Tilting at Windmills: Biased Benchmarks and the Risk-Taking Response of Mutual Funds*

Richard B. Evans† Sanghyun (Hugh) Kim‡ Thomas A. Maurer§

December 5, 2024

Abstract

Before the 2002 Morningstar ratings methodology change, we find that ranking all US domestic equity funds as a group disadvantages growth funds relative to value. In response, growth funds take significantly more risk. These results are more pronounced near ratings cutoffs and among funds with stronger incentives to attract investor flows. Greater risk-taking enables disadvantaged growth funds to generate returns comparable to advantaged value funds, mitigating the benchmark bias. Extending our analysis to socially responsible investing (SRI), we find similar patterns: morally constrained funds take greater risks to compete with advantaged peers that overweight sin stocks.

JEL-Classification: G11, G24.

Keywords: Mutual Funds, Peer Benchmarks, Risk Taking, Morningstar Ratings, Socially Responsible Investing.

---

*We thank Haoran Wang for outstanding work as our research assistant. Moreover, we are thankful to Roni Michaely, Shiyang Huang, Dunhong Jin, Thomas Schmid, Clemens Sialm Qifei Zhu, and Tim Riley for many helpful comments and suggestions. Finally, we are thankful for the generous funding (GRF grant) provided by the Research Grants Council of Hong Kong.

†Darden School of Business, The University of Virginia, evansr@darden.virginia.edu.

‡Lazaridis School of Business and Economics, Wilfrid Laurier University, shkim@wlu.ca.

§Faculty of Business and Economics, The University of Hong Kong, maurer@hku.hk.

1 Introduction

"I know who I am" said Don Quixote, "and who I may be. I will not be deterred.
I shall do battle with these windmills, for they are giants in my eyes."
— Miguel de Cervantes, Don Quixote

"...sometimes, we're like Don Quixote, tilting at windmills, convinced we can win,
even when it's clear the odds are stacked against us."
— Garth Stein,The Art of Racing in the Rain

Benchmarks are instrumental in assessing fund manager performance, influencing their compensation directly through bonus pay and indirectly through fund flows. Consequently, the choice of benchmarks can have important effects on fund manager behavior, and there is a large academic literature studying these effects. To mention a few, Brown et al. (1996) and Chevalier and Ellison (1997) initiated a large literature showing that mid-year losers increase fund volatility in the second half of the year in the hope to catch up with their peers and mitigate the current relative underperformance. In the context of bonus pay, Evans et al. (2023) document that fund managers exert more effort and perform better when their pay depends on peer rather than pure benchmarks. Finally, a recent forensic finance literature documents that funds often deceive investors and strategically choose mismatched prospectus benchmarks that are easier to beat than benchmarks that more accurately reflect their investment style (Sensoy, 2009; Elton et al., 2014; Chen et al., 2021; Cremers et al., 2022; Mullally and Rossi, 2024; Chen et al., 2024).

We contribute to this literature and investigate the importance of and incentives generated by the choice of a specific comparison group that is used in a peer benchmark evaluation. That is, rather than studying ex-post effects of competition among peers (Brown et al., 1996) or comparing peer versus pure benchmarks in general (Evans et al., 2023), we ask whether the choice of a fund peer group (the benchmark) that is heterogeneous in their systematic risk exposures creates unintended incentives thereby impacting the behavior and risk-taking of fund managers.[1]

Morningstar ratings provide an ideal setting to investigate our research question. First, Morningstar ratings, which are based on relative fund performance within Morningstar peer groups, have a significant and causal impact on fund flows (Del Guercio and Tkac, 2008; Reuter and Zitzewitz, 2021). Ben-David et al. (2022a) even suggest that Morningstar ratings are the main determinant of fund flows. More importantly, the refinement of Morningstar peer groups in June 2002 provides an exogenous event (e.g., Evans and Sun (2021a), Ben-David et al. (2022b), Han et al. (2021)) to estimate the differential effect of benchmarks on the risk-taking of mutual funds with heterogeneous investment styles.

The compensation of fund managers crucially hinges on the assets under management (AUM) and the fee revenue generated by funds (Ibert et al., 2018; Cen et al., 2023; Bai et al., 2023). Thus, the choice of peer groups for relative performance evaluation by Morningstar can affect fund flows. If that peer group is heterogenous in its systematic risk exposures (e.g., grouping value and growth funds), this may generate undesirable incentives for fund managers leading to unintended distortions in fund behavior. In particular, if certain funds are disadvantaged (e.g. growth funds) relative to other funds (e.g. value funds) due to these differences in their systematic risk exposures, they may have an incentive to increase risk taking to compete with their advantaged peers. This increase in risk-taking may result in an inefficient outcome for fund investors as a fund's risk profile may deviate from the desired level.

We argue that Morningstar's original peer group assignment (ranking all U.S. domestic equity funds within a single category) put growth funds at a significant disadvantage when competing with value funds. Evaluating all U.S. domestic equity funds within a single cat-

[1]Note that our economic mechanism works through fund flows rather than bonus pay as for instance in Evans et al. (2023). Moreover, we use Morningstar benchmarks, which are exogenously imposed on a fund, setting our analysis apart from the literature on prospectus benchmarks, which can be strategically chosen by funds.

egory group amounts to using a single-factor model (e.g., the CAPM) for risk adjustment (Evans and Sun, 2021b). It is well-documented in the literature that growth stocks, on average, earn a lower CAPM alpha than value stocks (Fama and French, 1993). Therefore, a sizable value premium places growth funds at a significant disadvantage when competing against value funds under the single-category-group (or single-factor) framework. This disadvantage was lifted in June 2002 when Morningstar began grouping U.S. domestic equity funds into nine (three-by-three) categories, classified by the size-by-value/growth Morningstar equity style box. As a result, growth funds are no longer compared to value funds under this new multi-category (or three-factor) framework.

Prior to the refinement of Morningstar peer groups, we expect that growth funds anticipate their disadvantage, and are concerned about net outflows due to their lower expected ratings. To boost their chances of improving ratings and mitigate net outflows, growth funds can increase their risk taking. For instance, Sharpe (1998) shows that when the Sharpe ratio of a fund is sufficiently large, then increasing the risk exposure improves Morningstar ratings to some extent. In addition, tilting a growth portfolio towards the market or value improves the Sharpe ratio and enables growth funds to catch up with value funds. While doing so, the overall risk exposure of funds usually increases. Another interpretation is that the disadvantaged growth fund managers need to exert greater effort and be more active to catch up with their advantaged value fund peers. A common argument in the literature is that activeness is associated with tracking error or idiosyncratic volatility.

Following the peer group refinement, however, growth and value funds are no longer compared against each other. Therefore, we expect that differences in risk taking across growth and value funds significantly decrease following the refinement.[2]

[2] We implicitly assume that growth fund managers choose to invest in growth stocks despite the fact that they are aware of the well-documented value premium. We assume that these managers intend to cater to investors who seek an exposure to growth stocks. At the same time, the fund managers are aware that Morningstar ratings are an important determinant of fund flows. Thus, they optimize their asset allocation within the set of growth stocks (such that their investment style qualifies for the growth category) to maximize

We begin our empirical analysis by showing that growth funds indeed had lower average Morningstar ratings than value funds when all equity funds were rated together in a single peer group. This finding is consistent with our argument that, in the presence of the value premium, growth funds are at a significant disadvantage when competing against value funds. Prior to the peer group refinement, growth funds had an average Morningstar rating of 2.98, whereas the corresponding figure for value funds was 3.18. The difference of a fifth of a star is economically significant. Following the refinement, the average ratings of growth versus value funds became indistinguishable as intended by the refinement of peer groups.

Moreover, prior to the refinement, Morningstar ratings were highly correlated with the past returns of the value factor (HML) of the Fama-French 3-factor model (Fama and French, 1993). The better the value factor had performed in the past 36 months, the worse the ratings for growth funds compared to value funds. This correlation also became muted following the refinement of Morningstar peer groups. This time-series test corroborates our argument that lower ratings of growth funds were driven by a superior performance of the value factor.

Next, we move onto testing our main hypothesis that inappropriate benchmarks have a significant effect on the risk-taking of mutual funds. Using a difference-in-differences (DiD) setting around the refinement of Morningstar peer groups in June 2002, we find that, prior to the refinement, disadvantaged growth funds took substantially greater risks compared to advantaged value funds. The annualized idiosyncratic volatility and tracking error of growth funds was, on average, 0.93% respectively 0.94% higher than that of value funds. This is an economically significant difference, given that the median idiosyncratic volatility and tracking error are 3.72% respectively 4.09%.

To further corroborate our mechanism, we exploit the time-variation in the expected value premium, using the value spread as a proxy. Cohen et al. (2003) find that the value

their ratings.

spread positively predicts the value premium. Thus, we expect that when the value spread was larger, the disadvantage of growth funds would be more severe. Indeed, we find that disadvantaged growth funds increased risk taking even more when the value spread was higher, in the pre-refinement period. In particular, a one standard deviation increase in the value spread lead to an increase in idiosyncratic volatility and tracking error of disadvantaged growth funds (relative to value funds) by 0.24% and 0.28%, respectively. In contract, this correlation between the value spread and greater risk-taking of growth funds relative to value funds became muted following the refinement of Morningstar peer groups.

Our mechanism implies that the increase in risk taking should be especially pronounced for growth funds that have stronger incentives to mitigate their disadvantage. We identify two such cases. First, since Morningstar ratings are discrete with only five possible outcomes, growth funds that are ranked close to one of the four rating thresholds are more concerned about their disadvantage compared to funds that are far from a threshold. Second, growth funds with manager compensations that are linear (rather than concave) in the fund size are more concerned about a disadvantage. We then show that unintended risk taking incentives induced by benchmarks indeed have a greater impact on growth funds if they have greater incentives to mitigate their disadvantages.

Studying growth funds that differ with respect to their incentives to mitigate a disadvantage further allows us to control for within style time fixed effects. Thus, we can sharpen our tests and control for all time-varying style specific unobservables. We confirm that our risk taking mechanism of funds in response to a disadvantage is robust.

In addition, this setting allows us to investigate the efficacy of risk taking. Specifically, we find that growth funds that increase risk taking more due to stronger incentives to mitigate a disadvantage have higher Morningstar ratings. Accordingly, the increase in risk taking is effective to mitigate the disadvantage. This is evidence that the additional risk taking is a manifestation of exerting more effort.

We emphasize that our results have implications beyond the specific context of growth versus value funds. To demonstrate that our main results can be generalized to other settings where benchmarks put some funds at a disadvantage, we extend our analysis to the context of socially responsible investing (SRI) in recent years. The Environmental, Social, and Governance (ESG) considerations have become important in the asset management industry (Krueger et al., 2020). Mutual funds are still evaluated relative to peers that may differ in their efforts in impact investing when receiving star ratings.[3] We explore how the existing benchmarks, which fail to consider a new important factor, will affect the risk-taking behavior of mutual funds in the post-2002 period.

To measure the extent of a fund's SRI, we use the fund's divestment from sin stocks (related to the alcohol, tobacco, and gambling industries) (Hong and Kacperczyk, 2009).[4] First, we show that morally constrained funds (those divesting from sin stocks) have a systematic disadvantage. Further, we find that morally constrained funds exhibit more risk taking. Thus, our SRI results corroborate our earlier results from growth versus value funds: disadvantaged funds tend to take greater risks, presumably in an effort to compete with their advantaged peers.

Our paper is closely related to a recent study by Evans et al. (2023), who contrast the use of pure versus peer benchmarks and find that peer benchmarks incentivize managers to be more active and exert more effort, resulting in stronger performance. Importantly, they point out that benchmarks used to determine a manager's bonus may differ from prospectus benchmarks. While Evans et al. (2023) focus on benchmarks used for bonus payments of fund managers, we investigate the implications of benchmarks used by fund investors to

[3] In response to growing interest in sustainability of the U.S. mutual fund market, Morningstar introduced sustainability ratings (also known as globe ratings) in March 2016 (Hartzmark and Sussman, 2019). Morningstar's globe ratings do not add a new dimension to Morningstar's original star ratings, which is the focus of our paper.

[4] Lo and Zhang (2023) estimate within the Treynor and Black (1973) framework that the cost of divesting from sin stocks ranges from 0.6% to 3.3% in forgone alpha per year.

guide their wealth allocation to funds. That is, mutual fund investors may use performance metrics and benchmarks to rank managers and decide in which fund to invest, which may differ from prospectus benchmarks or the evaluation methods used to determine manager bonus payments.

Our paper is also related to a large literature on mutual fund risk-taking. Roughly speaking, the literature attributes variations in risk-taking either to fund manager skills, risk-shifting due to non-linearities in the relationship between performance and manager compensation, or excessive risk-taking of past losers who try to catch up with past winners. Brown et al. (1996) and Chevalier and Ellison (1997) initiated a large literature showing that past losers take more risks than winners in the remainder of the year.[5] The economic channel in our study differs from the extant literature on risk-shifting. We are not investigating past losers. Rather, we focus on the effects of peer group assignments that are relevant for mutual fund investors when allocating their wealth to funds (i.e., fund flows). Specifically, funds that expect to be at a disadvantage due to an unfair peer group assignment anticipate a systematic disadvantage when competing against their advantaged peers. This incentivizes them to take more risks in an attempt to compete with their advantaged peers. In robustness checks, we further show that our channel is present when controlling for past (under)performance.

2 Background on Morningstar ratings and peer groups

At the end of every month, mutual fund share classes are rated by Morningstar on an integer scale of one star (the lowest rating) to five stars (the highest rating) on the basis of Morningstar Risk-Adjusted Returns (MRAR) over the prior three, five, and ten years, depending on data availability. Within Morningstar peer groups, the top 10% of funds

[5] Without getting into details, this literature includes Chen and Pennacchi (2009), Basak et al. (2007), Schwarz (2012), Busse (2001), Basak et al. (2008), Cullen et al. (2012), Spiegel and Zhang (2013), Kim (2019), Lee et al. (2019), Ma and Tang (2019), Ma et al. (2019), Del Guercio and Tkac (2002), Huang et al. (2011), Han et al. (2021).

receive five stars, the next 22.5% four stars, the middle 35% three stars, the next 22.5% two stars, and the bottom 10% one star.

'Overall' ratings are determined by the weighted averages of three, five, and ten-year ratings, depending on data availability, rounded to the nearest integer value. Share classes less than three years old are not rated. Share classes at least three years old but less than five years old are rated based only on three-year ratings. Share classes at least five years old but less than ten years old are rated based on three-year (40 percent weight) and five-year (60 percent weight) ratings. Share classes at least ten years old are rated based on three-year (20 percent weight), five-year (30 percent weight), and ten-year (50 percent weight) ratings. Thus, the past three-year performance (36 monthly returns) accounts for 53 $(= .2 \times (3/3) + .3 \times (3/5) + .5 \times (3/10))$ to 100 percent of the ratings.

On June 30, 2002, Morningstar refined its peer groups relative to which funds are rated. Prior to June 2002, Morningstar used category groups based on broader asset classes (e.g., U.S. domestic equity) to define a peer group. After the change, Morningstar began using specific categories that match mutual funds' investment styles (e.g., U.S. domestic equity large-cap value) define peer groups to assign ratings. Specifically, Morningstar began rating U.S. domestic equity funds within one of nine investment style categories defined based on the size tilt and growth/value orientation ($\{$Large, Mid-cap, Small$\}\times\{$Growth, Blend, Value$\}$) of the fund. Following prior studies (Evans and Sun, 2021b; Ben-David et al., 2022b; Han et al., 2021), we exploit the refinement of Morningstar peer groups in June 2002 for the main identification in our paper.

