating	1,454	2.97	2.70	0.90	1.67	2.09	2.58	3.18	3.87
Quantitative rating	12,480	2.50	2.13	0.91	1.14	1.61	2.10	2.55	3.20
All ratings	13,934	2.77	2.19	0.92	1.19	1.65	2.13	2.64	3.29
No rating	4,512	2.76	2.25	1.06	1.26	1.81	2.13	2.45	3.38
All	18,446	2.77	2.20	0.96	1.21	1.68	2.13	2.60	3.31
Panel D: Analyst alphas									
Analyst rating	1,454	1.29	0.60	1.35	−1.09	−0.25	0.69	1.42	2.24
Quantitative rating	12,480	−0.55	−1.62	2.48	−4.82	−3.21	−1.57	0.04	1.50
All ratings	13,934	0.51	−1.39	2.49	−4.67	−2.99	−1.24	0.34	1.66

*Notes.* The table shows value-weighted (V.W., by 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 percentage per year.

each point in time are given by

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

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

where  $\hat{\sigma}_{a,t+1}^2$  describes the uncertainty concerning the perceived skill,  $\hat{a}_{i,t+1}$ , given initial conditions  $a_0$  and  $\sigma_{a,0}^2$ . We assume a logarithmic specification for the decreasing returns to scale; that is,  $c(\text{AUM}) = \eta \log(\text{AUM})$ , where  $\eta$  is a parameter capturing the sensitivity of fund returns to an increase in AUM. We examine a more flexible functional form in Section 8.1 and in Section C.1 of the Online Appendix. The results suggest that the logarithmic specification fits the data well.<sup>7</sup>

We use the expectation maximization (EM) algorithm (Dempster et al. 1977, Watson and Engle 1983) to estimate the model on the fund level (using gross fund returns and fund size).<sup>8</sup> 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  $r_{i,t+1} + f_{i,t+1} = \hat{a}_i + \zeta_{i,t+1}$ , where  $\hat{a}_i$  is the sample average of realized gross abnormal returns.<sup>9</sup> 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.

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, because they focus on U.S.-domiciled

**Table 4.** Parameter Estimates of the Rational Fund Performance Model

Parameter	Description	Estimate
$\eta$	Decreasing returns to scale (%)	0.215*** (0.013)
$a_0$	Prior mean (%)	2.134*** (0.062)
$\sigma_{a,0}$	Prior standard deviation (%)	2.086*** (0.040)
$\sigma_\epsilon$	Residual standard deviation (%)	8.281*** (0.015)
$\rho$	Skill persistence	0.949*** (0.006)

*Notes.* The table shows the parameter estimates of the rational fund performance model in percentage per year. Standard errors are shown in parentheses. The model is estimated using fund-year observations from 1979 to 2020.

\*, \*\*, and \*\*\*Significance levels at 10%, 5%, and 1%, respectively, for the null hypothesis of a zero coefficient.

funds, whereas we also include funds from other domiciles to be consistent with Morningstar’s methodology. The estimated prior mean of managerial skill is 2.13% per year, the prior standard deviation is 2.09%, the residual volatility is 8.28%, and the persistence parameter is 0.95. With a standard deviation of  $\log(\text{AUM})$  of 1.90, the decreasing returns to scale parameter estimate of 0.215% implies that a one-standard-deviation increase in  $\log(\text{AUM})$  leads to a 0.41-percentage-point decrease in returns. Alternatively, a doubling of AUM, corresponding to a log increase of 0.69, leads to a 0.15-percentage-point decrease in returns.

The model laid out thus 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

$$E_t[r_{i,t+1}] = \hat{a}_{i,t} - \eta \log(\text{AUM}_{i,t}) - f_{i,t+1}, \quad (7)$$

which may or may not equal zero. If the rational learner also has rational expectations, the true parameter values of  $a_0$ ,  $\sigma_{a,0}$ ,  $\eta$ ,  $\sigma_\epsilon$ , and  $\rho$ , which are approximated by our estimates, are used to form expectations. We assume rational expectations to form the alphas in December 2020, for every fund according to Equation (7).<sup>10</sup>

## 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 toward 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.

Figure 3(a) shows the results for the sample of funds with an analyst rating, and Figure 3(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. Because funds with a quantitative rating constitute most of the sample in Figure 3(b), the average analyst alphas are significantly lower in Figure 3(b) than in Figure 3(a). The figure also shows that, although 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 therefore 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 perceived skill,  $\text{size} \times \eta$  (where size is measured as the logarithm of AUM, and  $\eta$  is scalability), and fees: The coefficient estimates should be 1,  $-1$ , and  $-1$ ,

respectively.<sup>11</sup> Table 5 presents two cross-sectional regressions: Specification (1) uses the sample of funds with an analyst rating; and 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.

**5.2.1. 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  $+1$  in both specifications.

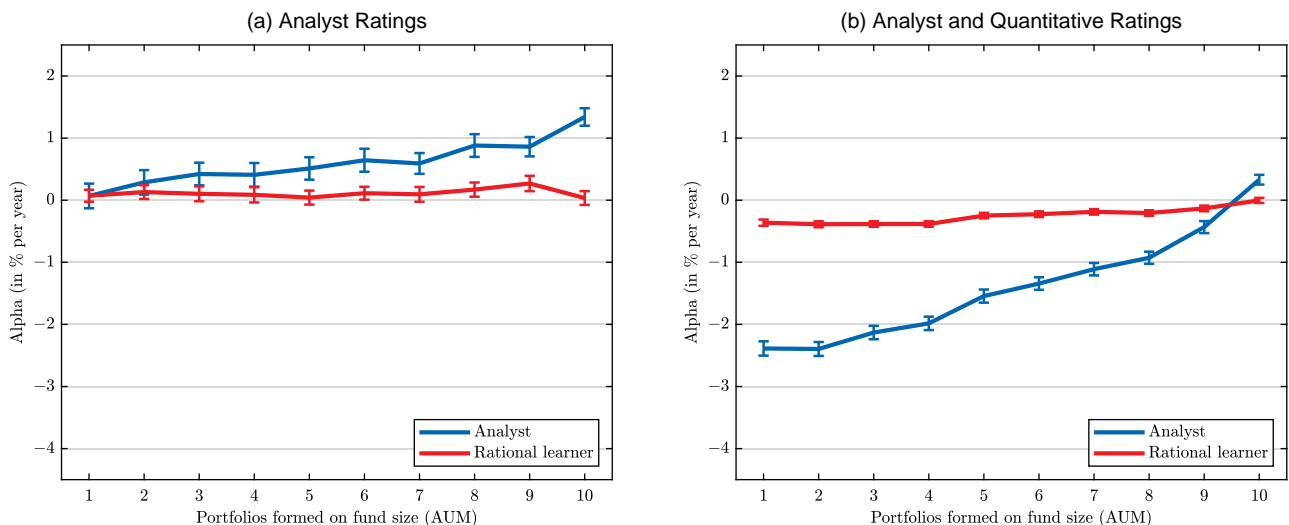
**5.2.2. Fund Size  $\times \eta$ .** The estimate on  $\text{size} \times \eta$  is statistically positive in both columns and has the opposite sign to that of the model's prediction, which leads us to reject the rational expectations model. For instance, in specification (2), the coefficient estimate on size is 0.64 as opposed to  $-1$ .

