Revealed Heuristics: Evidence from Investment Consultants' Search Behavior*

Sudheer Chava[a]
Soohun Kim[b]
Daniel Weagley[c]

First Draft: May 2018
This Draft: March 21, 2019

Abstract

Using proprietary data from a major fund data provider, we analyze the screening activity of investment consultants (ICs) who advise institutional investors with trillions of dollars in assets. We find that ICs frequently shortlist funds using threshold screens clustered at round, base 10 numbers: $500MM for AUM, 0% for the return net of a benchmark, and quartiles for return percentile rank screens. A fund's probability of being eliminated by a screen is significantly negatively related to its future fund attention and flows, with funds just above the $500MM AUM threshold, getting 20% more page views and 5%-9% greater flows over the next year compared to similar funds just below the threshold. Our results are consistent with ICs using a two-stage, consider-then-choose decision making process, and cognitive reference numbers in selecting screening thresholds.

Keywords: heuristics, consider-then-choose, consideration sets, cognitive reference points, investment consultants, investment screens, mutual funds

*We are grateful to Daniel Brickhouse at eVestment. We are grateful to Abhiroop Mukherjee, and seminar and conference participants at Georgia Tech, KAIST, the 2018 SFS Cavalcade Asia-Pacific, and the 2018 Conference on Asia-Pacific Financial Markets for helpful comments on the paper.

[a]Georgia Institute of Technology; email: `sudheer.chava@scheller.gatech.edu`.
[b]Georgia Institute of Technology; email: `soohun.kim@scheller.gatech.edu`.
[c]Georgia Institute of Technology; email: `daniel.weagley@scheller.gatech.edu`.

1 Introduction

What processes are used by investors to make investment decisions? Answering this question has been a goal of researchers for decades, yet there is very little *direct* evidence on the investment decision processes used by investors. Most research on investor behavior analyzes outcomes like flows, trading, holdings or asset price dynamics, and makes an inference on the investment decision process. Although we learn a lot about investors from those studies, much of how investors process information to arrive at a final investment decision remains a blackbox. In this paper, we provide direct evidence on how an important class of financial decision makers, investment consultants, process investment information.

Investment consultants (ICs) are a key feature of the asset management industry and advise institutional investors on their choice of fund managers. An overwhelming majority of the U.S. public plan sponsors use investment consultants (*Pensions and Investments* (2017); Goyal and Wahal (2008)). Moreover, investment consultants' recommendations have a significant impact on fund flows. Using survey data, one can observe some of the recommendations of ICs and derive useful insights on what factors drive their recommendations (Jenkinson, Jones, and Martinez (2016)). However, it is not clear *how* the consultants process information and decide on their recommendations. We provide *direct* evidence on how ICs process information by examining ICs behavior on the website of eVestment, a major fund data provider.[1]

Our analysis focuses on the fund screens used by the investment consultants on the eVestment website. We find consultants use simple heuristics and reference points to form their initial consideration sets of funds. In particular, we find they use: (i) a two-stage consider-then-choose decision making process in which they eliminate (or screen) a significant number of funds in the first stage of the process, and (ii) cognitive reference numbers in selecting the screening thresholds. They most frequently screen on fund-level assets under

[1] eVestment is a significant player in the fund data industry with their clients comprising 70% of the top 50 global consultants, 100% of the top 50 global managers and 72% of the top 50 largest U.S. plans. Currently, eVestment clients advise or manage over $38 trillion in assets. (Source: `http://www.evestment.com`)

management (AUM) and 3-year and 5-year past returns. We find the use of cognitive reference numbers for threshold values leads to significant clustering of screens at the same threshold values. This clustering results in large discontinuities in the probability a fund is eliminated from consideration around the commonly used values. After analyzing the screening behavior of consultants, we examine the impact of screens on fund outcomes. We show the probability a fund is eliminated by a screen is correlated with future fund attention and flows and exploit the commonly used \$500M AUM threshold to provide causal evidence of these effects.

For our analysis, we obtain data on the behavior of investment consultants on eVestment's website. eVestment provides both hard and soft information on traditional and alternative funds to investment consultants and institutional investors.[2] We examine how eVestment's investment consultant clients download datasets of funds from the website. Although the clients can choose to download the universe of funds for further analysis, they frequently apply filters or screens to the data. Screening the data involves choosing a fund aspect (e.g., AUM) and a threshold value (e.g., $\geq$\$500MM). A majority of these screens eliminate at least half of the relevant universe of funds from the IC's consideration set.

ICs seem to initially screen mainly on return percentile ranks, excess returns over a benchmark and fund-level AUM (see Figure 1). Screens on firm-level assets under management occur at a much lower frequency. Screens on returns over the past 3 years and 5 years account for 30% and 32%, respectively, of the return screens, while less than 2% of the return screens are at the 2-, 4- or 6-year horizons. Surprisingly, horizons less than or equal to one year are used 15% of the time. Even though the ICs consult for long-horizon investors (plan sponsors), they appear to consider short-term performance to some degree.[3]

[2] We use the term 'fund' to refer to the investment 'products' that investment consultants recommend. Each 'product' can have multiple 'vehicles' which follow the same strategy, but may differ on other dimensions like the fee schedule. A firm can have multiple 'products' invested in different strategies. We use the term 'fund', 'product', 'vehicle' and 'manager' interchangeably throughout the paper.