3 Data and construction of variables

Our primary data come from Morningstar Direct, from which we obtain information on Morningstar ratings, Morningstar (investment style) categories, daily and monthly returns, monthly total net assets (TNA), annual expense ratios, and annual turnover ratios for all open-ended U.S. domestic equity mutual funds from 1988 to 2022. We also obtain mutual fund holdings data directly from Morningstar. Fund holdings data from Morningstar are not only more frequently available than those from Thomson/Refinitiv, which are commonly used in the mutual fund literature, but also provide a more comprehensive representation of the actual composition of the funds' portfolios (Elton et al., 2011; Ammann et al., 2022).

We extract Form N-SAR semiannual reports through the SEC's Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system. To merge N-SAR data with Morningstar data, we first match funds from Morningstar data and those from the CRSP Mutual Funds Database using CUSIP, ticker, and fund name (Berk and van Binsbergen, 2015; Pástor et al., 2015). Then, we match funds from N-SAR data with those from the CRSP-Morningstar merged dataset based on CRSP historical fund names.

Idiosyncratic volatility

For each fund-month, we measure the fund's idiosyncratic volatility (iVol) as the standard deviation of the residuals from the regression of a fund's daily excess returns on the daily returns of the Fama-French 3 factors (Fama and French, 1993):

$$r_{i,t,d} - r_{f,t,d} = a_{i,t} + b_{i,t}\left(r_{m,t,d} - r_{f,t,d}\right) + s_{i,t}SMB_{t,d} + h_{i,t}HML_{t,d} + \varepsilon_{i,t,d}, \tag{1}$$

where $r_{i,t,d}$ is the fund's return, $r_{f,t,d}$ is the risk-free rate, $r_{m,t,d}$ is the market return, $SMB_{t,d}$ is the return on the small-minus-big portfolio, $HML_{t,d}$ is the return on the high-minus-low portfolio, and $\varepsilon_{i,t,d}$ is the residual on day d of month t. The Fama-French factor returns are obtained from Kenneth French's data library. We annualize the fund's idiosyncratic volatility by multiplying it by $\sqrt{252}$, and then report it as a percentage.

Tracking error

For each fund-month, we measure the tracking error (TE) as the standard deviation of the fund's benchmark-adjusted returns ($r_{i,t,d} - r_{b,t,d}$), where $r_{i,t,d}$ is the fund's return and $r_{b,t,d}$ is the benchmark index return on day d of month t. We select the benchmark index as the one that minimizes the fund's tracking error, choosing from the Russell 1000, Midcap, 2000, and their value/growth variants. Index daily returns are calculated as the weighted average of stock returns based on the most recent monthly index holdings, with constituent weights sourced from FTSE/Russell and stock returns from CRSP. We annualize the fund's tracking error by multiplying it by $\sqrt{252}$, and then report it as a percentage.

$\mathbb{1}$(*Close to a rating threshold*), or simply $\mathbb{1}$(*Close*)

For each fund-month, we identify whether the fund's performance ranking is moving closer to a Morningstar rating threshold, based on the fund's decile ranking its return at month $t-36$ relative to its peers as of month $t-1$. The intuition is that the fund's return at $t-36$ affects its three-year rating until month $t-1$, but not in month t. As a result, the fund's percentile ranking will move closer to an upper (lower) rating threshold when it rolls out of an extremely poor (good) performance month from 36 months ago (Kim, 2022). Specifically, $\mathbb{1}(Close_{i,t-1})$ takes a value of one if the fund's return at $t-36$ places it in the bottom (top) decile relative to its peers with the same rating as of month $t-1$ for ratings of one to four stars (two to five stars), and zero if the fund's return at $t-36$ falls within the middle eight deciles. To confirm the validity of our identification, we find in untabulated results that funds with $\mathbb{1}(Close_{i,t-1}) = 1$ are 48% more likely to experience rating changes in month t, compared to funds with $\mathbb{1}(Close_{i,t-1}) = 0$, suggesting that we are successfully identifying funds near rating thresholds. We note that this information is available to fund managers in month $t-1$, but the percentile rankings and ratings for month t will still be uncertain to them, as these depend on the fund's performance relative to its peers during month t. In addition, it is important to note that the return 36 months ago is unlikely to have an impact on the fund manager's behavior today other than through the Morningstar rating channel. As such this variable is exogenous, and it is a suitable instrument to classify funds as being located either close to or far from a rating threshold.

$\mathbb{1}$(*Linear advisory contract*), or simply $\mathbb{1}$(*Linear*)

For each fund-year, we measure Coles' incentive rate (*CIR*) as the difference between the marginal advisory fee rates between the highest TNA and lowest TNA breakpoints, scaled by the effective advisory fee rate (Coles et al., 2000; Massa and Patgiri, 2009). Advisory fee rates are obtained from N-SAR filings.[6] The variable *CIR* measures the concavity of a fund's advisory fee schedule. The variable $\mathbb{1}$(*Linear*) is an indicator variable that takes a value of one if a fund's advisory contract is 'high-incentive' – that is, the advisory fee rate is linear with respect to the fund's TNA ($CIR = 0$) – and zero if the fund's advisory contract is 'low-incentive' – that is, the advisory fee rate is concave with respect to the fund's TNA ($CIR < 0$). About 57% of the fund-years in our sample have a linear advisory contract.

$\mathbb{1}$(*Morally constrained*)

Following Hong and Kacperczyk (2009), we define sin stocks as those of firms operating in the alcohol, tobacco, and gambling industries. The variable $\mathbb{1}$(*Morally constrained*) is an indicator variable that takes a value of one if a fund's portfolio weight in sin stocks is smaller than that of its benchmark index ($\text{Weight}^{\text{Fund}}_{\text{Sin}} < \text{Weight}^{\text{Index}}_{\text{Sin}}$), and zero otherwise. About two-thirds of the fund-month observations in our sample are considered 'morally constrained'

[6] Form N-SAR filings are available for some funds starting in 1994 (and for most funds starting in late 1995) until 2018, after which they were replaced by Form N-CEN filings in 2019. For each fund, we backfill the data using the first available data point and forward-fill the data using the last available data point. Our results are insensitive to this backfilling method, even if we restrict our sample to the period from 1995 to 2018.

– that is, divesting from sin stocks, underweighting them by 0.48% on average (0.73% at the median) relative to their benchmark index. These figures are economically significant, given that sin stocks constitute only about 1–3% of the benchmark indices in our sample.

Other variables

The remaining variables are defined in the usual way. As is standard in the literature, we aggregate share-class-level variables to the fund level across all share classes belonging to the same fund. The summary statistics are reported in Table 1.

[Insert Table 1]

4 Growth versus value funds

Prior to the June 2002 change in the Morningstar ratings methodology, all U.S. equity mutual funds were ranked relative to one another. However, since style performance is a significant driver of fund performance but out of control of fund managers, Morningstar ratings were more a reflection of style performance rather than manager skills (Ben-David et al., 2022b). This issue became once again prevalent with the dot-com crash, resulting in a drastic underperformance of growth stocks and significantly lower ratings of growth funds. Recognizing this problem, Morningstar made a major change in their star ratings methodology in June 2002. Subsequently funds were ranked relative to only their peers within their respective style categories, e.g., small growth relative to other small growth, and large value relative to other large value funds.

On top of the issue that ratings were largely driven by style performance, we further argue that the peer group choice prior to June 2002 has put growth funds at a significant disadvantage when competing against value funds due to the existence of the value premium (Fama and French, 1993). This is an issue that has not been studied in the literature.

In the following we first document that growth funds were disadvantaged (value funds were advantaged) prior to June 2002 (Sections 4.1 and 4.2). Second, we show that this disadvantage has provided incentives for growth funds to increase risk taking (Sections 4.3 to 4.6).

4.1 Disadvantage of growth versus value funds

We begin our empirical analysis by examining whether growth funds had a disadvantage competing against value funds when differences in investment styles (and underlying factor exposures) were not adequately accounted for in Morningstar's relative performance evaluations prior to June 2002.

In Figure 1 we visually inspect whether growth funds had a disadvantage relative to value funds by plotting the time-series of Morningstar ratings averaged each month across all funds belonging to growth (Large Growth, Mid-Cap Growth, and Small Growth) respectively value categories (Large Value, Mid-Cap Value, and Small Value) from January 1988 to December 2022. Prior to the 2002 Morningstar peer group refinement, growth funds tended to have lower ratings compared to value funds. Following the change, this disadvantage has become muted. The average ratings for both growth and value funds hover around 3 stars, which is the midpoint in Morningstar ratings.

[Insert Figure 1]

Next, we estimate the extent to which growth funds were disadvantaged competing against value funds. In Table 2, we find that prior to the ratings methodology change, growth funds had average ratings of 2.98 while value funds had average ratings of 3.18, resulting in a statistically significant difference of −0.20 points in Morningstar ratings. The difference is also economically meaningful, representing 0.19 standard deviations in Morningstar ratings.

Following the change, the difference has decreased to virtually zero (0.003) and has become statistically insignificant.

[Insert Table 2]

The true extent of the disadvantage of growth funds is likely more severe than what we estimate in Table 2. In response to their disadvantage growth funds increase risks in an attempt to catch up with value funds and to mitigate their disadvantage. Indeed, this is the main finding of our paper. This mitigation results in a smaller gap in Morningstar ratings between growth and value funds in the data than if growth funds did not adjust their risk taking. It is difficult to quantify this mitigation effect as we do not observe ratings of disadvantaged growth funds that do not increase risk taking. However, in Section 4.7 we provide evidence that the mitigation effect is economically meaningful. In particular, we show that disadvantaged growth funds with stronger incentives to compete for fund flows (and thus Morningstar ratings) take more risk and subsequently achieve higher ratings compared to less incentivized growth funds.

4.2 Time-varying disadvantage

We argue that the disadvantage of growth relative to value funds is driven by factors unrelated to skills of fund managers and hence out of their control. In particular, we argue that the disadvantage stems from the differences in investment styles, and specifically factor exposures to the value premium. Accordingly, we expect that differences in ratings between growth and value funds are correlated with past performance of the value factor (HML) prior to the ratings methodology change.

To test this prediction, we estimate the following linear regression model:

$$\begin{aligned} \textit{Morningstar rating}_{i,t} = {} & \delta \mathrm{HML}_{t-35,t} \times \mathbb{1}\left(\mathrm{Growth}_{i,t-1}\right) \\ & + \beta_1 \mathrm{HML}_{t-35,t} + \beta_2 \mathbb{1}\left(\mathrm{Growth}_{i,t-1}\right) + \theta_i + \varepsilon_{i,t}. \end{aligned} \tag{2}$$

where i indexes mutual funds, and t indexes time (in months). $\textit{Morningstar ratings}_{i,t}$ represents fund i's Morningstar rating at time t. $\mathbb{1}(\textit{Growth}_{i,t-1})$ is an indicator variable that takes a value of one if fund i belongs to one of the growth categories (Large Growth, Mid-cap Growth, or Small Growth) at time $t-1$, and zero if fund i belongs to one of the value categories (Large Value, Mid-cap Value, or Small Value) at time $t-1$. $\textit{HML}_{t-35,t}$ is the cumulative return on the value factor (HML) from the Fama-French 3-factor model (Fama and French, 1993) over the past 36 months. θ_i denotes fund fixed effects. Standard errors $\varepsilon_{i,t}$ are double-clustered by fund and time.

Based on the Morningstar ratings methodology we choose the past 3-year period $(t-35, t)$ as the relevant return of HML. As described in Section 2, Morningstar ratings are determined by the weighted averages of the past 3-, 5-, and 10-year fund performance, depending on data availability. The past 3-year performance is the most relevant, accounting for 53% to 100% of the ratings.

For identification, we rely on fund fixed-effects, exploiting within-fund variations in the ratings that are correlated with time-series variations in the past performance of HML. To examine how risk-taking behavior changes as peer groups change, we run the regressions in Equation (2), separately for the pre-refinement period January 1988 to May 2002, and the post-refinement period June 2002 to December 2022.

We present the results in Table 3. The key coefficient of interest is δ, representing the interaction between growth category and past HML performance. As predicted, δ is negative and statistically significant, in the pre-change period (columns (1) and (2)). That is, the strong performance of the value factor tended to exacerbate the disadvantage of growth funds relative to value funds, driving the wedge in the ratings between growth funds and value funds. In- or exclusion of fund fixed effects has no material effect on our estimates. For brevity we only describe in detail the estimates that include fund fixed effects (column (2)).

The slope coefficient $\beta_1 = 1.64$ suggests that one standard deviation (SD) increase in the 36-month return $\text{HML}_{t-35,t}$ (0.23) is associated with an increase in value funds' ratings by 0.38 points, or a 0.36 SD increase in ratings. Furthermore, the sum $\delta + \beta_1 = -1.43$ suggests that one SD increase in $\text{HML}_{t-35,t}$ is associated with a decrease in growth funds' ratings of a similar magnitude (0.33 points or 0.31 SD decrease in ratings), driving the wedge in the ratings between growth funds and value funds by 0.71 points, or 0.67 SD of the ratings.

In contrast, δ becomes insignificant and virtually zero in the post-change period (columns (3) and (4)), suggesting that the Morningstar ratings methodology change on June 30, 2002 was effective in eliminating the disadvantage that growth funds were facing when competing with value funds.

[Insert Table 3]

4.3 Differential risk-taking of growth versus value funds

We now test our hypothesis that growth funds took more risk when they were placed at a disadvantage since Morningstar's benchmark peer group assignments did not account for the value premium prior to June 2002.

We consider two risk measures: a fund's idiosyncratic volatility (iVol) and tracking error (TE). The two measures are described in section 3. Both risk measures crucially differ from total volatility. While total volatility allows for various interpretations, iVol and TE provide more detail information about the nature of risk taking. For instance, an increase in total volatility could mean increasing (i) tilting the portfolio towards the market or other factors (e.g., over-weighting high beta stocks), (ii) deviating from the benchmark and gamble in hope for high realized returns, or (iii) exerting more effort. A common argument in the literature is that activeness and more effort is associated with higher TE and iVol. In contrast, iVol does not pick up (i) as long as we control for the relevant factors, and neither does TE as it is measured with respect to the index that best matches the fund's investment style. Accordingly, likely interpretations of an increase in iVol and TE is that a fund manager either gambles, or exerts more effort.

More gambling is an undesirable outcome and likely leads to subsequent underperformance, while exerting more effort is desirable and should come with a stronger future performance. In section 4.7 we provide evidence that managers exert effort as the increase in iVol and TE of disadvantaged growth funds appears to increase their Morningstar ratings on average.

In a first step we visually inspect in Figures 2 and 3 whether growth funds took more risks relative to value funds by plotting the time-series of annualized iVol and TE averaged each month across all funds belonging to growth (Large Growth, Mid-Cap Growth, and Small Growth) respectively value categories (Large Value, Mid-Cap Value, and Small Value).

Growth funds had higher iVol and TE compared to value funds prior to the 2002 ratings methodology change. Following the change the gap in iVol and TE across the two investment style categories have been clearly closing.

[Insert Figures 2 and 3]

More formally, we estimate the extent to which disadvantaged growth funds increased risk taking using the following difference-in-differences (DiD) regression that includes all mutual funds i and months t from 1988 and 2022:

$$Risk\ taking_{i,t} = \delta\left(\mathbb{1}(Growth_{i,t-1}) \times \mathbb{1}(Pre_t)\right) + \beta\mathbb{1}(Growth_{i,t-1}) + \gamma\Gamma_{i,t-1} + \theta_t + \varepsilon_{i,t}. \quad (3)$$

$Risk\ taking_{i,t}$ of fund i and month t is measured by the annualized iVol or TE. $\mathbb{1}(Pre_t)$ is an indicator variable that takes a value of one if time t is prior to the refinement of Morningstar peer groups in June 2002, and zero otherwise. $\mathbb{1}(Growth_{i,t-1})$ is defined in specification (2). We control for the lagged fund characteristics $\Gamma_{i,t-1}$ which include the natural logarithm of total net assets (TNA), expense ratio, turnover ratio, and percentile rank (in the cross-section) of a fund's year-to-date (YTD) return. θ_t denotes time fixed effects. Our sample spans the period from 1988 to 2022. Standard errors $\varepsilon_{i,t}$ are double-clustered by fund and time.