**5.2.3. 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  $-1$  in specification (1) but is statistically different from  $-1$  in specification (2).

A potential concern is 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 suggests that we do not need to identify the effect of size on analysts' expectations to argue that analysts' expectations are tilted too much

**Figure 3.** (Color online) Alphas Against Fund Size



*Notes.* Alphas against fund size (AUM) as of December 2020 for analysts and for a rational learner. (a) Funds with an analyst rating. (b) 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.

**Table 5.** Cross-Sectional Regressions of Alphas on Fund Characteristics

	Analyst ratings (1)	Analyst and quantitative ratings (2)
<i>Perceived skill</i>	0.395*** (0.067) [0.000]	0.729*** (0.042) [0.000]
<i>Size × η</i>	0.321** (0.149) [0.000]	0.644*** (0.114) [0.000]
<i>Fees</i>	−0.962*** (0.150) [0.799]	−1.536*** (0.059) [0.000]
Constant (×100)	0.062 (0.274) [0.821]	−1.509*** (0.174) [0.000]
<i>N</i>	1,454	13,934
Adjusted <i>R</i> <sup>2</sup>	0.15	0.32

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

\*, \*\*, and \*\*\*Significance levels at 10%, 5%, and 1%, respectively, for the null hypothesis of a zero coefficient.

toward larger funds. As usual, assume that alphas are a linear function of perceived skill and a term accounting for decreasing returns to scale:  $\alpha = \hat{a} - \eta \times \text{size}$ . The figure is consistent with two interpretations. Under a first interpretation, analyst alphas ( $\alpha$ ) for the funds that have grown to be the largest are too large because analysts perceive these funds to be much more skilled than they actually are (a too large  $\hat{a}$ )—while still believing that an increase in size deteriorates future returns. This could happen when analysts start off with a too high prior of  $a$ , perhaps with a high certainty around it, when a fund is born. 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  $\eta$ ). By imposing structure and modeling alphas as a linear function of perceived managerial skill and size, the results of this section support the latter interpretation.

Finally, we add additional variables to our empirical specifications in the next section 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.<sup>12</sup>

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 United States and Sweden (Khorana et al. 2007, Ibert 2023). However, because ownership information is only publicly available for U.S.-domiciled funds, our sample is restricted.<sup>13</sup>

For process, we include a fund’s top 10 assets (the percentage of AUM in the 10 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 Securities and Exchange Commission (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.<sup>14</sup> Top 10 assets and tracking error serve as measures of diversification and activeness, respectively. There is evidence that more active funds outperform (Cremers and Petajisto 2009). In contrast, broker-sold funds and funds held primarily by retail investors have underperformed on average (Bergstresser et al. 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 (Massa 2003, Gaspar et al. 2006, Ferreira et al. 2018).

Because 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

**Table 6.** Cross-Sectional Regressions of Alphas on Additional Fund Characteristics

	Analyst ratings		Analyst and quantitative ratings	
	(1)	(2)	(3)	(4)
Rational learner				
<i>Perceived skill</i>	0.280*** (0.068)	0.117 (0.073)	0.885*** (0.083)	0.354*** (0.060)
Size × $\eta$	0.771*** (0.205)	0.363** (0.175)	0.550*** (0.127)	0.254* (0.136)
Fees	-1.355*** (0.147)	-0.947*** (0.115)	-1.770*** (0.195)	-0.959*** (0.212)
People				
<i>Manager tenure</i>		0.111*** (0.040)		0.247*** (0.033)
<i>Manager ownership</i>		0.115** (0.054)		0.192*** (0.041)
<i>Managerial multitasking</i>		0.645*** (0.205)		0.576*** (0.193)
<i>Management team</i>		0.095 (0.109)		0.496*** (0.114)
Process				
<i>Top 10 assets (%)</i>		0.128 (0.131)		-0.027 (0.091)
<i>Tracking error</i>		-0.014 (0.064)		-0.155* (0.093)
<i>Turnover ratio</i>		-0.485*** (0.156)		-0.108 (0.081)
<i>Retail</i>		-0.290*** (0.092)		-0.156* (0.088)
<i>Broker-sold</i>		-0.266** (0.116)		-0.067 (0.105)
N	698	650	2,830	2,626
Adjusted R <sup>2</sup>	0.26	0.62	0.29	0.64
Sustainability fixed effects	No	Yes	No	Yes
Star rating fixed effects	No	Yes	No	Yes
Morningstar category fixed effects	No	Yes	No	Yes
Fund family fixed effects	No	Yes	No	Yes

*Notes.* 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 10 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 \*\*\*Significance levels at 10%, 5%, and 1%, respectively, for the null hypothesis of a zero coefficient.

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 ×  $\eta$ , and fees unstandardized for comparison with previous tables.

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 ×  $\eta$  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 ×  $\eta$  remain positive in all columns. Most importantly, the point estimates are still far

from the  $-1$  point estimate implied by the rational expectations model. The  $p$ -value for the null hypothesis that the coefficient equals  $-1$  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  $-1$  and not statistically different from  $-1$ , 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 et al. 2018, Ma et al. 2019).

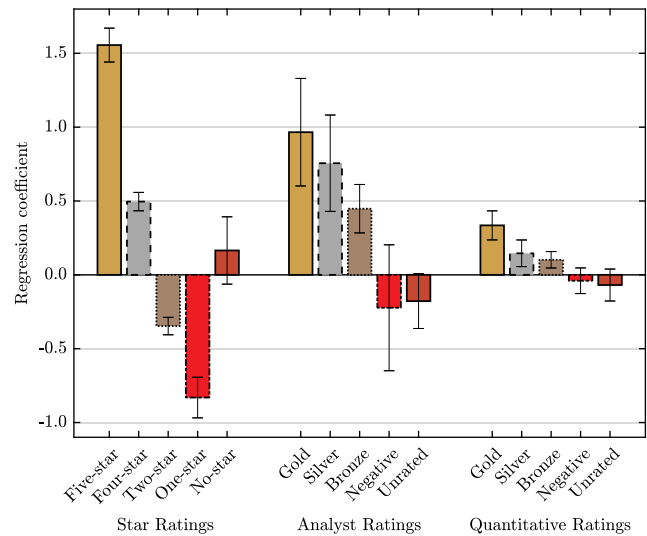
## 6. Analysts' Expectations and Investors' Expectations

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. Indeed, the results of Giglio et al. (2021) suggest that the relation between actions and expectations is economically weak for retail investors.

Although we do not observe investors' subjective expectations, one advantage of working with mutual fund data are that we can test whether better ratings are associated with larger investor fund flows. Online Appendix F shows that they are, 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 Online Appendix F. The figure shows coefficient estimates on star rating dummies, analyst rating

Figure 4. (Color online) Fund Flows and Ratings



Notes. 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 F.1 in the Online Appendix. The regression omits the three-star, the Neutral-analyst, and the Neutral-quantitative rating dummy variables. The bars indicate 90% confidence bands.

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 (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, although statistically significant, the effect of quantitative ratings on flows is considerably smaller.

## 7. Discussion and Summary

### 7.1. Analyst Ratings Fail to Account for Decreasing Returns to Scale

The previous sections robustly show that analyst ratings fail to account for the decreasing returns to scale that we document in actual fund returns. This section aims to further assess this error and, ultimately, assess the implicit value loss for investors who follow analyst ratings.

### 7.2. Quantifying Analysts' Failure to Account for Decreasing Returns to Scale

One way to quantify analysts' failure to account for decreasing returns to scale is to further compare analysts' expectations with those of the rational learner. To set the stage, similar to Figure 3, Table 7 tabulates

**Table 7.** Analyst vs. Rational Learner Expectations and Expected Value Loss for Investors

Size decile	Alphas				Ratings				Expected value loss			
	AUM (in billions of USD)	Rational learner alpha	Analyst alpha	Alpha difference	Rational learner rating	Analyst rating	Rating difference	Flow rating sensitivity	Alpha inference × Flow rating sensitivity	Value loss (in millions of USD)	Aggregate value loss (in billions of USD)	
Panel A: analyst ratings												
1	0.08	0.07	0.07	0.00	2.83	2.81	0.01	3.87	-0.04	-0.04	-0.01	
2	0.22	0.13	0.29	0.16	2.86	3.01	-0.16	4.61	-0.05	-0.12	-0.02	
3	0.42	0.10	0.42	0.32**	2.77	3.01	-0.25**	4.75	-0.05	-0.21	-0.03	
4	0.68	0.09	0.41	0.32**	2.77	3.07	-0.29**	5.01	-0.05	-0.37	-0.05	
5	1.04	0.04	0.51	0.47***	2.71	3.08	-0.37***	4.96	-0.06	-0.62	-0.09	
6	1.56	0.11	0.64	0.53***	2.73	3.16	-0.43***	5.38	-0.07	-1.01	-0.15	
7	2.39	0.09	0.59	0.50***	2.76	3.14	-0.38***	5.38	-0.06	-1.36	-0.20	
8	3.90	0.17	0.88	0.71***	2.88	3.40	-0.52***	6.29	-0.09	-3.30	-0.48	
9	7.10	0.27	0.86	0.59***	3.00	3.37	-0.37***	6.26	-0.07	-4.94	-0.72	
10	30.28	0.04	1.34	-1.31***	2.68	3.77	-1.09***	7.85	-0.13	-47.33	-6.86	
Panel B: Analyst and quantitative ratings												
1	0.01	-0.36	-2.39	2.03***	2.48	1.93	0.54***	0.18	-0.01	-0.00	-0.00	
2	0.02	-0.39	-2.40	2.01***	2.41	1.93	0.48***	0.18	-0.01	-0.00	-0.00	
3	0.03	-0.38	-2.13	1.75***	2.38	1.99	0.39***	0.22	-0.01	-0.00	-0.01	
4	0.06	-0.38	-1.98	1.60***	2.39	2.06	0.33***	0.31	-0.01	-0.01	-0.01	
5	0.10	-0.25	-1.54	1.30***	2.48	2.22	0.26***	0.51	-0.01	-0.01	-0.02	
6	0.16	-0.23	-1.34	1.12***	2.52	2.26	0.26***	0.62	-0.01	-0.02	-0.03	
7	0.27	-0.19	-1.11	0.92***	2.55	2.37	0.18***	0.85	-0.02	-0.04	-0.06	
8	0.47	-0.21	-0.93	0.72***	2.50	2.43	0.07*	1.07	-0.02	-0.08	-0.11	
9	0.97	-0.13	-0.43	0.30***	2.57	2.66	-0.10**	1.65	-0.02	-0.24	-0.34	
10	6.54	-0.00	0.33	-0.33***	2.67	3.08	-0.40***	3.52	-0.05	-6.36	-8.85	

*Notes.* The table compares expectations of analysts with a Bayesian rational learner in December 2020 and shows averages per fund size decile. 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 and expressed in percentage per year. Rational learner ratings are constructed following Morningstar's analyst rating methodology but replace alphas as expected by analysts with rational learner alphas. As flow-rating sensitivities, we consider the case of a rating change from Neutral to the corresponding actual rating of the fund (the coefficients on ratings estimated in monthly data shown in Figure 4 multiplied by 12). We report the statistical significance of mean alpha and rating differences.