We present the results in Table 4. Columns (1) and (2) report results for iVol, and columns (3) and (4) for TE. First, we estimate the DiD model (3) only controlling for time fixed effects but excluding all other fund controls $\Gamma_{i,t-1}$ (columns (1) and (3)). The coefficient of interest is δ, representing the interaction between the growth category and the pre-June 2002 period. δ is positive and significant across both risk measures. This implies that growth funds took more risks relatively to value funds before June 2002 when they had to compete with value funds despite their disadvantage due to the value premium.

The estimates are not only statistically significant but also economically meaningful. We find that the annualize iVol and TE of disadvantaged growth funds compared to advantaged value funds were 0.93% respectively 0.94% higher prior to June 2002. This amounts to over a quarter or two-fifths the median iVol of 3.72% respectively TE of 4.09% across all fund-month observations in our sample.

Controlling in addition for the log(TNA), expense ratio, turnover ratio, and percentile ranking of the YTD return does not have any noticeable effect on our estimates (columns

(2) and (4)). The point estimates of δ only slightly (and statistically insignificantly) change from 0.93 to 0.92 in the case of iVol, and from 0.94 to 0.95 in case of TE.

[Insert Table 4]

Among the fund level control variables the percentile ranking of a fund's YTD return is the most relevant for our purpose. A large literature initiated by Brown et al. (1996) and Chevalier and Ellison (1997) shows that past loser funds increase risk taking in an attempt to catch up with winners. This means that the YTD return should negatively predict risk taking. While the literature has shown that this is true for total volatility as a risk measure, interestingly, the effect of the YTD return is small and slightly positive when we analyze iVol and TE. Most important for our analysis, our mechanism is not absorbed or driven by the past loser risk taking channel. This alleviates the potential concern that growth funds might be past losers as they on average underperform value funds, and therefore, they increase risk-taking in hope to catch up with past winners.

Note further that the channel of Brown et al. (1996) and Chevalier and Ellison (1997) is fundamentally different from our economic mechanism. In their story a poor past performance (either due to bad luck or low skill) creates ex-post incentives to take risk. In contrast, our mechanism is novel as it describes ex-ante risk taking incentives as an unintended by-product of a benchmark choice which leaves certain funds at a disadvantage. To further set our mechanism apart from the past loser/winner channel in section 4.7 we provide evidence that the additional risk taking of disadvantaged growth funds appears to be efficient to the extent that it mitigates the underperformance relative to value funds. In contrast, Huang et al. (2011) show that risk taking of past loser funds is inefficient and leads to subsequent underperformance.

Our identification strategy relies on the DiD setting by comparing fund risks in different investment style categories before versus after June 2002. A possible concern is that our time window from 1988 to 2022 is too long to attribute our risk taking findings to the Morningstar peer group refinement in June 2002. In other words, there could be other changes (even slow trends) in our long time window, such as a weakening performance of the value factor (HML) for example, and we might be merely picking up the effects of that. To address this concern we experiment with different time windows to establish the robustness of our results.

Table 5 provides the estimates of our model (3) (including fund controls and time fixed effects) for the following alternative time windows: 2001 to 2003, 2000 to 2004, and 2000 to 2004 while removing 2002 (i.e., the final year of the dot-com crash, and the year of the Morningstar peer group refinement). These windows are narrow around June 2002 and make it unlikely that other events drive our results. We find that, if anything, our results are stronger in these sub-samples. The estimated δ is always significant. The point estimates of δ are comparable and slightly larger in the case of iVol (δ is 1.00, 1.05, and 1.26 in the three narrow windows in Table 5 compared to 0.93 in the baseline estimation in Table 4), and they are substantially larger in case of TE (δ is 1.22, 1.43, and 1.58 in Table 5 compared to 0.95 in Table 4).

[Insert Table 5]

To sum up, our results suggest that benchmark or peer group choices can distort mutual funds' behavior. We document that disadvantaged growth funds have significantly higher iVol and TE relative to value funds prior to the June 2002 Morningstar peer group refinement, presumably to catch up with the value funds that enjoy an advantage due to the value premium.

4.4 Time-varying risk taking

When examining the risk-taking behavior, we have thus far implicitly assumed that the value premium, which puts growth funds at a disadvantage when competing against value funds, is constant over time. However, the asset pricing literature suggests that the value premium is time-varying (Cohen et al., 2003). As shown in Section 4.2, ratings were highly correlated with the past performance of the value factor in the pre-June 2002 period. Accordingly, when the value premium is expected to be larger, and growth funds anticipate a greater disadvantage, then they should take even more risks in an attempt to compete with value funds.

Building on Cohen et al. (2003), we use the value spread at time $t-1$ as a proxy for the conditional value premium. The value spread, $\log(B/M)^{\text{Hi}-\text{Lo}}_{t-1}$ is defined as the log difference of book-to-market ratios of the value portfolio (i.e., the top book-to-market decile) and the growth portfolio (i.e., the bottom book-to-market decile) at the end of the previous June.[7]

To test our prediction, we estimate the following linear regression model:

$$\textit{Risk taking}_{i,t} = \delta\left(\log(B/M)^{\text{Hi}-\text{Lo}}_{t-1} \times \mathbb{1}(\textit{Growth}_{i,t-1})\right) + \beta\log(B/M)^{\text{Hi}-\text{Lo}}_{t-1} + \gamma\Gamma_{i,t-1} + \theta_i + \varepsilon_{i,t}. \tag{4}$$

All variables are defined in specification (4), and standard errors are double-clustered by fund and time.

For identification, we rely on fund fixed-effects, exploiting within-fund variations in iVol and TE that are correlated with time-series variations in the value premium. To examine how risk-taking behavior changes as peer groups change, we run the regressions in Equation (4), separately for the pre-change period (January 1988 to May 2002) and the post-change period (June 2002 to December 2022).

We present the results in Table 6. The key coefficient δ, representing the interaction between the growth category and the value premium. As predicted, δ is positive and statis-

[7]We thank Ken French for making these data available on his website. To construct the value spread, stocks are first sorted into decile portfolios at the end of June based on book-to-market ratios. The aggregate book value and market value of each portfolio are then computed by summing book values and market values, respectively, across all stocks belonging to the portfolio.

tically significant for iVol and TE in the pre-June 2002 period (columns (1) and (3)). When the the value spread is larger (and we expect the value premium to be high and value stocks to outperform growth stocks by a larger margin), then growth funds tended to take even more risks, presumably to keep up with value funds. In case of iVol we estimate $\delta = 0.54$ suggesting that a one SD increase in the value spread (0.44) drives the wedge in iVol between growth and value funds by 0.24%. This is economically significant given that the median fund iVol is 3.72%. Similarly, the effect is meaningful and similar in magnitude in the case of TE. We estimate $\delta = 0.64\%$ implying that a one SD increase in the value spread (0.44) increases the tracking error by 0.28%. This is economically meaningful compared to the median TE of 4.09%.

In contrast to the pre-change period, for both risk measures δ is indistinguishable from zero in the post-change period (columns (2) and (4)). This suggests that the refinement of Morningstar's peer groups in June 2002 was effective in eliminating unintended incentives of disadvantaged growth funds to increase risk taking.

[Insert Table 6]

To sum up, we document that growth funds increase their risk taking relative to value funds even more in the pre-June 2002 period when the value premium is high and we expect valeu funds to outperform growth stocks by a larger margin. This is consistent with the idea that growth funds are at an even larger disadvantage when the value premium is high (prior to June 2002), and are required to increase risk even more to compete with advantaged value funds.

4.5 Morningstar rating cutoffs

Our mechanism postulates that disadvantaged growth funds worry about their relative performance as it affects their Morningstar ratings. Recall that Morningstar ratings are discrete with only five possible outcomes. Thus, our mechanism is arguably most relevant for funds that are ranked close to (either slightly below or above) one of the four cutoffs. To illustrate this consider two disadvantaged growth funds that currently have a four-star rating. The lower and upper cutoffs for the four-star rating are the 67.5- and 90-percentile when ranking all funds. Suppose the first fund ranks just slightly below the cutoff to the five-star rating (e.g., 89.9-percentile), while the second fund is far from both cutoffs (e.g., 78.75-percentile). The first fund is more worried about its disadvantage then the second fund. Accordingly, the first fund will more aggressively increase its risk compared to the second fund.

We exploit this insight and test whether disadvantaged growth funds that are ranked closer to a Morningstar rating cutoff take more risk. We can measure the distance of a fund to the Morningstar cutoffs in month $t-1$ as the absolute difference between the fund's percentile rank and the percentile of the closest rating cutoff. A potential concern about this measure is that fund managers may not know their percentile rank as it is difficult to keep track of all their peers, especially in real time, and there is a lot of uncentainty about their rankings and ratings for month t as these depend on the fund's performance relative all of its peer funds in month t. Moreover, the (relative) performance of a fund within the past few months may be related to its current risk taking behavior, and thus, using the current ranking of a fund based on its 36-month performance may not cleanly identify the effect of the fund's distance to a rating cutoff.

Following Kim (2022) we use the one-month return 3 years ago (i.e., the return at $t-36$) $r_{i,t-36}$, and classify a fund as close to a Morningstar cutoff ($\mathbb{1}(\text{Close}_{i,t-1}) = 1$) if $r_{i,t-36}$ is in either the bottom or the top decile relative to its peers with the same rating in month $t-1$. The idea of the instrument is as follows. Recall that the past 3 years of monthly returns are the most relevant and contribute between 53% and 100% to the rating (Section 2). Accordingly, the one-month return 3 years ago is relevant for the current rating but will be significantly less relevant (or even irrelevant in case the fund is less than 5 years old) for the rating next month. If a fund had a large negative (positive) one-month return 3 years ago, then this likely leads to a boost (drag) in the rating next month, and the fund is more likely to get a chance to gain an additional star (face the risk to lose one star) next month. Note that in case of a current one- or five-star rating the only possible rating change is an upgrade respectively downgrade. Therefore, in these cases we adjust our aforementioned classification and a fund is identified as being close to a cutoff if $r_{i,t-36}$ is in the bottom respectively top decile.

To confirm the validity of our identification, we find in untabulated results that funds with $\mathbb{1}(Close_{i,t-1}) = 1$ are 48% more likely to experience rating changes in month t, compared to funds with $\mathbb{1}(Close_{i,t-1}) = 0$, suggesting that we are successfully identifying funds near rating thresholds. Finally, note that the one-month return 3 years ago is unlikely to be related to current risk taking, except through our Morningstar rating channel. Accordingly, $r_{i,t-36}$ is suitable to cleanly identify a fund's distance to a rating cutoff, while it is unlikely to have an effect on the current risk taking of the fund through any other channel than what we intend to measure.

We test the implications of a fund's distance to rating cutoffs on the risk-taking of disadvantaged growth funds in two steps. First, we re-run the DiD specification in (3) separately for the sub-samples of funds which are close to ($\mathbb{1}(\text{Close}_{i,t-1}) = 1$) respectively distant ($\mathbb{1}(\text{Close}_{i,t-1}) = 0$) from a rating cutoff. We expect that δ is larger (smaller) for the sub-sample of funds that are close to (distant from) a rating cutoff.

In a second step we estimate a combined regression with all funds and a triple interaction between $\mathbb{1}(\text{Growth}_{i,t-1})$, $\mathbb{1}(\text{Pre}_t)$ and $\mathbb{1}(\text{Close}_{i,t-1})$:

$$\begin{aligned} \textit{Risk taking}_{i,t} = {} & \rho\left(\mathbb{1}(\textit{Growth}_{i,t-1}) \times \mathbb{1}(\textit{Pre}_t) \times \mathbb{1}(\textit{Close}_{i,t-1})\right) + \delta_1\left(\mathbb{1}(\textit{Growth}_{i,t-1}) \times \mathbb{1}(\textit{Pre}_t)\right) \\ & + \delta_2\left(\mathbb{1}(\textit{Growth}_{i,t-1}) \times \mathbb{1}(\textit{Close}_{i,t-1})\right) + \delta_3\left(\mathbb{1}(\textit{Pre}_t) \times \mathbb{1}(\textit{Close}_{i,t-1})\right) \\ & + \beta_1\mathbb{1}(\textit{Growth}_{i,t-1}) + \beta_2\mathbb{1}(\textit{Close}_{i,t-1}) + \gamma\Gamma_{i,t-1} + \theta_t + \varepsilon_{i,t}. \end{aligned} \tag{5}$$

All variables are defined in (4), and standard errors are double-clustered by fund and time. The coefficient of interest is ρ and provides a formal test for the difference between the risk taking of disadvantaged growth funds that are close to versus distant from a rating cutoff. Note that $\rho > 0$ is similar to $\delta_{\text{close}} > \delta_{\text{distant}}$.

Table 7 reports the results and confirms our expectations. In Panel A columns (1) and (3) show the estimates for iVol and TE for funds which are close to a rating cutoff, and columns (2) and (4) for funds which are far away. When we measure risk taking by iVol, $\delta = 1.07$ for funds that are close to a rating threshold. This compares to $\delta = 0.78$ for funds that are far away. While the risk taking effect of disadvantaged growth funds is present for both type of funds, it is almost one-third larger for funds that are close to a rating cutoff. Consistent with this the coefficient on the triple interaction term $\rho = 0.29$ (column (1) in Panel B) confirms again that the higher risk taking of disadvantaged growth funds in the pre-June 2002 period is particularly pronounced for funds close to a rating cutoff. Moreover, the triple interaction regression informs us that the difference in the effect is statistically significant when we compare funds that are close to versus distant from a rating threshold.

Similarly, in the case of TE these numbers are $\delta = 1.01$ for funds that are close to a rating cutoff, and $\delta = 0.81$ for funds that are far away (columns (3) and (4) in Panel A). Furthermore, $rho = 0.22$ (column (3) in Panel B) is again statistically significant.

[Insert Table 7]

A potential concern in our DiD specification in (3) is that our results might be driven by changes in risk profiles of the underlying stocks rather than changes in funds' risk taking behaviors. Indeed, the average growth stock was more volatile relative to the average value stock prior to June 2002. A possible explanation of the differential change in volatilities of growth versus value stocks around 2002 might be that many risky growth stocks have vanished after the dot-com crash. Note that iVol and TE are only partially affected by changes in total volatility, but this does not fully resolve the concern.

One possibility is to control for the volatility of the relevant Russell style index associated with fund i in month t. However, this would only address differential changes in risk profiles of stocks across styles but not other style specific time-varying unobservables. Moreover, controlling for volatilities of Russell style indices comes with econometrics challenges. Since the six Russell style indices ($\{$growth, value$\} \times \{$large, mid-cap, small$\}$) are highly correlated, we worry about near-multicollinearity if we control for time fixed effects at the same time. But dropping time fixed effects in favor of controlling for the volatilities of the Russell style indices would be worrisome as time fixed effects capture all time-varying unobservables (volatility as well as others) that are common across all funds.

Note that our analysis of funds that are close to versus distant from a rating threshold allows us to exploit within-style differences. Thus, the funds that we contrast are exposed to the same style specific unobservables. This provides a perfect setting to control for time variations in any style specific characteristics (as for instance, but not limited to, risk profiles of growth versus or stocks), and allows us to cleanly analyze the risk taking behavior of funds. To be more specific we replace the time fixed effects by within style time effects in our triple interaction model in (5). This sharpens our analysis and allows us to control for any time-varying unobservables that are common to funds within a style (close to or distant from a rating cutoff), but these unobservables may differ across styles. For instance, this fully alleviate the concern that differential changes in risk profiles of growth versus value stocks could potentially drive our results.

Columns (2) and (4) of Panel B in Table 7 report the results of the triple interaction test with within style time fixed effects. ρ is significant across both risk measures. In the case of iVol it only slightly decreases from 0.29 to 0.21, suggesting that time-varying style specific unobservables do not drive much part of our baseline results. Similarly, in the case of TE ρ hardly changes from 0.22 to 0.21. That is, accounting for the style specific time-varying unobservables has no material effect on our results.

To sum up, while section 4.3 shows that the difference in risk taking between growth and value funds was higher prior to June 2002 when growth funds were at a disadvantage, in this section we document that this difference is significantly larger for funds that close to a rating threshold compared to funds that are far away. In other words, unintended risk taking incentives induced by benchmarks have a greater impact on disadvantaged funds if they have greater incentives to mitigate their disadvantages due to their proximity to a rating cutoff. Moreover, we use a triple interaction and study within style differences to sharpen our analysis and control for all time-varying style specific unobservables. We confirm that our risk taking mechanism of funds in response to a disadvantage is robust.

4.6 Contractual incentives

The compensation of portfolio managers crucially hinges on the AUM and the fee revenues funds generate (Ibert et al., 2018; Cen et al., 2023; Bai et al., 2023). Mutual fund advisory fees are typically based solely on a percentage of the fund's assets (Deli, 2002). Thus, we proxy the incentives of portfolio managers by those of the advisers. Specifically, we use Coles' incentive rate (*CIR*) (Coles et al., 2000) to measure fund manager incentives implied by the concavity of the advisory fee schedule (Massa and Patgiri, 2009).[8] The

[8] Massa and Patgiri (2009) find that mutual funds with high-incentive advisory contracts take greater risks, have lower survival rates, but earn higher risk-adjusted returns.

intuition is that mutual funds with linear (i.e., high-incentive) advisory contracts ($CIR = 0$) have a greater incentive to exert more effort to improve their Morningstar ratings and attract fund flows compared to funds with concave (i.e., relatively low-incentive) advisory contracts ($CIR < 0$) in which case the advisory fee rate decreases as the fund size increases.

We show the distribution of CIR in Figure 4. A large mass of about half of the funds is located at $CIR = 0$. Moreover, most funds with $CIR < 0$ have contracts that are not far from linear. Accordingly, we expect a relatively moderate difference in the behavior of fund managers with linear versus concave contracts.

We follow the approach in section 4.5 and examine the effect of contractual incentives on the risk-taking of disadvantaged growth funds in two steps. In the first step we re-run the DiD specification in (3) separately for the sub-samples of funds with linear respectively concave contracts. In the second step we run a combined regression with all funds and a triple interaction between $\mathbb{1}(\text{Growth}_{i,t-1})$, $\mathbb{1}(\text{Pre}_t)$ and the dummy $\mathbb{1}(\text{Linear}_{i,t-1})$ where the latter indicates whether a fund has a linear contract ($CIR = 0$):

$$\begin{aligned} \textit{Risk taking}_{i,t} = {} & \rho\left(\mathbb{1}(\textit{Growth}_{i,t-1}) \times \mathbb{1}(\textit{Pre}_t) \times \mathbb{1}(\textit{Linear}_{i,t-1})\right) + \delta_1\left(\mathbb{1}(\textit{Growth}_{i,t-1}) \times \mathbb{1}(\textit{Pre}_t)\right) \\ & + \delta_2\left(\mathbb{1}(\textit{Growth}_{i,t-1}) \times \mathbb{1}(\textit{Linear}_{i,t-1})\right) + \delta_3\left(\mathbb{1}(\textit{Pre}_t) \times \mathbb{1}(\textit{Linear}_{i,t-1})\right) \\ & + \beta_1 \mathbb{1}(\textit{Growth}_{i,t-1}) + \beta_2 \mathbb{1}(\textit{Linear}_{i,t-1}) + \gamma \Gamma_{i,t-1} + \theta_t + \varepsilon_{i,t} \end{aligned} \tag{6}$$

where all variables are defined in (4), and standard errors are double-clustered by fund and time.

Table 8 reports the results. Columns (1) and (3) in Panel A show the estimates of (3) for iVol and TE for funds with linear contracts, and columns (2) and (4) for funds with concave contracts. Among funds with linear contracts the iVol of growth relative to value funds was on average 1.21% higher pre-June 2002 (Panel A, column (1)). Consistent with our expectation the same figure decreases by more than half to 0.55% for the sub-sample of funds with concave contracts (Panel A, column (2)). The same pattern holds for TE with a $\delta = 1.31$ for funds with linear contracts versus $\delta = 0.51$ for funds with concave contracts.

Panel B in Table 8 shows that ρ in the triple interaction model is significant and positive for both risk measures. This confirms our expectation that the higher risk taking of growth funds in the pre-June 2002 period is particularly prevalent funds that have a higher incentive to boost their Morningstar rating and maximize fund flows.

[Insert Table 8]

As in the analysis in section 4.5, the triple interaction setting with within style time fixed effects allows us to control for all time-varying style specific characteristics (such as risk profiles of growth or value stocks among others) when we study the risk taking effect across growth funds which differ in their incentives to mitigate their disadvantage prior to June 2002. For iVol adn TE ρ only slightly decreases from 0.62 to 0.55 respectively from 0.77 to 0.68 when we replace time fixed effects by within style time fixed effects in the specification (6). This suggests that the style specific unobservables explain very little (if anything) of our baseline results. In other words, our risk taking effect appears to be in response to the disadvantage that growth funds were facing prior to June 2002, and we rule out alternative channels.

In summary, unintended risk taking incentives induced by benchmarks have a greater impact on funds if they have greater incentives to generate fund flows as measured by the curvature of advisory fee schedules.

4.7 Efficacy of risk taking by growth funds

It is an important question whether the increase in risk taking by disadvantaged growth funds is effective to mitigate the inherent disadvantage to better compete with advantaged value funds in the pre-June 2002 era. If the answer is yes and fund performance improves, then this provides solid evidence that the increase in risk taking is a proxy for managers exerting more effort. The difficulty is that we cannot directly measure the counter-factual performance of disadvantaged growth funds had they not increased their risk prior to June 2002.

However, we can compare the performance of growth funds which differ in the amount of risk taking due to differences in incentives, i.e., being close to versus far from Morningstar cutoffs, or using linear versus concave contracts. Recall that disadvantaged growth funds that are (i) closer to a Morningstar cutoff or (ii) use linear contracts are taking relatively more risks than their counter-parts which are far form Morningstar thresholds or use concave contracts (sections 4.5 and 4.6). Accordingly, an outperformance (underperformance) of these disadvantaged growth funds which are close to a cutoff or use linear contracts relative to their counter-parts which are far from cutoffs or use concave contracts informs us that the increase in risk taking is efficient (counter-productive).

Table 9 compares the Morningstar ratings of growth versus value funds close to versus far from a rating cutoff (Panel A), and funds with linear versus concave contracts (Panel B) for both pre- and post-June 2002. In the pre-June 2002 period growth funds that were close to a rating cutoff have star ratings that are 0.12 points lower than equivalent value funds. However, this difference is almost twice, i.e., 0.23 between growth and value funds that are far from a rating cutoff. Moreover, we confirm that the difference between the 0.12 and the 0.23 points is significant. This pattern is even more striking for funds with linear versus concave contracts. Growth funds with linear contracts receive statistically indistinguishable ratings (the point estimate is only 0.08 points lower) than comparable value funds. In contrast, the ratings of growth funds with concave contracts are on average 0.36 points lower than the ratings of equivalent value funds. Moreover, this difference of 0.36 points as well as the difference between the 0.08 and 0.36 points are statistically significant and economically meaningful. As a sanity check we further confirm that after June 2002 there is no difference across growth and value funds, either close or far from a rating cutoff, or with linear or concave contracts.

To sum up, we provide evidence that the additional risk taking of growth funds, which have more incentives to mitigate their disadvantage, leads to a significant improvement in the Morningstar rating. Accordingly, increasing risk taking appears to be an effective action for fund managers to tackle the inherent disadvantage. This is consistent with the interpretation that risk taking is a proxy for exerting effort.

5 Extension to Socially Responsible Investing (SRI)

In this section, we demonstrate that our results can be generalized beyond the context of value versus growth funds, and is more broadly valid whenever benchmarks are relevant for fund flows and put some funds at a disadvantage. To this end, we extend our analysis to the context of socially responsible investing (SRI).

5.1 Disadvantage of morally constrained versus unconstrained funds

Hong and Kacperczyk (2009) find that stocks in alcohol, tobacco, and gaming industries – so called sin stocks – have positive risk-adjusted returns as morally responsible investors stay way from investing in these stocks. We use divesting from sin stocks to measure the extent of SRI.

We argue that more socially responsible funds will avoid investing in sin stocks due to moral constraints, explicit or otherwise. As a result, these morally constrained funds will have a disadvantage competing against unconstrained funds if they were to be evaluated against each other. To measure the extent of the disadvantage of socially responsible investing, we use Morningstar ratings where funds are ranked within respective size-by-value categories. We note that for star ratings, morally constrained funds will be evaluated against unconstrained funds without any regard to funds' efforts in SRI.

Figure 5 plots the time-series of Morningstar ratings averaged each month across morally constrained respectively unconstrained funds from June 2002 to December 2022. It is striking that unconstrained funds have received higher ratings than morally constrained funds in almost every month. Furthermore, the outperformance of unconstrained funds has become stronger in recent years with the rise of ESG investing, which may be the cause of additional underpricing of sin stocks.

In Table 11 we further report the average ratings of morally constrained (3.05) versus unconstrained funds (3.12). The difference of 0.07 points is statistically significant. To put our estimates into perspective we note that our numbers are in line with Lo and Zhang (2023) who estimate in the Treynor and Black (1973) framework that the cost of avoiding or divesting sin stocks ranges from 0.6% to 3.3% in forgone alpha per year.

Finally, we would like to emphasize that the true extent of the disadvantage that socially responsible funds are facing may be larger than what we estimate. In equilibrium, morally constrained funds should anticipate their disadvantage, and in response, increase their risk taking in an attempt to be more competitive with unconstrained funds. Indeed, we show this in the remainder of this section. This additional risk taking may lead (in the data) to a smaller underperformance of morally constrained funds.

Next, we investigate whether morally constrained funds increase risk taking in response to their disadvantage as they have to compete against unconstrained funds in Morningstar ratings. Figures 6 and 7 plot the time-series of monthly iVol respectively TE for morally constrained and unconstrained funds. It is obvious that morally constrained funds are on average riskier than unconstrained funds. Interestingly, this is true only after 2009, which coincides with the publication date of Hong and Kacperczyk (2009). A possible explanation is that the positive alpha of sin stocks was not well-known prior to the publication of Hong and Kacperczyk (2009), and thus, morally constrained funds were unaware that they were at a disadvantage (or at least the extent of it) when competing with unconstrained funds. It is reasonable to believe that after 2009 constrained funds learned about their disadvantage, and in response they started to increase risk taking.

Formally, we estimate the following linear regression model:

$$\textit{Risk taking}_{i,t} = \beta \mathbb{1}(\textit{Morally constrained}_{i,t-1}) + \gamma \Gamma_{i,t-1} + \theta_{i,t-1} + \varepsilon_{i,t}. \tag{7}$$

$\mathbb{1}(\textit{Morally constrained}_{i,t-1})$ is an indicator variable that takes a value of one if fund i underweights sin stocks at time $t-1$, meaning its weight in sin stocks is smaller than that of its benchmark index, and zero otherwise. Sin stocks are defined as firms that operate in the alcohol, tobacco, and gambling industries (Hong and Kacperczyk, 2009). $\theta_{i,t-1}$ denotes lagged category $\times$ time fixed effects, controlling for style specific time-varying unobservables. All other variables are defined in (3). Our sample spans the period from June 2002 to December 2022. Standard errors are double-clustered by fund and time

We present the results in Table 12. The coefficient of interest is β, which measures the additional risk taking of morally constrained funds. β is statistically significant across all specifications. β is 0.25 in the case of iVol, which means that morally constrained funds chose an iVol that exceeds the iVol of unconstrained funds by 0.25%. This is a meaningful effect considering the median iVol of 3.21% for all fund-month observations. Similarly, the TE of morally constrained funds is 0.29% larger than the TE of an equivalent unconstrained fund. This compares to the median TE of 3.58%. Adding fund level controls to our estimation leaves our results unchanged.

Overall, our results suggest that benchmarks may distort mutual funds' behavior, forcing disadvantaged socially responsible funds to take more risks, presumably to catch up with unconstrained peer funds.

[Insert Table 12]

6 Conclusion

We find that growth funds are at a disadvantage compared to value funds due to the value premium and Morningstar's pre-2002 rating methodology. This disadvantage incentivizes growth fund managers to increase risk-taking behavior to improve their ratings and attract investor flows. Our analysis reveals that this increased risk-taking is effective in mitigating the disadvantage and is more pronounced for funds with stronger incentives to improve their ratings. These findings highlight the significant impact of rating methodologies on fund manager behavior and the importance of well-designed systems to avoid unintended consequences. Finally, we show the recent relevance of our findings in the context of socially responsible investing. Specifically, we document that morally constrained funds are at a disadvantage when competing with unconstrained funds in Morningstar ratings. In response the morally constrained funds increase risk taking compared to the unconstrained funds.

References

Ammann, Manuel, Alexander Cochardt, Lauren Cohen, and Stephan Heller, 2022, Hidden alpha, Technical report, Working paper, University of St. Gallen and Harvard Business School.

Bai, John Jianqiu, Linlin Ma, Kevin Mullally, and Yuehua Tang, 2023, Are mutual fund managers paid for performance? evidence from us administrative earnings data, *Evidence from US Administrative Earnings Data (October 10, 2023)* .

Basak, Suleyman, Anna Pavlova, and Alex Shapiro, 2007, Optimal asset allocation and risk shifting in money management, *Review of Financial Studies* 20, 1583–1621.

Basak, Suleyman, Anna Pavlova, and Alex Shapiro, 2008, Offsetting the implicit incentives: Benefits of benchmarking in money management, *Journal of Banking and Finance* 32, 1882–1993.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2022a, what do mutual fund investors really care about?, *Review of Financial Studies* 35, 1723–1774.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2022b, ratings-driven demand and systematic price fluctuations, *Review of Financial Studies* 35, 2790–2838.

Berk, Jonathan B, and Jules H van Binsbergen, 2015, Measuring skill in the mutual fund industry, *Journal of financial economics* 118, 1–20.

Brown, Keith C., W. V. Harlow, and Laura T. Starks, 1996, Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry, *Journal of Finance* 51, 85–110.

Busse, Jeffrey A., 2001, Another look at mutual fund tournaments, *The Journal of Financial and Quantitative Analysis* 36, 53–73.

Cen, Xiao, Winston Wei Dou, Leonid Kogan, and Wei Wu, 2023, Fund flows and income risk of fund managers, Technical report, National Bureau of Economic Research.

Chen, Hsiu-lang, and George G. Pennacchi, 2009, Does prior performance affect a mutual fund's choice of risk? theory and further empirical evidence, *Journal of Financial and Quantitative Analysis* 44, 745–775.

Chen, Huaizhi, Lauren Cohen, and Umit G. Gurun, 2021, Don't take their word for it: The misclassification of bond mutual funds, *Journal of Finance* 76, 1699–1730.

Chen, Huaizhi, Richard Evans, and Yang Sun, 2024, Self-declared benchmarks and fund manager intent: "cheating" or competing?, *Journal of Financial Economics* forthcoming.

Chevalier, Judith, and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105, 1167–1200.

Cohen, Randolph B., Christopher Polk, and Tuomo Vuolteenaho, 2003, The value spread, *Journal of Finance* 58, 609–642.

Coles, Jeffrey L, Jose Suay, and Denise Woodbury, 2000, Fund advisor compensation in closed-end funds, *The Journal of Finance* 55, 1385–1414.

Cremers, K. J. Martijn, Jon A. Fulkerson, and Timothy B. Riley, 2022, Benchmark discrepancies and mutual fund performance evaluation, *Journal of Financial and Quantitative Analysis* 57, 543–571.

Cullen, Grant, Dominic Gasbarro, Gary S. Monroe, and J. Kenton Zumwalt, 2012, Changes to mutual fund risk: Intentional or mean reverting?, *Journal of Banking and Finance* 36, 112–120.