\*, \*\*, and \*\*\*Significance levels at 10%, 5%, and 1%, respectively, for the null hypothesis of a zero coefficient.

analyst and the rational learner's alphas by size decile. The table shows that analysts' errors relative to the rational model are quantitatively largest for the smallest (decile 1 in Panel B) and the largest funds (decile 10 in Panel A), as one would expect of forecasts that do not incorporate decreasing returns to scale.

We can also translate the rational learner's alphas into ratings and compare analysts' actual ratings with counterfactual ratings implied by the rational model. To translate the rational learner's alphas into ratings, we follow Morningstar's methodology but simply replace analysts' net-of-fee alphas with the rational learner's net-of-fee alphas. Table 7 also shows the average counterfactual rational learner ratings and analyst ratings by size decile, assuming ratings are expressed on an ordinal scale from one (Negative) to five (Gold). The table shows that the largest analyst-rated funds are, on average, rated Silver (with an average rating of 3.77) and that, in general, analyst ratings strongly increase with size. In contrast, the ratings derived from the rational learner's alpha imply that the largest analyst-rated funds are barely rated Bronze (with an average rating of 2.68).

To further measure analysts' failure to account for decreasing returns to scale, we ask what the weight on a hypothetical size pillar in Morningstar's methodology should be if size were explicitly accounted for. The analysis in Online Appendix G suggests that a size pillar would receive a nonnegligible weight of 43% in a rational methodology that incorporates decreasing returns to scale. The effect of a one-standard-deviation increase in size on expectations of future returns under such a rational methodology would be slightly larger than the effect of a one-standard-deviation increase in fees.

### 7.3. Loss for Investors

Having documented that analysts' failure to incorporate decreasing returns is nonnegligible, we next attempt to assess the implicit expected loss for investors who follow analyst ratings.

For the largest analyst-rated funds (decile 10 in Panel A of Table 7), the rational alpha forecast is 1.31 percentage points lower than the analyst forecast. If the rational learner is correct, an investor following analyst ratings will, in expectation, lose 1.31% in abnormal returns on investments in the largest funds. As we have discussed in Section 6, not all investors follow analysts' ratings and  $-1.31\%$  is thus unlikely to be the aggregate alpha loss to investors. To get a crude measure of the aggregate loss to investors, we multiply the difference in analyst alpha and rational learner alpha by the flow-rating sensitivity. Intuitively, if no investors followed analysts, the flow-rating sensitivity would be zero and the expected abnormal return loss to investors in aggregate would be zero, too.<sup>15</sup> Our baseline flow-rating sensitivity for each fund is the sensitivity of a rating change from Neutral to the actual rating of the fund as shown

in Figure 4, and we annualize these estimates by multiplying them with 12. This sensitivity then captures the incremental increase in flows solely due to analysts rating the fund the way they do.

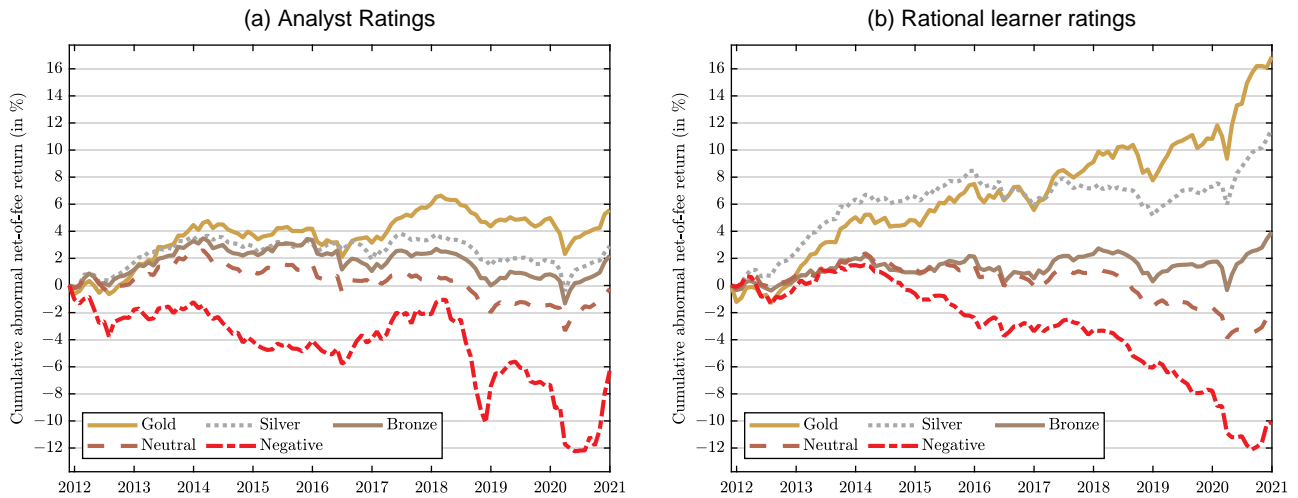
For the largest funds, investors' annualized fund flows are on average 7.85% larger than if these funds were counterfactually rated with Neutral. Table 7 also shows the aforementioned product of the flow-rating sensitivity and the alpha difference. For instance, investors lose 13 basis points per year on average when investing in the largest funds. This is an economically meaningful loss as it occurs for the largest funds. When we compute a value loss measure akin to the value added measure of Berk and van Binsbergen (2015), the value loss is \$47 million per year for the average fund in the largest decile of analyst-rated funds. The last column of Table 7 aggregates these value losses across funds within size deciles and shows that, in total, investors are expected to lose around \$9 billion per year by following analysts' recommendations relative to the rational model, with most of the loss concentrated in the largest funds.