Del Guercio, Diane, and Paula A. Tkac, 2002, The determinants of the flow of funds of managed portfolios: Mutual funds vs. pension funds, *Journal of Financial and Quantitative Analysis* 37, 523–557.

Del Guercio, Diane, and Paula A. Tkac, 2008, Star power: The effect of morningstar ratings on mutual fund flow, *The Journal of Financial and Quantitative Analysis* 43, 907–936.

Deli, Daniel N, 2002, Mutual fund advisory contracts: An empirical investigation, *The journal of finance* 57, 109–133.

Elton, Edwin J, Martin J Gruber, and Christopher R Blake, 2011, Holdings data, security returns, and the selection of superior mutual funds, *Journal of Financial and Quantitative Analysis* 46, 341–367.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, 2014, The performance of separate accounts and collective investment trusts, *Review of Finance* 18, 1717–1742.

Evans, Richard, Juan-Pedro Gomez, Linlin Ma, and Yuehua Tang, 2023, Peer versus pure benchmarks in the compensation of mutual fund managers, *Journal of Financial and Quantitative Analysis* forthcoming.

Evans, Richard, and Yang Sun, 2021a, Models or stars: The role of asset pricing models and heuristics in investor risk adjustment, *The Review of Financial Studies* 34, 67–107.

Evans, Richard B., and Yang Sun, 2021b, Models or stars: The role of asset pricing models and heuristics in investor risk adjustment, *Review of Financial Studies* 34, 67–107.

Fama, Eugene, and Kenneth French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Han, Xiao, Nikolai Roussanov, and Hongxun Ruan, 2021, Mutual Fund Risk Shifting and Risk Anomalies, Working paper.

Hartzmark, Samuel M, and Abigail B Sussman, 2019, Do investors value sustainability? a natural experiment examining ranking and fund flows, *The Journal of Finance* 74, 2789–2837.

Hong, Harrison, and Marcin Kacperczyk, 2009, The price of sin: The effects of social norms on markets, *Journal of financial economics* 93, 15–36.

Huang, Jennifer, Clemens Sialm, and Hanjiang Zhang, 2011, Risk shifting and mutual fund performance, *Review of Financial Studies* 24, 2575–2616.

Ibert, Markus, Ron Kaniel, Stijn Van Nieuwerburgh, and Roine Vestman, 2018, Are mutual fund managers paid for investment skill?, *The Review of Financial Studies* 31, 715–772.

Kim, Min S., 2019, Changes in Mutual Fund Flows and Managerial Incentives, Working paper, Boston University.

Kim, Sanghyun (Hugh), 2022, Do Mutual Funds Manipulate Star Ratings? Evidence from Portfolio Pumping, Working paper, Wilfrid Laurier University.

Krueger, Philipp, Zacharias Sautner, and Laura T Starks, 2020, The importance of climate risks for institutional investors, *The Review of financial studies* 33, 1067–1111.

Lee, Jung Hoon, Charles Trzcinka, and Shyam Venkatesan, 2019, Do portfolio manager contracts contract portfolio management?, *Journal of Finance* 74, 2543–2577.

Lo, Andrew W, and Ruixun Zhang, 2023, Quantifying the impact of impact investing, *Management Science* .

Ma, Linlin, and Yuehua Tang, 2019, Portfolio manager ownership and mutual fund risk taking, *Management Science* 65, 5449–5956.

Ma, Linlin, Yuehua Tang, and Juan-Pedro Gomez, 2019, Portfolio manager compensation in the u.s. mutual fund industry, *Journal of Finance* 74, 587–638.

Massa, Massimo, and Rajdeep Patgiri, 2009, Incentives and mutual fund performance: higher performance or just higher risk taking?, *The Review of Financial Studies* 22, 1777–1815.

Mullally, Kevin, and Andrea Rossi, 2024, Moving the goalposts? mutual fund benchmark changes and relative performance manipulation, Working paper, University of Arizona.

Pástor, L'uboš, Robert F Stambaugh, and Lucian A Taylor, 2015, Scale and skill in active management, *Journal of Financial Economics* 116, 23–45.

Reuter, Jonathan, and Eric Zitzewitz, 2021, How much does size erode mutual fund performance? a regression discontinuity approach, *Review of Finance* 25, 1395–1432.

Schwarz, Christopher G., 2012, Mutual fund tournaments: The sorting bias and new evidence, *Review of Financial Studies* 25, 913–936.

Sensoy, Berk A., 2009, Performance evaluation and self-designated benchmark indexes in the mutual fund industry, *Journal of Financial Economics* 92, 25–39.

Sharpe, William F., 1998, Morningstar's risk-adjusted ratings, *Financial Analysts Journal* 54, 21–33.

Spiegel, Matthew, and Hong Zhang, 2013, Mutual fund risk and market share-adjusted fund flows, *Journal of Financial Economics* 108, 506–528.

Treynor, Jack L, and Fischer Black, 1973, How to use security analysis to improve portfolio selection, *The journal of business* 46, 66–86.

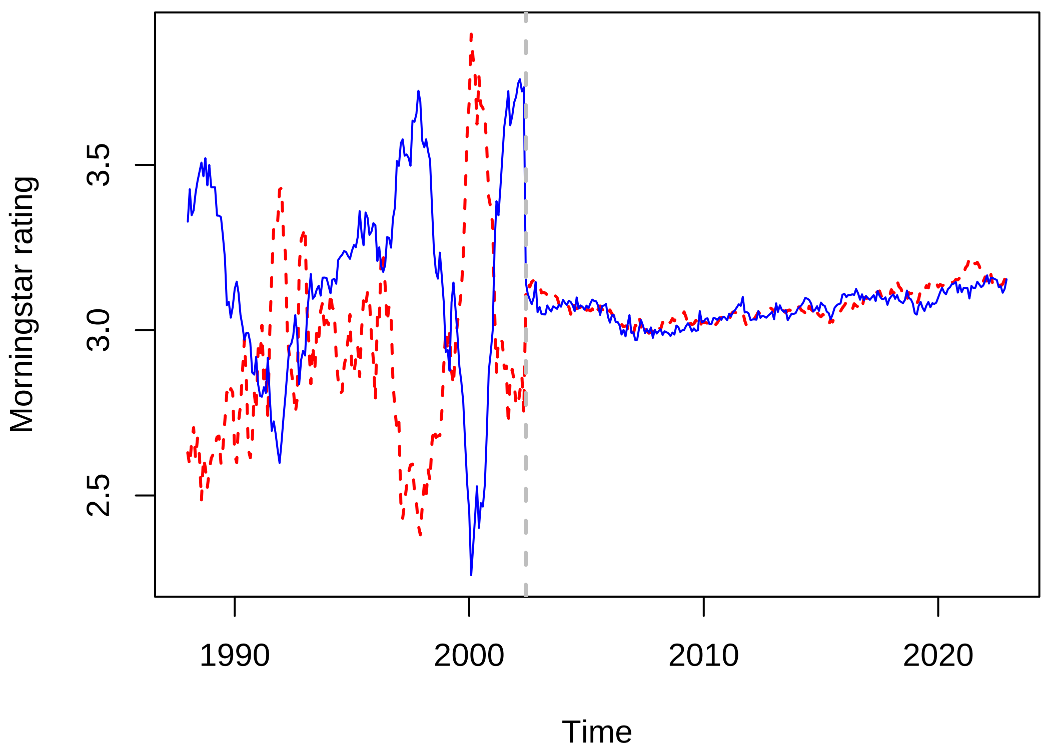

Figure 1: The average Morningstar rating of growth funds (red dashed line) compared to that of value funds (blue solid line).

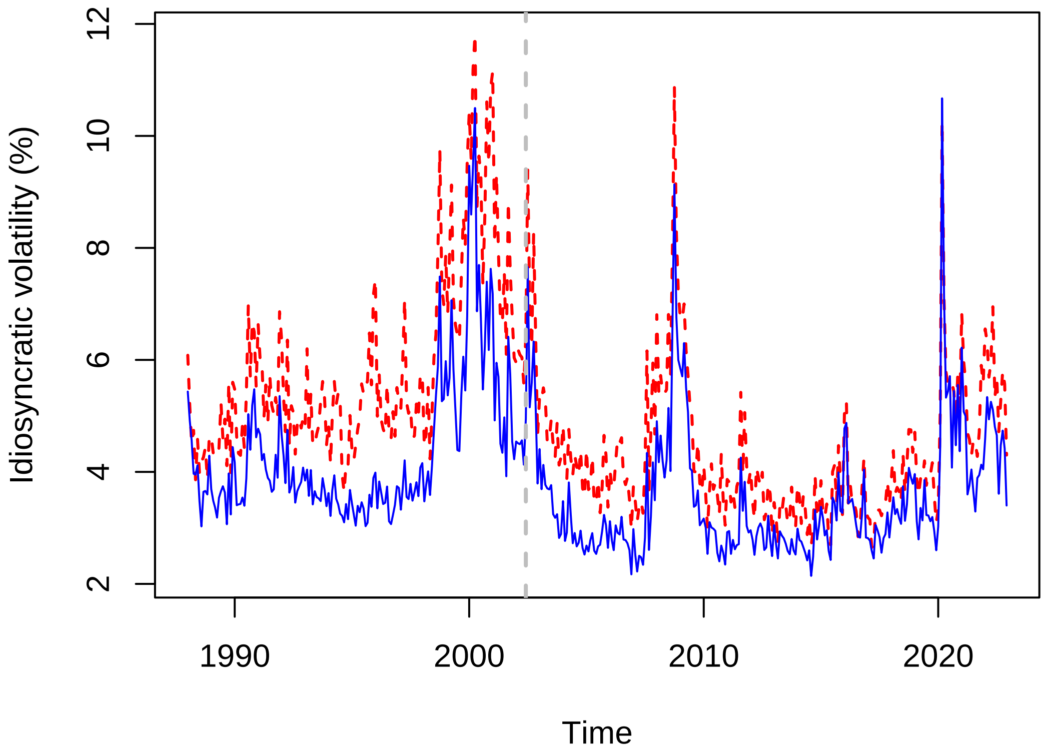

Figure 2: The average idiosyncratic volatility of growth funds (red dashed line) compared to that of value funds (blue solid line).

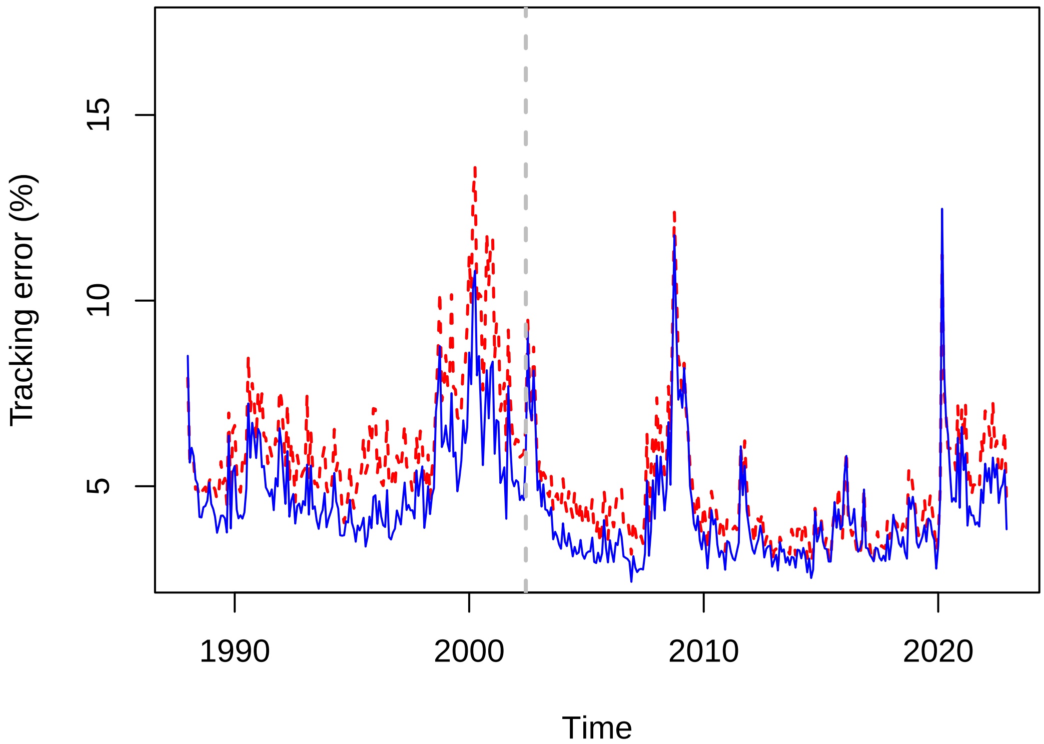

Figure 3: The average tracking error of growth funds (red dashed line) compared to that of value funds (blue solid line).

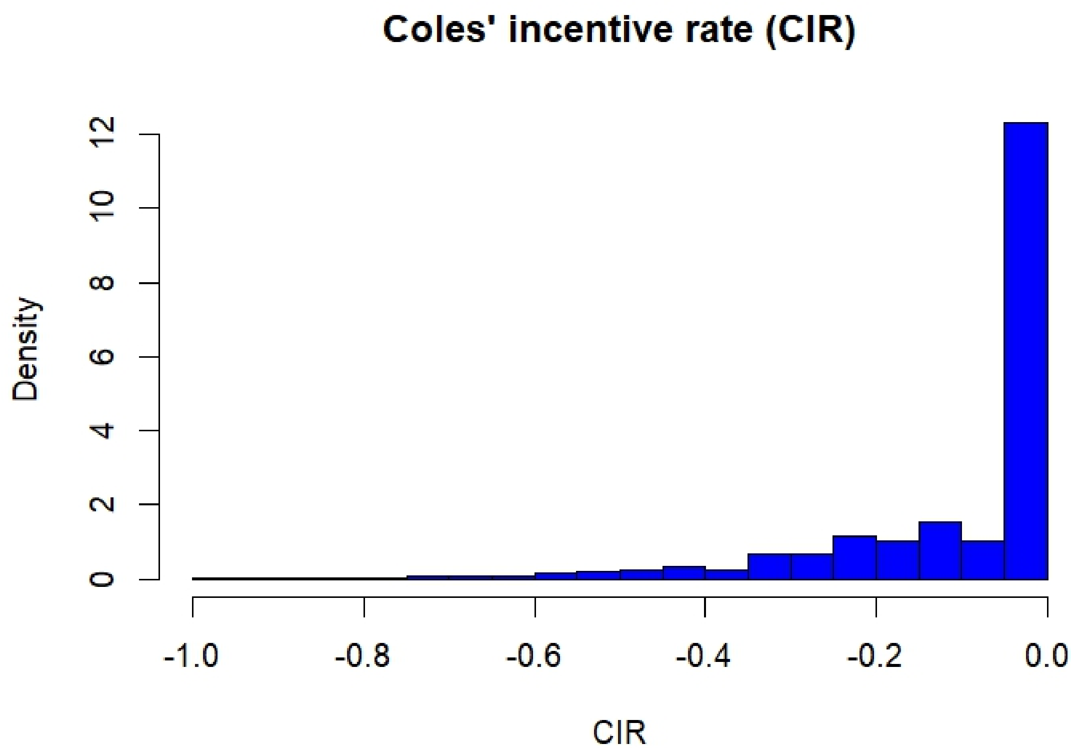

Figure 4: Distribution Cole's contract curvature measure of funds, CIR.