Finally, we compare how the counterfactual rational learner ratings would have performed out of sample relative to the actual Morningstar analyst ratings. To get a meaningfully long sample period, for this analysis we also use the ordinal analyst ratings that are available before 2019 (before Morningstar's methodology change). As mentioned above, to construct counterfactual rational ratings we simply replace Morningstar's net-of-fee alpha estimate with the rational learner's net-of-fee alpha estimate. We then sort the rational learner's net alpha estimates into five different bins within a Morningstar category in a given month as in Morningstar's new methodology.<sup>16</sup> Then, we restrict the set of funds to the ones that have a Morningstar analyst rating at a given point in time in order to compare the same sets of funds. Finally, in each month we construct equal-weighted portfolios of analyst-rated funds by analyst rating and by rational learner rating.

Figure 5 shows the results and plots cumulative out-of-sample abnormal returns over time by rating portfolio. Figure 5(a) shows that analyst ratings had some predictive power for mutual fund returns from 2011 to 2020, consistent with the results in Armstrong et al. (2019). Figure 5(b) shows that portfolios formed based on the counterfactual rational learner ratings would have performed better. The difference in performance is particularly apparent for Gold-rated funds, which is expected as Morningstar analysts rate many of the largest funds with Gold. These funds tend to be impacted the most by decreasing returns to scale. In contrast, the rational learner realizes that the largest funds tend to be impacted the most by decreasing returns to scale and rates smaller funds with Gold (Table 7), which then perform better out of sample.

Table 8 tabulates the results and provides  $p$ -values for the null hypotheses that there is no return difference

**Figure 5.** (Color online) Performance of Portfolios Sorted on Ratings



*Notes.* Cumulative abnormal net-of-fee returns of portfolios sorted on ratings using monthly data from November 2011 to December 2020. The sample includes funds with an analyst rating. (a) Sorts funds on Morningstar’s analyst ratings. (b) Sorts on our counterfactual rational learner ratings. Every month, we sort funds into portfolios based on their ratings and record the average portfolio return in the subsequent month. Abnormal returns in a given month for a given fund are computed by subtracting a benchmark return (computed relative to the fund’s Morningstar category benchmark in expanding window regressions) from the fund return.

between the portfolio of analyst-rated funds and the portfolio of rational-rated funds for a given rating. Consistent with Figure 5, the table shows that the return

difference between Gold-rated funds as identified by the rational learner and Gold-rated funds as identified by analysts is statistically significant with a  $p$ -value of

**Table 8.** Portfolio Sorts and Out-of-Sample Performance

	Sorting variable: analyst rating		Sorting variable: Rational learner rating		Alpha difference $p$ -value
	Alpha	Standard error	Alpha	Standard error	
Panel A: analyst ratings					
Gold	0.60	0.47	1.74***	0.64	[0.02]
Silver	0.32	0.39	1.21**	0.49	[0.00]
Bronze	0.25	0.42	0.43	0.40	[0.17]
Neutral	-0.02	0.39	-0.23	0.38	[0.22]
Negative	-0.67	1.16	-1.16**	0.57	[0.56]
Gold–Silver	0.28	0.25	0.52	0.33	
Gold–Bronze	0.36	0.22	1.30***	0.37	
Gold–Neutral	0.63***	0.23	1.97***	0.53	
Gold–Negative	1.28	0.89	2.89***	0.75	
Panel B: Analyst and Quantitative Ratings					
Gold	0.74	0.50	1.40**	0.63	[0.03]
Silver	0.37	0.42	0.97**	0.50	[0.01]
Bronze	0.15	0.43	0.19	0.41	[0.71]
Neutral	-0.27	0.42	-0.33	0.40	[0.53]
Negative	-1.04*	0.57	-1.16**	0.52	[0.74]
Gold–Silver	0.37	0.23	0.42	0.26	
Gold–Bronze	0.59***	0.19	1.21***	0.32	
Gold–Neutral	1.01***	0.20	1.73***	0.40	
Gold–Negative	1.79***	0.36	2.56***	0.50	

*Notes.* The table presents results of portfolio sorts with an out-of-sample performance evaluation using monthly data from November 2011 to December 2020. The statistics shown are annualized average abnormal net-of-fee returns (in percentage) of portfolios formed on lagged ratings, standard errors of the mean returns, and in brackets,  $p$ -values from a test for the null hypothesis of a zero difference between the average abnormal return of a rating portfolio formed on analyst ratings and the portfolio of the same rating category formed on rational learner ratings. Panel A includes funds with an analyst rating. Panel B includes funds with an analyst rating or a quantitative rating. Both panels report results for sorting funds on Morningstar’s analyst ratings and on our counterfactual rational learner ratings, respectively. Abnormal returns in a given month for a given fund are computed by subtracting a benchmark return (computed in expanding window regressions) from the fund return. Standard errors are as in Newey and West (1987), allowing for serial correlation up to three lags.

\*\*\*, \*\*, and \*Significance at 1%, 5%, and 10% levels, respectively.

0.02. On average, Gold-rated funds as identified by the rational model outperform Gold-rated funds as identified by analysts by 1.14% per year. The return differences for the remaining ratings all have the expected sign (i.e., they are positive for funds rated better than Neutral and negative for funds rated worse than Neutral), but only the returns of the Silver-rated funds as identified by the rational model are statistically different from the corresponding returns of the Silver-rated funds as identified by analysts. For completeness, Panel B of Table 8 reports the results for the set of funds with an analyst or a quantitative rating. However, quantitative ratings are only available since 2017.

On a final note, the out-of-sample predictability of fund returns is at odds with models of competitive capital provision (Berk and Green 2004). In such models, no variable whatsoever predicts fund returns in real time. The fact that, of all variables, incorporating fund size improves return prediction in Figure 5 is consistent with the notion that not only analysts misjudge returns to scale in active management, but investors do so too. Had investors instead correctly judged the effect of size on returns, adding size as a predictor should not lead to an improvement in return prediction.

In sum, although we do not provide counterfactuals through the lens of a structural model, the results suggest that investors would have been better off had they followed ratings that incorporate decreasing returns to scale as opposed to actual analyst ratings.

## 8. Additional Issues

### 8.1. Robustness

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

Section C.5 in the Online Appendix shows that 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 (Roussanov et al. 2020). Section C.1 in the Online Appendix suggests 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 Pollet and Wilson (2008), Pástor et al. (2015), and Busse et al. (2021), 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 (see Section C.2 in the Online Appendix). We also

extend the baseline model to account for uncertainty in the decreasing returns to scale parameter, which can vary over time, and industry size (Pástor and Stambaugh 2012). In the former case, the effect of size on returns varies fund-by-fund (Barras et al. 2022), just like managerial skill. Sections C.3 and C.4 in the Online Appendix show that the results are robust. Section C.6 in the Online Appendix estimates the rational model by global category and shows that the results are robust. 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 inception since 2000, which is the first year of data that enters Morningstar's methodology through the SIQR computation. Again, the results are robust (see Section C.7 in the Online Appendix). Section C.9 in the Online Appendix shows that the results are robust to controlling for alternative measures of perceived skill based on a diagnostic Kalman filter (Bordalo et al. 2019) and to controlling for several lags of past returns.

Our regressions of expectations on fund characteristics identify the coefficient estimate on size using cross-sectional variation. Section E.3 in the Online 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 (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.<sup>17</sup> We also document robust evidence of decreasing returns to scale in realized returns using the fund fixed effects recursive demeaning estimator of Zhu (2018) in Section E.1 of the Online Appendix. 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. Section E.4 in the Online Appendix shows that the results are robust.

### 8.2. 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.

Although we cannot rule out misguided incentives, 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 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 (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.

### 8.3. Size in Morningstar's Methodology

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 (e.g., Morningstar could have told its analysts to ignore fund size in constructing the ratings). 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.

We believe that it is unlikely that the analyst ratings do not account for size by design. First, Morningstar's methodology document outlining the construction of analyst ratings suggests that fund size could be a factor in determining the process pillar score as the evaluation of a fund's process seeks to understand "the fit of the process with the resources backing the strategy and with the size of the asset base tied to the strategy" (Morningstar 2021, p. 8). Second, we have performed a textual analysis of more than 20,000 reports and notes that analysts wrote to accompany the ratings (following the analysis in Wilke (2025)). Online Appendix G reports the results and suggests that, even though there is no explicit pillar rating for fund size, analysts do discuss fund size in the written reports and notes. 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.

Ultimately, we cannot look inside analysts' minds. One possibility is that humans are inherently drawn to predict future returns with past returns. Of course, such beliefs are misguided when there are decreasing returns to scale. Drawing an analogy with equities again, such beliefs are similarly misguided when equity returns are predictable, as they are (Cochrane 2011). For mutual funds, high past returns tend to imply large size, which implies lower returns going forward. For equities, high past returns tend to imply high prices relative to fundamentals, which implies lower returns

going forward, too. In both cases, even professionals do not seem to recognize that returns going forward are lower (for evidence on equities, see Greenwood and Shleifer (2014), who, among other forecasts, study chief financial officer's equity return forecasts).

## 9. Conclusion

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

Given that investors follow analysts' recommendations, our results have implications for investors. Specifically, investors would have been better off by following ratings as implied by the workhorse model—which incorporates decreasing returns to scale—as opposed to following actual analyst ratings—which do not incorporate decreasing returns to scale.

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 (Barberis et al. 2015, Adam et al. 2017, Nagel and Xu 2022). The models of active management of Gennaioli et al. (2015) and Spiegel (2020) allow for deviations from rational expectations. Huang et al. (2007) allow for heterogeneity among investors. These models 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.

## Acknowledgments

The authors 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, École Supérieure des Sciences Economiques (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 comments and suggestions. The paper was previously circulated under the title “Expectations of Active Mutual Fund Performance.”

## Endnotes

<sup>1</sup> Fama and French (2010, p. 1933) 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).”

<sup>2</sup> 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 one would expect a fund’s alpha going forward to be bounded by a reasonable prior, say zero, and the historically realized alpha—unless one believes increases in size actually increase future returns.

<sup>3</sup> In fact, estimating the model corresponds to the definition of “rational expectations” in this literature (Berk and Green 2004, p. 1274). 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.

<sup>4</sup> 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 McLemore (2019), Pástor et al. (2020, 2022), Roussanov et al. (2020), Dyakov et al. (2020), Busse et al. (2021), Reuter and Zitzewitz (2021), and Harvey et al. (2021).

<sup>5</sup> Armstrong et al. (2019) examine the ability of analyst ratings to predict fund performance from 2011 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, in particular for the worst-performing funds.

<sup>6</sup> Morningstar also operates a similar methodology for passive mutual funds and ETFs. However, this methodology is separate from the one for actively managed funds and, thus, consistent with Morningstar’s methodology, we exclude passive funds and ETFs from the sample.

<sup>7</sup> In the most general version of our model with indexing in Section C.2 of the Online 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).

<sup>8</sup> 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.

<sup>9</sup> One concern is that this procedure could create a bias toward finding decreasing returns to scale similar to the bias that troubles

finite-sample fixed effects regressions (Pástor et al. 2015; note that  $\hat{\alpha}_i$  is a fund fixed effect that is computed using information over the entire life of a fund). In Section C.8 of the Online Appendix, we alternatively estimate  $\hat{\alpha}_i$  using three-year rolling window averages, which eliminates this potential bias. The results are similar.

<sup>10</sup> We do not model the possibility that the rational learner could send a signal about the quality of the fund to other investors (as analysts can). If the rational learner could send a signal, the rational learner would take into account that any recommendation could affect fund flows, and hence the fund’s size and in turn the alpha that the rational learner signals.

<sup>11</sup> 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  $-1$ .

<sup>12</sup> 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.

<sup>13</sup> 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. Following 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.

<sup>14</sup> We winsorize fund turnover at the 1st and 99th percentiles as in Pástor et al. (2017) and do the same with the top 10 assets. The retail dummy takes the value of one if more than two-thirds 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 two-thirds 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%.

<sup>15</sup> We compute the alpha loss using comparative statics. We note that to compute the loss of investors in a structural sense, one needs to know investors’ utility function and how they allocate capital across funds. For instance, our computations neglect that a stronger-flow rating sensitivity will lead to increases in size that will change the initial alpha difference. That said, for the smaller increases in size that we consider, the effect of decreasing returns to scale is likely to be negligible, consistent with the results in Reuter and Zitzewitz (2021). Moreover, the effect of flows on size for the rational forecast would further increase the alpha difference relative to analysts, so our computations in that sense provide a lower bound.