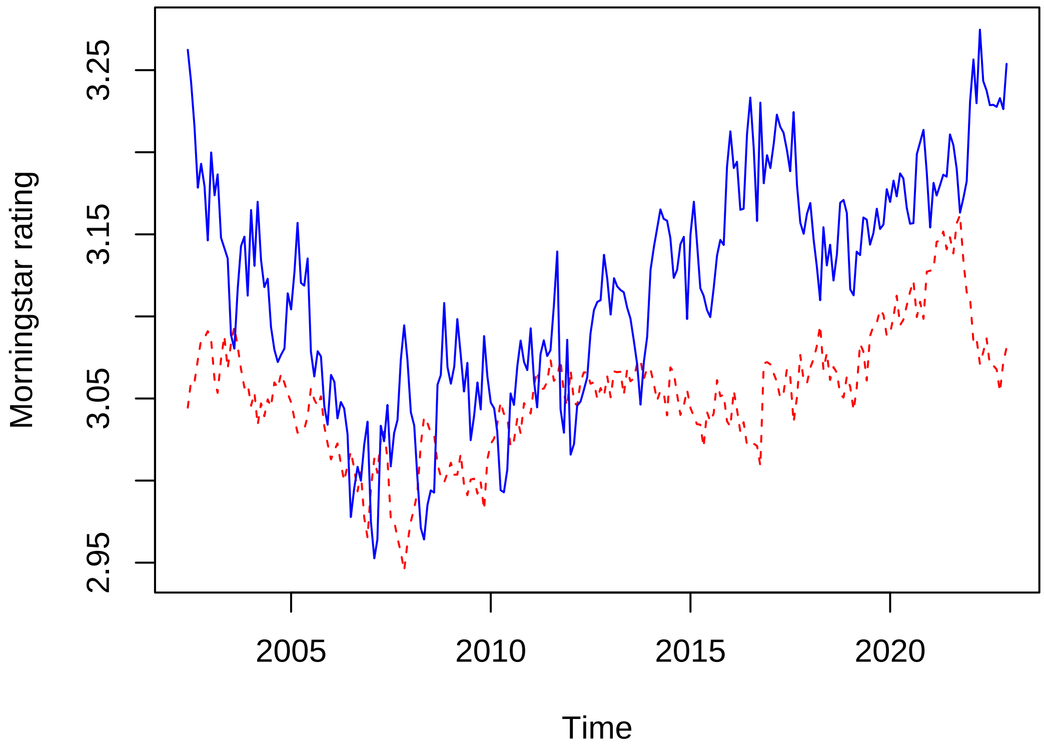

Figure 5: The average Morningstar rating of morally constrained funds (red dashed line) compared to that of unconstrained funds (blue solid line).

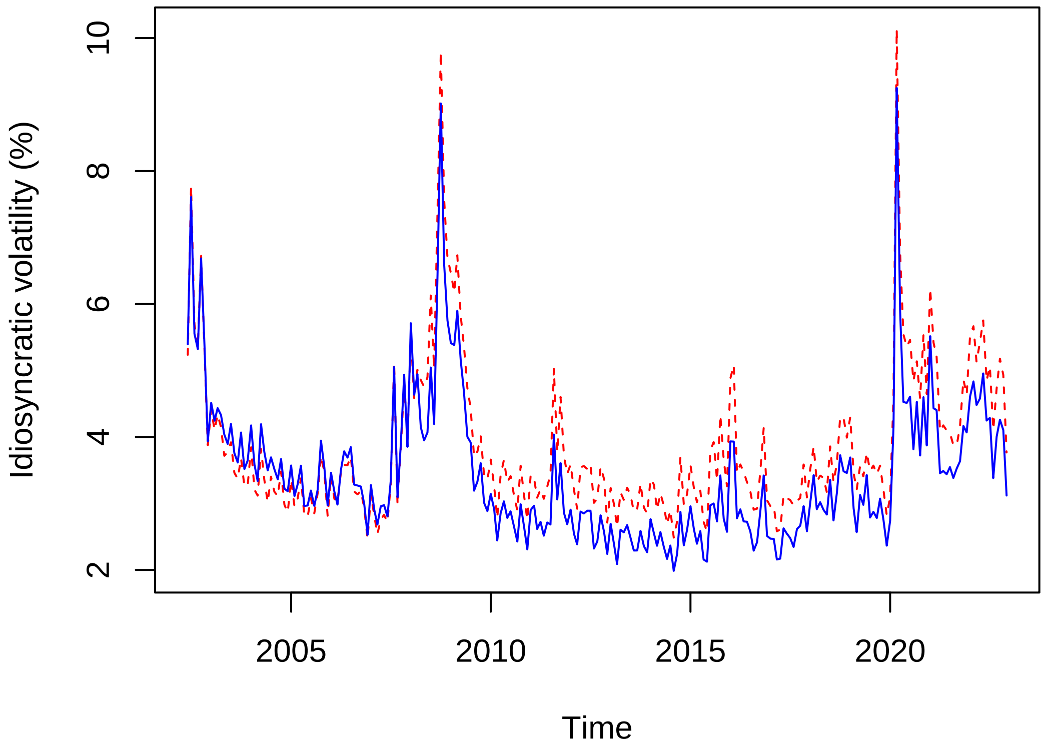

Figure 6: The average idiosynractic volatility of morally constrained funds (red dashed line) compared to that of unconstrained funds (blue solid line).

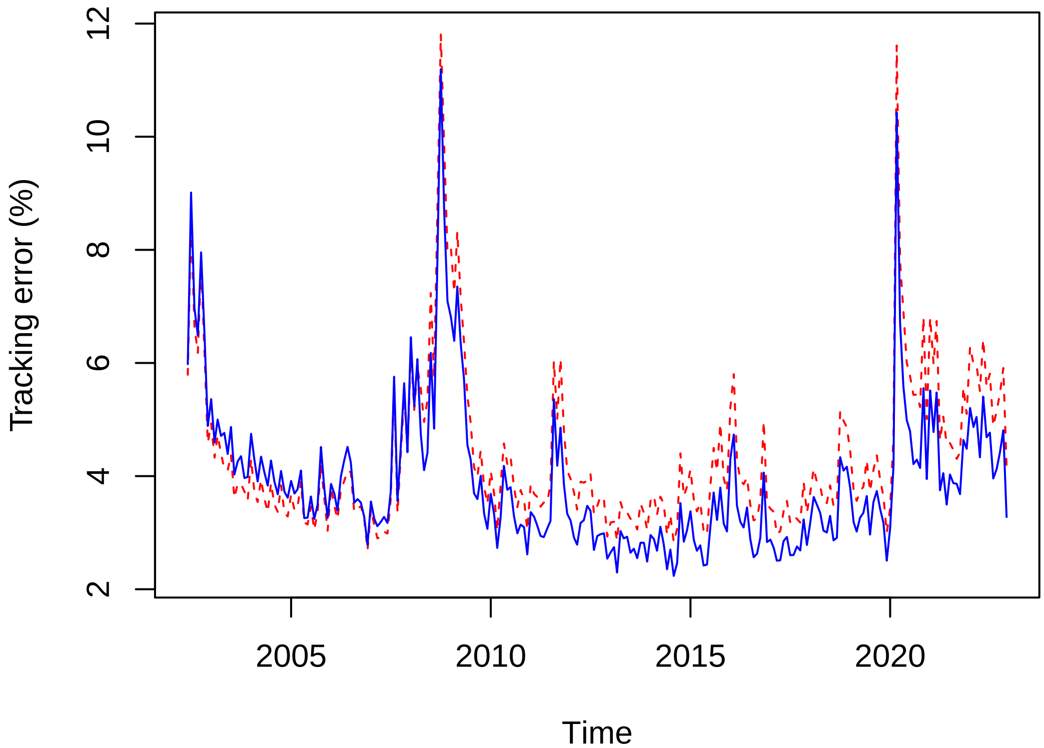

Figure 7: The average tracking error of morally constrained funds (red dashed line) compared to that of unconstrained funds (blue solid line).

Table 1: Summary statistics

Panel A of this table presents the summary statistics for the key variables used in our analyses of growth vs. value funds. All variables are indexed by fund i and time t (in months). $\mathbb{1}(Growth_{i,t-1})$ is an indicator variable that takes a value of one if fund i belongs to one of the growth categories (Large Growth, Mid-cap Growth, or Small Growth) at time $t-1$, and zero if fund i belongs to one of the value categories (Large Value, Mid-cap Value, or Small Value) at time $t-1$. $\mathbb{1}(Pre_t)$ is an indicator variable that takes a value of one if time t is prior to the refinement of Morningstar peer groups in June 2002, and zero otherwise. $Morningstar\ ratings_{i,t}$ is fund i's Morningstar rating at time t. $Idiosyncratic\ volatility_{i,t}$ is computed as the standard deviation of the residuals from the regression of fund i's daily returns (in excess of the T-bill rate) during month t on the Fama-French 3-factor daily returns (Fama and French, 1993), annualized and reported as a percentage. $Tracking\ error_{i,t}$ is computed as the standard deviation of fund i's daily returns in excess of the benchmark index returns during month t, annualized and reported as a percentage. $HML_{t-35,t}$ is the cumulative return on the value factor from the Fama-French 3-factor model (Fama and French, 1993) over the past 36 months. $\log(B/M)_{t-1}^{\mathrm{Hi-Lo}}$ is the value spread, computed as the log difference of the book-to-market ratios of the value portfolio (top book-to-market decile) and the growth portfolio (bottom book-to-market decile) at the end of the previous June (Cohen et al., 2003). $\mathbb{1}(Close_{i,t-1})$ is an indicator variable that takes a value of one if fund i's return at time $t-36$ falls within the bottom decile relative to its peers with the same rating at time $t-1$, and zero otherwise. Coles' incentive rate (CIR) is defined as the difference between the marginal advisory fee rates of the highest and lowest total net asset (TNA) breakpoints, divided by the effective advisory fee rate (Coles et al., 2000). $\mathbb{1}(\mathrm{Linear}_{i,t-1})$ is an indicator variable that takes a value of one if fund i's advisory fee rate at time $t-1$ is linear with respect to its TNA ($CIR = 0$), and zero otherwise. $\mathbb{1}(Close_{i,t-1})$ takes the value of one if fund i is moving closer to a rating threshold for a rating update at time t, as measured at time $t-1$, and zero otherwise (Kim, 2022). Our sample spans the period from 1988 to 2022.

Panel A

Statistic	N	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
$\mathbb{1}$(Growth)	395, 165	0.60	0.49	0	1	1
$\mathbb{1}$(Pre)	395, 165	0.20	0.40	0	0	0
Morningstar rating	395, 165	3.07	1.05	2.00	3.00	4.00
Idiosyncratic volatility	395, 165	4.38	2.54	2.67	3.72	5.29
Tracking error	395, 165	4.85	2.92	2.92	4.09	5.86
HML	395, 165	0.03	0.23	−0.10	−0.01	0.16
$\log(\mathrm{B/M})^{\mathrm{Hi-Lo}}$	395, 165	2.60	0.44	2.26	2.53	2.91
Coles' incentive rate (CIR)	349, 105	−0.10	0.15	−0.16	0.00	0.00
$\mathbb{1}$(Linear)	349, 105	0.57	0.50	0	1	1
$\mathbb{1}$(Close)	385, 652	0.21	0.40	0.00	0.00	0.00

Table 1–*Continued*

Panel B of this table presents the summary statistics for the variables used in our analyses of morally constrained vs. unconstrained funds. All variables are indexed by fund i and time t (in months). $\text{Weight}^{\text{Fund}}_{\text{Sin,i,t-1}}$ and $\text{Weight}^{\text{Index}}_{\text{Sin,i,t-1}}$ represent fund i's weight and its benchmark index weight in sin stocks, respectively. Sin stocks are defined as those of firms operating in the alcohol, tobacco, and gambling industries (Hong and Kacperczyk, 2009). $\mathbb{1}(\text{Morally constrained}_{i,t-1})$ is an indicator variable that takes a value of one if fund i's weight in sin stocks is smaller than that of its benchmark index ($\text{Weight}^{\text{Fund}}_{\text{Sin}} - \text{Weight}^{\text{Index}}_{\text{Sin}} < 0$), and zero otherwise. The remaining variables are defined as in Panel A of this table. Our sample spans the period from June 2002 to December 2022.

Panel B

Statistic	N	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
$\text{Weight}^{\text{Fund}}_{\text{Sin}} - \text{Weight}^{\text{Index}}_{\text{Sin}}$	471,184	−0.49	2.49	−1.69	−0.73	0.34
$\mathbb{1}$(Morally constrained)	471,184	0.67	0.47	0	1	1
Morningstar rating	471,184	3.08	1.04	2	3	4
Idiosyncratic volatility	471,184	3.70	2.24	2.24	3.21	4.56
Tracking error	471,184	4.19	2.66	2.49	3.58	5.12

Table 2: Disadvantage of growth funds compared to value funds

This table compares the average Morningstar ratings of growth funds and value funds, separately for the period before the refinement of Morningstar peer groups (January 1988 to May 2002) and the period after the refinement (June 2002 to December 2022). Standard errors are double-clustered by fund and time, and t-statistics are reported in parentheses, with statistical significance at the 10%, 5%, and 1% levels indicated by *, **, and ***, respectively.

Morningstar rating					
January 1988 – May 2002			June 2002 – December 2022		
Growth	Value	Growth–Value	Growth	Value	Growth–Value
(1)	(2)	(3)	(4)	(5)	(6)
2.98	3.18	-0.20^{***}	3.07	3.07	0.003
		(-3.74)			(0.11)

Table 3: Time-varying disadvantage of growth funds compared to value funds

This table presents the results of the following linear regression model:

$$\begin{aligned} \textit{Morningstar rating}_{i,t} = &\ \delta \text{HML}_{t-35,t} \times \mathbb{1}\left(\text{Growth}_{i,t-1}\right) \\ &+ \beta_1 \text{HML}_{t-35,t} + \beta_2 \mathbb{1}\left(\text{Growth}_{i,t-1}\right) + \theta_i + \varepsilon_{i,t}, \end{aligned}$$

where i indexes mutual funds, and t indexes time (in months). $\textit{Morningstar rating}_{i,t}$ represents fund i's Morningstar rating at time t. $\mathbb{1}(\textit{Growth}_{i,t-1})$ is an indicator variable that takes a value of one if fund i belongs to one of the growth categories (Large Growth, Mid-cap Growth, or Small Growth) at time $t-1$, and zero if fund i belongs to one of the value categories (Large Value, Mid-cap Value, or Small Value) at time $t-1$. $\textit{HML}_{t-35,t}$ is the cumulative return on the value factor (HML) from the Fama-French 3-factor model (Fama and French, 1993) over the past 36 months. θ_i denotes fund fixed effects. The regressions are run separately for the period before the refinement of Morningstar peer groups (January 1988 to May 2002) and the period after the refinement (June 2002 to December 2022). Standard errors are double-clustered by fund and time, and t-statistics are reported in parentheses, with statistical significance at the 10%, 5%, and 1% levels indicated by *, **, and ***, respectively.

	Morningstar rating			
	January 1988 – May 2002		June 2002 – December 2022	
	(1)	(2)	(3)	(4)
$\text{HML}_{t-35,t} \times \mathbb{1}(\text{Growth})$	-3.20^{***}	-3.07^{***}	0.01	0.01
	(−27.96)	(−26.89)	(0.13)	(0.14)
$\text{HML}_{t-35,t}$	1.75^{***}	1.64^{***}	−0.04	0.30^{***}
	(22.82)	(21.82)	(−0.61)	(4.51)
$\mathbb{1}(\text{Growth})$	0.08	0.05	0.003	0.03
	(1.49)	(0.13)	(0.12)	(0.23)
Constant	3.02^{***}		3.07^{***}	
	(83.46)		(133.10)	
Fund fixed effects		Yes		Yes
Observations	79,557	79,557	315,608	315,608
Adjusted R^2	0.07	0.52	0.0001	0.45

Table 4: Risk-taking of growth versus value funds

This table presents the results of the following linear regression model:

$$Risk\ taking_{i,t} = \delta\left(\mathbb{1}(Growth_{i,t-1}) \times \mathbb{1}(Pre_t)\right) + \beta\mathbb{1}(Growth_{i,t-1}) + \gamma\Gamma_{i,t-1} + \theta_t + \varepsilon_{i,t}$$

where i indexes mutual funds and t indexes time (in months). $Risk\ taking_{i,t}$ represents fund i's idiosyncratic volatility or tracking error during time t, annualized and reported as a percentage. $\mathbb{1}(Growth_{i,t-1})$ is an indicator variable that takes a value of one if fund i belongs to one of the growth categories (Large Growth, Mid-cap Growth, or Small Growth) at time $t-1$, and zero if fund i belongs to one of the value categories (Large Value, Mid-cap Value, or Small Value) at time $t-1$. $\mathbb{1}(Pre_t)$ is an indicator variable that takes a value of one if time t is prior to the refinement of Morningstar peer groups in June 2002, and zero otherwise. $\Gamma_{i,t-1}$ is a vector of lagged fund characteristics, including the natural logarithm of total net assets (TNA), expense ratio, turnover ratio, and percentile rank of year-to-date (YTD) return. θ_t denotes time fixed effects. Our sample spans the period from 1988 to 2022. Standard errors are double-clustered by fund and time, and t-statistics are reported in parentheses, with statistical significance at the 10%, 5%, and 1% levels indicated by *, **, and ***, respectively.