<sup>16</sup> Specifically, conditional on a positive rational learner net alpha within a particular Morningstar Category, the top 15% of funds receive a Gold rating, and so on, see page 12. There are two slight look-ahead biases in this exercise. First, the size of the bins (e.g., 15% of funds in the top bin) is only known with Morningstar’s new methodology, but not with Morningstar’s old methodology. Thus, the size of these bins was technically not known in, say, 2011. Second, to be consistent with our main analysis, we form rational learner alphas using the parameter estimates of Table 4. These are estimated using the entire sample until 2020. However, the parameter estimates are largely unchanged when we estimate the model with data stopping in 2011.

<sup>17</sup> For a similar reason, the predictability of forecast errors would not be powerful evidence against rational expectations models of

active management. To see this, consider the investors in the model of 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.

## References

- Adam K, Marcet A, Beutel J (2017) Stock price booms and expected capital gains. *Amer. Econom. Rev.* 107(8):2352–2408.
- Armstrong WJ, Genc E, Verbeek M (2019) Going for gold: An analysis of Morningstar analyst ratings. *Management Sci.* 65(5):2310–2327.
- Barberis N, Greenwood R, Jin L, Shleifer A (2015) X-CAPM: An extrapolative capital asset pricing model. *J. Financial Econom.* 115(1):1–24.
- Barras L, Gagliardini P, Scaillet O (2022) Skill, scale, and value creation in the mutual fund industry. *J. Finance* 77(1):601–638.
- Ben-David I, Li J, Rossi A, Song Y (2022) What do mutual fund investors really care about? *Rev. Financial Stud.* 35(4):1723–1774.
- Bender S, Choi JJ, Dyson D, Robertson AZ (2022) Millionaires speak: What drives their personal investment decisions? *J. Financial Econom.* 146(1):305–330.
- Bergstresser D, Chalmers JMR, Tufano P (2009) Assessing the costs and benefits of brokers in the mutual fund industry. *Rev. Financial Stud.* 22(10):4129–4156.
- Berk JB, Green RC (2004) Mutual fund flows and performance in rational markets. *J. Political Econom.* 112(6):1269–1295.
- Berk JB, Tonks I (2007) Return persistence and fund flows in the worst performing mutual funds. NBER Working Paper No. 13042, National Bureau of Economic Research, Cambridge, MA.
- Berk JB, van Binsbergen JH (2015) Measuring skill in the mutual fund industry. *J. Financial Econom.* 118(1):1–20.
- Berk JB, van Binsbergen JH (2017) Mutual funds in equilibrium. *Ann. Rev. Financial Econom.* 9:147–167.
- Blake CR, Morey MR (2000) Morningstar ratings and mutual fund performance. *J. Financial Quant. Anal.* 35(3):451–483.
- Bordalo P, Gennaioli N, La Porta R, Shleifer A (2019) Diagnostic expectations and stock returns. *J. Finance* 74(6):2839–2874.
- Bordalo P, Gennaioli N, Ma Y, Shleifer A (2020) Overreaction in macroeconomic expectations. *Amer. Econom. Rev.* 110(9):2748–2782.
- Busse JA, Chordia T, Jiang L, Tang Y (2021) Transaction costs, portfolio characteristics, and mutual fund performance. *Management Sci.* 67(2):1227–1248.
- Carhart MM (1997) On persistence in mutual fund performance. *J. Finance* 52(1):57–82.
- Chen J, Hong H, Huang M, Kubik JD (2004) Does fund size erode mutual fund performance? The role of liquidity and organization. *Amer. Econom. Rev.* 94(5):1276–1302.
- Choi JJ, Robertson AZ (2020) What matters to individual investors? Evidence from the horse's mouth. *J. Finance* 75(4):1965–2020.
- Cochrane JH (2011) Presidential address: Discount rates. *J. Finance* 66(4):1047–1108.
- Cochrane JH (2017) Macro-finance. *Rev. Finance* 21(3):945–985.
- Coibion O, Gorodnichenko Y (2012) What can survey forecasts tell us about information rigidities? *J. Political Econom.* 120(1):116–159.
- Coibion O, Gorodnichenko Y (2015) Information rigidity and the expectations formation process: A simple framework and new facts. *Amer. Econom. Rev.* 105(8):2644–2678.
- Cookson G, Jenkinson T, Jones H, Martinez JV (2021) Best buys and own brands: Investment platforms' recommendations of mutual funds. *Rev. Financial Stud.* 34(1):227–263.
- Cremers MKJ, Petajisto A (2009) How active is your fund manager? A new measure that predicts performance. *Rev. Financial Stud.* 22(9):3329–3365.
- Dangl T, Wu Y, Zechner J (2008) Market discipline and internal governance in the mutual fund industry. *Rev. Financial Stud.* 21(5):2307–2343.
- Del Guercio D, Reuter J (2014) Mutual fund performance and the incentive to generate alpha. *J. Finance* 69(4):1673–1704.
- Del Guercio D, Tkac PA (2008) Star power: The effect of Morningstar ratings on mutual fund flow. *J. Financial Quant. Anal.* 43(4):907–936.
- Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B (Methodological)* 39(1):1–38.
- Dyakov T, Jiang H, Verbeek M (2020) Trade less and exit overcrowded markets: Lessons from international mutual funds. *Rev. Finance* 24(3):677–731.
- Evans AL (2008) Portfolio manager ownership and mutual fund performance. *Financial Management* 37(3):513–534.
- Evans RB, Sun Y (2021) Models or stars: The role of asset pricing models and heuristics in investor risk adjustment. *Rev. Financial Stud.* 34(1):67–107.
- Fama EF, French KR (2010) Luck versus skill in the cross-section of mutual fund returns. *J. Finance* 65(5):1915–1947.
- Ferreira MA, Matos P, Pires P (2018) Asset management within commercial banking groups: international evidence. *J. Finance* 73(5):2181–2227.
- Franzoni F, Schmalz MC (2017) Fund flows and market states. *Rev. Financial Stud.* 30(8):2621–2673.
- Gârleanu N, Pedersen LH (2018) Efficiently inefficient markets for assets and asset management. *J. Finance* 73(4):1663–1712.
- Gaspar JM, Massa M, Matos P (2006) Favoritism in mutual fund families? Evidence on strategic cross-fund subsidization. *J. Finance* 61(1):73–104.
- Gennaioli N, Shleifer A, Vishny R (2015) Money doctors. *J. Finance* 70(1):91–114.
- Giglio S, Maggiori M, Stroebel J, Utkus SP (2021) Five facts about beliefs and portfolios. *Amer. Econom. Rev.* 111(5):1481–1522.
- Glode V, Green RC (2011) Information spillovers and performance persistence for hedge funds. *J. Financial Econom.* 101(1):1–17.
- Greenwood R, Nagel S (2009) Inexperienced investors and bubbles. *J. Financial Econom.* 93(2):239–258.
- Greenwood R, Shleifer A (2014) Expectations of returns and expected returns. *Rev. Financial Stud.* 27(3):714–746.
- Hartzmark SM, Sussman AB (2019) Do investors value sustainability? A natural experiment examining ranking and fund flows. *J. Finance* 74(6):2789–2837.
- Harvey CR, Liu Y (2022) Luck versus skill in the cross-section of mutual fund returns: Reexamining the evidence. *J. Finance* 77(3):1921–1966.
- Harvey CR, Liu Y, Tan EKM, Zhu M (2021) Crowding: Evidence from fund managerial structure. Working paper, Duke University, Durham, NC.
- Huang J, Wei KD, Yan H (2007) Participation costs and the sensitivity of fund flows to past performance. *J. Finance* 62(3):1273–1311.
- Ibert M (2023) What do mutual fund managers' private portfolios tell us about their skills? *J. Financial Intermediation* 53:100999.
- Ibert M, Kaniel R, Van Nieuwerburgh S, Vestman R (2018) Are mutual fund managers paid for investment skill? *Rev. Financial Stud.* 31(2):715–772.
- Jenkinson T, Jones H, Martinez JV (2016) Picking winners? Investment consultants' recommendations of fund managers. *J. Finance* 71(5):2333–2370.
- Jones H, Martinez JV (2017) Institutional investor expectations, manager performance, and fund flows. *J. Financial Quant. Anal.* 52(6):2755–2777.