	Idiosyncratic volatility		Tracking error	
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Growth) × $\mathbb{1}$(Pre)	0.93***	0.92***	0.94***	0.95***
	(7.18)	(7.53)	(6.34)	(6.87)
$\mathbb{1}$(Growth)	0.83***	0.68***	0.58***	0.40***
	(13.84)	(12.24)	(7.75)	(5.71)
log(TNA)		−0.08***		−0.11***
		(−6.76)		(−7.27)
Expense ratio		1.12***		1.52***
		(16.32)		(17.88)
Turnover ratio		0.16***		0.10**
		(3.64)		(2.03)
Rank(YTD return)		0.02		0.11*
		(0.36)		(1.82)
Time fixed effects	Yes	Yes	Yes	Yes
Observations	395,165	394,963	395,165	394,963
Adjusted R^2	0.40	0.45	0.36	0.43

Table 5: Increased risk-taking by growth funds over shorter time windows

This table presents the results of the following linear regression model:

$$Risk\ taking_{i,t} = \delta\left(\mathbb{1}(Growth_{i,t-1}) \times \mathbb{1}(Pre_t)\right) + \beta\mathbb{1}(Growth_{i,t-1}) + \gamma\Gamma_{i,t-1} + \theta_t + \varepsilon_{i,t}$$

where i indexes mutual funds and t indexes time (in months). $Risk\ taking_{i,t}$ represents fund i's idiosyncratic volatility or tracking error during time t, annualized and reported as a percentage. $\mathbb{1}(Growth_{i,t-1})$ is an indicator variable that takes a value of one if fund i belongs to one of the growth categories (Large Growth, Mid-cap Growth, or Small Growth) at time $t-1$, and zero if fund i belongs to one of the value categories (Large Value, Mid-cap Value, or Small Value) at time $t-1$. $\mathbb{1}(Pre_t)$ is an indicator variable that takes a value of one if time t is prior to the refinement of Morningstar peer groups in June 2002, and zero otherwise. $\Gamma_{i,t-1}$ is a vector of lagged fund characteristics, including the natural logarithm of total net assets (TNA), expense ratio, turnover ratio, and percentile rank of year-to-date (YTD) return. θ_t denotes time fixed effects. The regressions are run separately for the periods from 2001 to 2003, from 2000 to 2004, and from 2000 to 2004, excluding 2002. Standard errors are double-clustered by fund and time, and t-statistics are reported in parentheses, with statistical significance at the 10%, 5%, and 1% levels indicated by *, **, and ***, respectively.

	Idiosyncratic volatility			Tracking error		
	01–03	00–04	00–01, 03–04	01–03	00–04	00–01, 03–04
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Growth) × $\mathbb{1}$(Pre)	1.00***	1.05***	1.26***	1.22***	1.43***	1.58***
	(4.78)	(6.03)	(6.68)	(6.00)	(7.55)	(8.11)
$\mathbb{1}$(Growth)	1.01***	0.95***	0.90***	0.36**	0.49***	0.57***
	(8.41)	(10.30)	(10.77)	(2.72)	(4.31)	(5.54)
log(TNA)	−0.14***	−0.11***	−0.10***	−0.19***	−0.15***	−0.14***
	(−5.34)	(−4.71)	(−4.45)	(−6.27)	(−5.54)	(−5.30)
Expense ratio	0.96***	0.89***	0.85***	1.25***	1.23***	1.19***
	(8.63)	(8.91)	(8.58)	(9.30)	(10.22)	(9.86)
Turnover ratio	0.30***	0.32***	0.35***	0.21**	0.25***	0.27***
	(4.21)	(5.27)	(5.82)	(2.37)	(3.38)	(3.85)
Rank(YTD return)	−0.71**	−0.47**	−0.17	−0.06	0.03	0.14
	(−2.56)	(−2.59)	(−0.90)	(−0.29)	(0.16)	(0.80)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	38,883	64,735	51,917	38,883	64,735	51,917
Adjusted R^2	0.38	0.51	0.55	0.34	0.48	0.52

Table 6: Time-varying excess risk-taking of growth funds compared to value funds

This table presents the results of the following linear regression model:

$$Risk\ taking_{i,t} = \delta\left(\log(B/M)_{t-1}^{\text{Hi}-\text{Lo}} \times \mathbb{1}(Growth_{i,t-1})\right) + \beta \log(B/M)_{t-1}^{\text{Hi}-\text{Lo}} + \gamma\Gamma_{i,t-1} + \theta_i + \varepsilon_{i,t}$$

where i indexes mutual funds, and t indexes time (in months). $Risk\ taking_{i,t}$ represents fund i's idiosyncratic volatility or tracking error during time t, annualized and reported as a percentage. $\log(B/M)_{t-1}^{\text{Hi}-\text{Lo}}$ is the value spread, defined as the log difference of the book-to-market ratios of the value portfolio (top book-to-market decile) and the growth portfolio (bottom book-to-market decile) at the end of the previous June. $\mathbb{1}(Growth_{i,t-1})$ is an indicator variable that takes a value of one if fund i belongs to one of the growth categories (Large Growth, Mid-cap Growth, or Small Growth) at time $t-1$, and zero if fund i belongs to one of the value categories (Large Value, Mid-cap Value, or Small Value) at time $t-1$. $\Gamma_{i,t-1}$ is a vector of lagged fund characteristics, including the natural logarithm of total net assets (TNA), expense ratio, turnover ratio, and percentile rank of year-to-date (YTD) return. θ_i denotes fund fixed effects. The regressions are run separately for the period before the refinement of Morningstar peer groups (January 1988 to May 2002) and the period after the refinement (June 2002 to December 2022). Standard errors are double-clustered by fund and time, and t-statistics are reported in parentheses, with statistical significance at the 10%, 5%, and 1% levels indicated by *, **, and ***, respectively.

	Idiosyncratic volatility		Tracking error	
	Pre	Post	Pre	Post
	(1)	(2)	(3)	(4)
$\log(\text{B/M})^{\text{Hi–Lo}} \times \mathbb{1}(\text{Growth})$	0.54***	−0.09	0.64***	0.08
	(2.74)	(−0.91)	(3.17)	(0.74)
$\log(\text{B/M})^{\text{Hi–Lo}}$	2.02***	1.28***	1.91***	1.24***
	(7.54)	(7.60)	(7.31)	(6.30)
$\mathbb{1}(\text{Growth})$	−0.44	0.26	−0.73	−0.69*
	(−0.52)	(0.80)	(−0.80)	(−1.71)
log(TNA)	0.05	−0.13***	−0.02	−0.19***
	(0.71)	(−3.50)	(−0.28)	(−4.13)
Expense ratio	−0.06	0.36**	−0.10	0.44**
	(−0.38)	(2.02)	(−0.56)	(2.23)
Turnover ratio	0.08	0.30***	0.02	0.37***
	(1.14)	(6.77)	(0.29)	(6.94)
Rank(YTD return)	−0.09	0.04	0.13	0.14***
	(−1.03)	(1.00)	(1.46)	(2.95)
Fund fixed effects	Yes	Yes	Yes	Yes
Observations	79,480	315,483	79,480	315,483
Adjusted R^2	0.53	0.44	0.51	0.45

Table 7: Increased risk-taking of growth funds around rating thresholds

Panel A of this table presents the results of the following linear regression model:

$$Risk\ taking_{i,t} = \delta\left(\mathbb{1}(Growth_{i,t-1}) \times \mathbb{1}(Pre_t)\right) + \beta\mathbb{1}(Growth_{i,t-1}) + \gamma\Gamma_{i,t-1} + \theta_t + \varepsilon_{i,t}$$

where i indexes mutual funds and t indexes time (in months). $Risk\ taking_{i,t}$ represents fund i's idiosyncratic volatility or tracking error during time t, annualized and reported as a percentage. $\mathbb{1}(Growth_{i,t-1})$ is an indicator variable that takes a value of one if fund i belongs to one of the growth categories (Large Growth, Mid-cap Growth, or Small Growth) at time $t-1$, and zero if fund i belongs to one of the value categories (Large Value, Mid-cap Value, or Small Value) at time $t-1$. $\mathbb{1}(Pre_t)$ is an indicator variable that takes a value of one if time t is prior to the refinement of Morningstar peer groups in June 2002, and zero otherwise. $\Gamma_{i,t-1}$ is a vector of lagged fund characteristics, including the natural logarithm of total net assets (TNA), expense ratio, turnover ratio, and percentile rank of year-to-date (YTD) return. θ_t denotes time fixed effects. The regressions are run separately for sub-samples of funds split by an indicator variable, $\mathbb{1}(Close_{i,t-1})$, which indicates whether a fund is moving closer to a rating threshold for an upgrade. Specifically, $\mathbb{1}(Close_{i,t-1})$, which indicates whether fund i is moving closer to a rating threshold for a rating update at time t, as measured at time $t-1$ (Kim, 2022). Our sample spans the period from 1988 to 2022. Standard errors are double-clustered by fund and time, and t-statistics are reported in parentheses, with statistical significance at the 10%, 5%, and 1% levels indicated by *, **, and ***, respectively.

Panel A

	Idiosyncratic volatility		Tracking error	
	Close	Distant	Close	Distant
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Growth) × $\mathbb{1}$(Pre)	1.07***	0.78***	1.01***	0.81***
	(7.17)	(6.71)	(5.84)	(6.29)
$\mathbb{1}$(Growth)	0.72***	0.68***	0.38***	0.44***
	(10.31)	(13.57)	(4.23)	(7.04)
log(TNA)	−0.08***	−0.08***	−0.11***	−0.10***
	(−4.88)	(−7.07)	(−5.46)	(−7.57)
Expense ratio	1.15***	1.02***	1.59***	1.38***
	(13.00)	(16.68)	(14.69)	(18.34)
Turnover ratio	0.12**	0.16***	0.08	0.10*
	(2.58)	(3.86)	(1.48)	(1.95)
Rank(YTD return)	0.02	0.03	0.10	0.10*
	(0.31)	(0.44)	(1.38)	(1.83)
Time fixed effects	Yes	Yes	Yes	Yes
Observations	93,559	291,952	93,559	291,952
Adjusted R^2	0.44	0.46	0.42	0.44

Table 7–*Continued*

Panel B of this table presents the results of the following linear regression model:

$$\begin{aligned} \textit{Risk taking}_{i,t} &= \rho\,(\mathbb{1}(\textit{Growth}_{i,t-1}) \times \mathbb{1}(\textit{Pre}_t) \times \mathbb{1}(\textit{Close}_{i,t-1})) + \delta_1\,(\mathbb{1}(\textit{Growth}_{i,t-1}) \times \mathbb{1}(\textit{Pre}_t)) \\ &+ \delta_2\,(\mathbb{1}(\textit{Growth}_{i,t-1}) \times \mathbb{1}(\textit{Close}_{i,t-1})) + \delta_3\,(\mathbb{1}(\textit{Pre}_t) \times \mathbb{1}(\textit{Close}_{i,t-1})) \\ &+ \beta_1 \mathbb{1}(\textit{Growth}_{i,t-1}) + \beta_2 \mathbb{1}(\textit{Close}_{i,t-1}) + \gamma \Gamma_{i,t-1} + \theta_t + \varepsilon_{i,t} \end{aligned}$$

where i indexes mutual funds and t indexes time (in months). $\textit{Risk taking}_{i,t}$ represents fund i's idiosyncratic volatility or tracking error during time t, annualized and reported as a percentage. $\mathbb{1}(\textit{Growth}_{i,t-1})$ is an indicator variable that takes a value of one if fund i belongs to one of the growth categories (Large Growth, Mid-cap Growth, or Small Growth) at time $t-1$, and zero if fund i belongs to one of the value categories (Large Value, Mid-cap Value, or Small Value) at time $t-1$. $\mathbb{1}(\textit{Pre}_t)$ is an indicator variable that takes a value of one if time t is prior to the refinement of Morningstar peer groups in June 2002, and zero otherwise. $\mathbb{1}(\textit{Close}_{i,t-1})$ is an indicator variable that takes a value of one if fund i is moving closer to a rating threshold for a rating update at time t, as measured at time $t-1$, and zero otherwise (Kim, 2022). Our sample spans the period from 1988 to 2022. $\Gamma_{i,t-1}$ is a vector of lagged fund characteristics, including the natural logarithm of total net assets (TNA), expense ratio, turnover ratio, and percentile rank of year-to-date (YTD) return. θ_t denotes time fixed effects. In columns (2) and (4), time fixed effects are replaced by lagged category × time fixed effects. Our sample spans the period from 1988 to 2022. Standard errors are double-clustered by fund and time, and t-statistics are reported in parentheses, with statistical significance at the 10%, 5%, and 1% levels indicated by *, **, and ***, respectively.

	Idiosyncratic volatility		Tracking error	
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Growth) × $\mathbb{1}$(Pre) × $\mathbb{1}$(Close)	0.29***	0.21**	0.22*	0.21**
	(2.82)	(2.31)	(1.86)	(2.01)
$\mathbb{1}$(Growth) × $\mathbb{1}$(Pre)	0.78***		0.81***	
	(6.69)		(6.26)	
$\mathbb{1}$(Growth) × $\mathbb{1}$(Close)	0.03	0.05	−0.05	−0.02
	(0.80)	(1.28)	(−0.87)	(−0.40)
$\mathbb{1}$(Pre) × $\mathbb{1}$(Close)	−0.06	−0.12*	−0.06	−0.17**
	(−0.89)	(−1.81)	(−0.63)	(−2.21)
$\mathbb{1}$(Growth)	0.68***		0.44***	
	(13.62)		(7.04)	
$\mathbb{1}$(Close)	0.43***	0.42***	0.64***	0.63***
	(12.31)	(12.44)	(12.83)	(13.20)
Fund characteristics	Yes	Yes	Yes	Yes
Time fixed effects	Yes		Yes	
Category × Time fixed effects		Yes		Yes
Observations	385,511	385,511	385,511	385,511
Adjusted R^2	0.46	0.52	0.44	0.48

Table 8: Risk-taking of growth funds across different advisory fee structures

Panel A of this table presents the results of the following linear regression model:

$$\mathit{Risk\ taking}_{i,t} = \delta\left(\mathbb{1}(\mathit{Growth}_{i,t-1}) \times \mathbb{1}(\mathit{Pre}_t)\right) + \beta\mathbb{1}(\mathit{Growth}_{i,t-1}) + \gamma\Gamma_{i,t-1} + \theta_t + \varepsilon_{i,t}$$

where i indexes mutual funds and t indexes time (in months). $\mathit{Risk\ taking}_{i,t}$ represents fund i's idiosyncratic volatility or tracking error during time t, annualized and reported as a percentage. $\mathbb{1}(\mathit{Growth}_{i,t-1})$ is an indicator variable that takes a value of one if fund i belongs to one of the growth categories (Large Growth, Mid-cap Growth, or Small Growth) at time $t-1$, and zero if fund i belongs to one of the value categories (Large Value, Mid-cap Value, or Small Value) at time $t-1$. $\mathbb{1}(\mathit{Pre}_t)$ is an indicator variable that takes a value of one if time t is prior to the refinement of Morningstar peer groups in June 2002, and zero otherwise. $\Gamma_{i,t-1}$ is a vector of lagged fund characteristics, including the natural logarithm of total net assets (TNA), expense ratio, turnover ratio, and percentile rank of year-to-date (YTD) return. θ_t denotes time fixed effects. The regressions are run separately for sub-samples of funds split by the fund's advisory fee structure, as measured by Coles' incentive rate (CIR) (Coles et al., 2000). A fund's advisory fee rate is linear with respect to its total net assets if $\mathit{CIR} = 0$, and concave if $\mathit{CIR} < 0$. Our sample spans the period from 1988 to 2022. Standard errors are double-clustered by fund and time, and t-statistics are reported in parentheses, with statistical significance at the 10%, 5%, and 1% levels indicated by *, **, and ***, respectively.

Panel A

	Idiosyncratic volatility		Tracking error	
	Linear	Concave	Linear	Concave
	(1)	(2)	(3)	(4)
𝟙(Growth) × 𝟙(Pre)	1.21***	0.55***	1.31***	0.51**
	(7.89)	(3.16)	(7.38)	(2.57)
𝟙(Growth)	0.70***	0.73***	0.39***	0.46***
	(8.72)	(10.38)	(3.82)	(4.91)
log(TNA)	−0.08***	−0.07***	−0.11***	−0.09***
	(−4.50)	(−3.29)	(−4.73)	(−3.55)
Expense ratio	1.22***	0.94***	1.68***	1.22***
	(13.35)	(7.69)	(15.16)	(7.97)
Turnover ratio	0.11*	0.17***	0.05	0.12*
	(1.90)	(3.10)	(0.69)	(1.66)
Rank(YTD return)	0.02	0.05	0.14**	0.10
	(0.26)	(0.75)	(1.99)	(1.32)
Time fixed effects	Yes	Yes	Yes	Yes
Observations	198,988	150,001	198,988	150,001
Adjusted R^2	0.45	0.46	0.43	0.42

Table 8–*Continued*

Panel B of this table presents the results of the following linear regression model:

$$\begin{aligned} \textit{Risk taking}_{i,t} = &\ \rho\,(\mathbb{1}(\textit{Growth}_{i,t-1}) \times \mathbb{1}(\textit{Pre}_t) \times \mathbb{1}(\textit{Linear}_{i,t-1})) + \delta_1\,(\mathbb{1}(\textit{Growth}_{i,t-1}) \times \mathbb{1}(\textit{Pre}_t)) \\ &+ \delta_2\,(\mathbb{1}(\textit{Growth}_{i,t-1}) \times \mathbb{1}(\textit{Linear}_{i,t-1})) + \delta_3\,(\mathbb{1}(\textit{Pre}_t) \times \mathbb{1}(\textit{Linear}_{i,t-1})) \\ &+ \beta_1 \mathbb{1}(\textit{Growth}_{i,t-1}) + \beta_2 \mathbb{1}(\textit{Linear}_{i,t-1}) + \gamma \Gamma_{i,t-1} + \theta_t + \varepsilon_{i,t} \end{aligned}$$

where i indexes mutual funds and t indexes time (in months). $\textit{Risk taking}_{i,t}$ represents fund i's idiosyncratic volatility or tracking error during time t, annualized and reported as a percentage. $\mathbb{1}(\textit{Growth}_{i,t-1})$ is an indicator variable that takes a value of one if fund i belongs to one of the growth categories (Large Growth, Mid-cap Growth, or Small Growth) at time $t-1$, and zero if fund i belongs to one of the value categories (Large Value, Mid-cap Value, or Small Value) at time $t-1$. $\mathbb{1}(\textit{Pre}_t)$ is an indicator variable that takes a value of one if time t is prior to the refinement of Morningstar peer groups in June 2002, and zero otherwise. $\mathbb{1}(\textit{Linear}_{i,t-1})$ is an indicator variable that takes a value of one if fund i's advisory fee rate is linear with respect to its total net assets at time $t-1$, and zero otherwise. $\Gamma_{i,t-1}$ is a vector of lagged fund characteristics, including the natural logarithm of total net assets (TNA), expense ratio, turnover ratio, and percentile rank of year-to-date (YTD) return. θ_t denotes time fixed effects. In columns (2) and (4), time fixed effects are replaced by lagged category × time fixed effects. Our sample spans the period from 1988 to 2022. Standard errors are double-clustered by fund and time, and t-statistics are reported in parentheses, with statistical significance at the 10%, 5%, and 1% levels indicated by *, **, and ***, respectively.

Panel B

	Idiosyncratic volatility		Tracking error	
	(1)	(2)	(3)	(4)
𝟙(Growth) × 𝟙(Pre) × 𝟙(Linear)	0.62***	0.55***	0.77***	0.68***
	(2.99)	(2.89)	(3.16)	(2.99)
𝟙(Growth) × 𝟙(Pre)	0.57***		0.53***	
	(3.25)		(2.65)	
𝟙(Growth) × 𝟙(Linear)	−0.01	−0.10	−0.05	−0.07
	(−0.12)	(−1.17)	(−0.41)	(−0.60)
𝟙(Pre) × 𝟙(Linear)	−0.15	−0.09	−0.26	−0.21
	(−1.08)	(−0.74)	(−1.53)	(−1.36)
𝟙(Growth)	0.72***		0.45***	
	(10.33)		(4.78)	
𝟙(Linear)	0.09	0.11*	0.25**	0.20**
	(1.28)	(1.70)	(2.34)	(2.07)
Fund characteristics	Yes	Yes	Yes	Yes
Time fixed effects	Yes		Yes	
Category × Time fixed effects		Yes		Yes
Observations	348,989	348,989	348,989	348,989
Adjusted R^2	0.45	0.52	0.43	0.48

Table 9: Efficacy of risk taking by disadvantaged growth funds

This table compares the average Morningstar ratings of growth funds and value funds, separately for the period before the refinement of Morningstar peer groups (January 1988 to May 2002) and the period after the refinement (June 2002 to December 2022), as well as for sub-samples of funds split by the indicator variable $\mathbb{1}(Close_{i,t-1})$ which indicates whether a fund is moving closer to a rating threshold (Panel A), or the indicator variable $\mathbb{1}(Linear_{i,t-1})$ which indicates whether a fund uses a linear contract (Panel B). Standard errors are double-clustered by fund and time, and t-statistics are reported in parentheses, with statistical significance at the 10%, 5%, and 1% levels indicated by *, **, and ***, respectively.

	Morningstar rating					
	January 1988 – May 2002			June 2002 – December 2022		
	Growth	Value	Growth–Value	Growth	Value	Growth–Value
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
Close	3.03	3.15	−0.12*	3.05	3.05	−0.001
			(−1.77)			(−0.04)
Distant	2.95	3.18	−0.23***	3.08	3.08	0.002
			(−4.19)			(0.06)
Close–Distant			0.11***			−0.003
			(3.10)			(−0.17)
Panel B						
Linear	3.09	3.17	−0.08	3.05	3.06	−0.01
			(−1.14)			(−0.29)
Concave	2.79	3.15	−0.36***	3.06	3.02	0.04
			(−4.29)			(0.81)
Linear–Concave			0.28***			−0.05
			(2.69)			(−0.81)

Table 10: Risk-adjusted performance of disadvantaged growth funds

This table presents the results of the following linear regression model:

$$r_{p,t} - r_{f,t} = \alpha_{Pre} + \Delta\alpha_{Post}\mathbb{1}(Post_t) + b_{Pre}MKT_t + \Delta b_{Post}MKT_t \times \mathbb{1}(Post_t) + s_{Pre}SMB_t + \Delta s_{Post}SMB_t \times \mathbb{1}(Post_t) + h_{Pre}HML_t + \Delta h_{Post}HML_t \times \mathbb{1}(Post_t) + \varepsilon_{i,t}$$

where $r_{p,t}$ is the portfolio return and $r_{f,t}$ is the risk-free rate at time t (in months). MKT_t, SMB_t, and HML_t are the returns on the market, size, and value factors of the Fama-French 3-factor model (Fama and French, 1993). $\mathbb{1}(Pre_t)$ is an indicator variable that takes the value of one if time t is after the refinement of Morningstar peer groups in June 2002, and zero otherwise. Mutual funds (indexed by i) are sorted into 2×2 portfolios, first split by the indicator variable $\mathbb{1}(Growth_{i,t-1})$, which takes the value of one if fund i belongs to one of the growth categories (Large Growth, Mid-cap Growth, or Small Growth) at time $t-1$, and zero if fund i belongs to one of the value categories (Large Value, Mid-cap Value, or Small Value) at time $t-1$. Funds are further split by the indicator variable $\mathbb{1}(Close_{i,t-1})$, which indicates whether a fund is close to a rating threshold at time $t-1$ (Panel A), or the indicator variable $\mathbb{1}(Linear_{i,t-1})$, which indicates whether a fund has a linear advisory contract at time $t-1$ (Panel B). Heteroskedasticity robust t-statistics are reported in parentheses, with statistical significance at the 10%, 5%, and 1% levels indicated by *, **, and ***, respectively.

Panel A

	Close to rating thresholds			Distant from rating thresholds		
	Growth	Value	Growth–Value	Growth	Value	Growth–Value
	(1)	(2)	(3)	(4)	(5)	(6)
α_{Pre}	0.13	−0.21**	0.34**	0.05	−0.19**	0.24**
	(1.04)	(−2.56)	(2.11)	(0.69)	(−2.15)	(2.19)
$\Delta\alpha_{Post}$	−0.24*	0.11	−0.36**	−0.16*	0.11	−0.27**
	(−1.83)	(1.16)	(−2.03)	(−1.80)	(1.07)	(−2.07)
b_{Pre}	1.11***	0.87***	0.24***	1.07***	0.90***	0.16***
	(41.16)	(31.43)	(5.38)	(55.61)	(44.47)	(5.16)
Δb_{Post}	0.40***	0.12*	0.27***	0.30***	0.08	0.22***
	(8.47)	(1.90)	(2.72)	(12.56)	(1.25)	(3.91)
s_{Pre}	−0.29***	0.42***	−0.70***	−0.22***	0.43***	−0.65***
	(−4.82)	(10.35)	(−10.33)	(−4.74)	(9.21)	(−10.01)
Δs_{Post}	−0.07**	0.08**	−0.15***	−0.04	0.05*	−0.09**
	(−2.32)	(2.55)	(−3.16)	(−1.51)	(1.88)	(−2.30)
h_{Pre}	−0.08	0.02	−0.10	−0.01	0.05	−0.06
	(−1.44)	(0.33)	(−0.95)	(−0.23)	(0.79)	(−0.97)
Δh_{Post}	0.08	−0.09*	0.17**	0.02	−0.11**	0.12*
	(1.29)	(−1.80)	(2.32)	(0.37)	(−1.99)	(1.79)

Table 10–*Continued*

Panel B

	Linear contract			Concave contract		
	Growth	Value	Growth–Value	Growth	Value	Growth–Value
	(1)	(2)	(3)	(4)	(5)	(6)
α_{Pre}	0.14	-0.20^{**}	0.34^{***}	0.004	-0.20^{**}	0.20
	(1.63)	(−2.40)	(2.75)	(0.04)	(−2.35)	(1.62)
$\Delta\alpha_{Post}$	-0.26^{**}	0.11	-0.38^{***}	−0.12	0.11	−0.23
	(−2.59)	(1.17)	(−2.69)	(−1.17)	(1.10)	(−1.54)
b_{Pre}	1.09^{***}	0.90^{***}	0.19^{***}	1.07^{***}	0.89^{***}	0.18^{***}
	(51.40)	(41.09)	(5.36)	(49.70)	(39.96)	(4.87)
Δb_{Post}	0.34^{***}	0.11	0.23^{***}	0.32^{***}	0.07	0.25^{***}
	(11.04)	(1.55)	(3.26)	(11.81)	(1.19)	(3.96)
s_{Pre}	-0.26^{***}	0.44^{***}	-0.70^{***}	-0.24^{***}	0.41^{***}	-0.65^{***}
	(−5.06)	(9.11)	(−10.34)	(−4.90)	(9.80)	(−9.74)
Δs_{Post}	-0.06^{**}	0.05^{*}	-0.11^{***}	−0.04	0.06^{**}	-0.10^{**}
	(−2.27)	(1.96)	(−2.85)	(−1.42)	(2.16)	(−2.24)
h_{Pre}	−0.01	0.04	−0.05	−0.04	0.03	−0.07
	(−0.25)	(0.60)	(−0.69)	(−1.15)	(0.47)	(−1.05)
Δh_{Post}	0.09^{*}	-0.11^{**}	0.20^{***}	−0.001	-0.10^{**}	0.10
	(1.67)	(−2.00)	(2.82)	(−0.01)	(−2.00)	(1.38)

Table 11: Disadvantage of morally constrained funds

This table compares the average Morningstar ratings of morally constrained funds with those of unconstrained funds. A fund is considered morally constrained if it underweights sin stocks, meaning its weight in sin stocks is smaller than that of its benchmark index; it is unconstrained if it overweights sin stocks. Sin stocks are defined as those of firms operating in the alcohol, tobacco, and gambling industries (Hong and Kacperczyk, 2009). Our sample spans the period from June 2002 to December 2022. Standard errors are double-clustered by fund and time, and t-statistics are reported in parentheses, with statistical significance at the 10%, 5%, and 1% levels indicated by *, **, and ***, respectively.

Morningstar rating		
Constrained	Unconstrained	Difference
(1)	(2)	(3)
3.05	3.12	-0.07^{***}
		(-3.95)

Table 12: Increased risk-taking of morally constrained funds

This table presents the results of the following linear regression model:

$$Risk\ taking_{i,t} = \beta \mathbb{1}(Morally\ constrained_{i,t-1}) + \gamma\Gamma_{i,t-1} + \theta_{i,t-1} + \varepsilon_{i,t}$$

where i indexes mutual funds and t indexes time (in months). $Risk\ taking_{i,t}$ represents fund i's total volatility, idiosyncratic volatility, or tracking error during time t, annualized and reported as a percentage. $\mathbb{1}(Morally\ constrained_{i,t-1})$ is an indicator variable that takes a value of one if fund i underweights sin stocks at time $t-1$, meaning its weight in sin stocks is smaller than that of its benchmark index, and zero otherwise. Sin stocks are defined as those of firms operating in the alcohol, tobacco, and gambling industries (Hong and Kacperczyk, 2009). $\Gamma_{i,t-1}$ is a vector of lagged fund characteristics, including the natural logarithm of total net assets (TNA), expense ratio, turnover ratio, and percentile rank of year-to-date (YTD) return. $\theta_{i,t-1}$ denotes lagged category × time fixed effects. Our sample spans the period from June 2002 to December 2022. Standard errors are double-clustered by fund and time, and t-statistics are reported in parentheses, with statistical significance at the 10%, 5%, and 1% levels indicated by *, **, and ***, respectively.

	Idiosyncratic volatility		Tracking error	
	(1)	(2)	(3)	(4)
𝟙(Morally constrained)	0.25***	0.21***	0.29***	0.24***
	(7.74)	(7.69)	(6.98)	(6.86)
log(TNA)		−0.05***		−0.06***
		(−4.23)		(−4.28)
Expense ratio		1.19***		1.61***
		(19.56)		(19.91)
Turnover ratio		0.08*		0.06
		(1.92)		(1.15)
Rank(YTD return)		0.02		0.06
		(0.29)		(1.15)
Category × Time FEs	Yes	Yes	Yes	Yes
Observations	471,184	470,922	471,184	470,922
Adjusted R^2	0.41	0.47	0.38	0.45