- Khorana A, Nelling E (1998) The determinants and predictive ability of mutual fund ratings. *J. Investing* 7(3):61–66.
- Khorana A, Servaes H, Wedge L (2007) Portfolio manager ownership and fund performance. *J. Financial Econom.* 85(1):179–204.
- Kim JH (2022) Investor learning and the aggregate allocation of capital to active management. Working paper, Florida State University, Tallahassee, FL.
- Kosowski R, Timmermann A, Wermers R, White H (2006) Can mutual fund “stars” really pick stocks? New evidence from a bootstrap analysis. *J. Finance* 61(6):2551–2595.
- Linnainmaa JT (2013) Reverse survivorship bias. *J. Finance* 68(3):789–813.
- Ma L, Tang Y, Gómez JP (2019) Portfolio manager compensation in the U.S. mutual fund industry. *J. Finance* 74(2):587–638.
- Mankiw GN, Reis R (2002) Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve. *Quart. J. Econom.* 117(4):1295–1328.
- Massa M (2003) How do family strategies affect fund performance? When performance-maximization is not the only game in town. *J. Financial Econom.* 67(2):249–304.
- McLemore P (2019) Do mutual funds have decreasing returns to scale? Evidence from fund mergers. *J. Financial Quant. Anal.* 54(4):1683–1711.
- Morningstar (2021) Morningstar analyst rating for funds: Methodology. White paper, Morningstar research, Chicago.
- Nagel S, Xu Z (2022) Asset pricing with fading memory. *Rev. Financial Stud.* 35(5):2190–2245.
- Newey WK, West KD (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55(3):703–708.
- Pástor L, Stambaugh RF (2012) On the size of the active management industry. *J. Political Econom.* 120(4):740–781.
- Pástor L, Stambaugh RF, Taylor LA (2017) Do funds make more when they trade more? *J. Finance* 72(4):1483–1528.
- Pástor L, Stambaugh RF, Taylor LA (2020) Fund tradeoffs. *J. Financial Econom.* 138(3):614–634.
- Pástor L, Stambaugh RF, Taylor LA (2015) Scale and skill in active management. *J. Financial Econom.* 116(1):23–45.
- Pástor L, Stambaugh RF, Taylor LA, Zhu M (2022) Diseconomies of scale in active management: Robust evidence. *Critical Finance Rev.* 11(3–4):593–611.
- Pollet JM, Wilson M (2008) How does size affect mutual fund behavior? *J. Finance* 63(6):2941–2969.
- Reuter J, Zitzewitz E (2021) How much does size erode mutual fund performance? A regression discontinuity approach. *Rev. Finance* 25(5):1395–1432.
- Roussanov N, Ruan H, Wei Y (2020) Mutual Fund flows and performance in (imperfectly) rational markets? Working paper, University of Pennsylvania, Philadelphia.
- Roussanov NL, Ruan H, Wei Y (2021) Marketing mutual funds. *Rev. Financial Stud.* 34(6):3045–3094.
- Sharpe WF (1998) Morningstar’s risk-adjusted ratings. *Financial Analysts J.* 54(4):21–33.
- Song Y (2020) The mismatch between mutual fund scale and skill. *J. Finance* 75(5):2555–2589.
- Spiegler R (2020) A simple model of a money-management market with rational and extrapolative investors. *Eur. Econom. Rev.* 127:1–17.
- van Binsbergen JH, Kim JH, Kim S (2023) Capital allocation and the market for mutual funds: Inspecting the mechanism. Working paper, University of Pennsylvania, Philadelphia.
- Watson MW, Engle RF (1983) Alternative algorithms for the estimation of dynamic factor, mimic and varying coefficient regression models. *J. Econom.* 23(3):385–400.
- Wilke F (2025) Learning about managerial skill and fund scale from mutual fund analysts. Working paper, Nova School of Business and Economics, Carcavelos, Portugal.
- Zhu M (2018) Informative fund size, managerial skill, and investor rationality. *J. Financial Econom.* 130(1):114–134.