[3] We can also shed light on the timing of ICs screening behavior and the performance "as of" dates used for screens. We find there is a median three month lag between the date ICs screen and download fund data and the performance "as of" date. Lags of 4+ months are not uncommon. This indicates there is likely to be a lengthy time lag before performance affects investor attention and flows. Performance as of the fourth quarter is used more often than performance as of quarters one, two or three. This does not seem to be

Why do investment consultants screen funds on AUM or past returns in the first stage of analysis? Screening is a form of consider-then-choose decision making (henceforth, CTC). CTC is a process used to choose an object from a choice set. A decision maker faced with a set of options first forms a smaller, consideration set of options, then evaluates the options within the consideration set and makes their selection. Objects outside of the consideration set are immediately eliminated from contention. The motivation for using CTC is quite straightforward - we do not evaluate all options when evaluation is costly. Hence, the decision maker trades-off the benefit of a larger consideration set against the additional evaluation costs. Jenkinson et al. (2016) provide evidence that investment consultants' recommendations are driven primarily by "soft" factors. These "soft" factors cannot be easily quantified and are costly for consultants to evaluate. Evaluating the "soft" factors of all funds in the investment universe would be extremely costly. By screening managers on a relatively costless signal (like past returns or AUM) in the initial stage of the decision making process, consultants can optimize their evaluation costs and their overall utility.

The use of CTC by consumers has been extensively studied in the marketing literature and many recent papers in economics as well as finance endogenize the information acquisition process to explain decision maker behavior.[4] In particular, Caplin, Dean, and Leahy (forthcoming) find that a consumer may optimally consider only alternatives above an endogenously determined threshold, resembling the observed behavior of ICs. We provide a descriptive model to illustrate the trade-offs facing investment consultants in Appendix B. Our model shows that the use of a CTC process with a cutoff-rule (i.e., eliminating funds below a certain threshold value) can be boundedly rational if the decision maker faces costs to evaluate each fund. To the best of our knowledge, the active use of CTC as a decision making process has

driven by seasonal variation in screening frequency; we find consultant screening activity is fairly uniform across months with small increases in June and August.

[4] Hoyer (1984), Hauser and Wernerfelt (1990) and Roberts and Lattin (1991) are seminal articles on CTC. See Hauser, Ding, and Gaskin (2009) for a survey of the CTC literature in the context of consumer choice. The economics literature on endogenous information acquisition started with the introduction of 'rational inattention' by Sims (2003). See Van Nieuwerburgh and Veldkamp (2009, 2010) and Li (2018) for the application of rational inattention to investment decisions.

not been studied in the context of investment management. CTC stands in contrast to most rational asset pricing theories, including CAPM, which assume that investors consider all traded assets in an economy.[5]

After documenting the frequency different fund aspects are used by ICs, we next examine how ICs pick screening thresholds. We find evidence consistent with the use of a cognitive reference numbers heuristic. If an IC chooses to screen on a specific aspect, they are free to manually input the threshold value (i.e., there is no drop down menu of choices). Even so, there is significant commonality in the threshold values used across ICs with clustering at round numbers especially of base 5 and base 10. There is clear clustering at the $100MM, $500MM and $1B fund-level AUM thresholds (Figure 2). Threshold values of $99MM, $499MM or $999MM are used zero times. Funds with AUM of $499MM have a 44% chance of being eliminated, while funds with AUM of $501M have only a 29% chance of being eliminated when an AUM threshold is used. A similar plot for return percentile rank thresholds shows that there is significant clustering at rank quartiles (Figure 3). For example, the probability of elimination decreases by over fifty percentage points at the 50th percentile rank threshold. In other words, more than half of the return rank screens use the median as the cut-off value. For excess returns over a benchmark screens, there is significant clustering at the 0 percentage points threshold (see Figure 4).

These patterns are consistent with the use of a cognitive reference number heuristic in forming consideration sets. Rosch (1975) shows that, given a wide range of granular choices, people tend to categorize the potential choice sets into typical types based on their own cognitive reference points. Frequently, if the choice variable is expressed in numbers, they tend to categorize the numbers around multiples of ten. These human tendencies yield interesting patterns such as the discontinuous frequency of retaking the SAT around scores

[5] Merton (1987) provided the first model of limited attention in which investors only invest in a subset of securities either due to informational frictions or institutional structures. We show that even when investors have access to enormous amounts of information for the near-universe of funds, they actively limit their attention to a subset of funds.

of 900, 1000, ... , 1400 (Pope and Simonsohn, 2010) and an uneven distribution of rightmost digits in prices (Schindler and Kirby, 1997).[6] We find a similar bias when the investment consultants select the threshold level for AUM or past returns.

The clustering of screens at cognitive reference numbers leads to significant differences in the probability of a fund being eliminated from consideration around these threshold values. This difference in elimination rate has the potential to affect the amount of attention and future capital funds receive. We test for these effects using AUM screens.[7] Our first set of analysis documents the correlation between future fund outcomes and a fund's elimination rate. We include all funds and regress future fund outcomes on a fund's elimination rate plus control variables and different combinations of fixed effects. We find funds with a greater probability of being eliminated by a screen receive less attention (as proxied by page views on the eVestment website) and lower future fund flows. The effects are economically meaningful. We find a 10 percentage point increase in elimination rate is associated with 3.1 to 3.6 fewer page views over the next four quarters (which is 12.5% of the median number of page views). For flows, a 10 percentage point increase in elimination rate is associated with 5.7 to 6.1 percentage points lower flows over the next four quarters.

To more precisely estimate the effect of screening behavior on fund outcomes and to assess the impact of clustering at cognitive reference numbers, we analyze outcomes near a widely used fund threshold, \$500MM. The \$500MM value is the second most highly used threshold and is distanced enough from other highly used thresholds to allow for a clean analysis of the effect of thresholds on fund outcomes. We find funds just above the \$500MM threshold receive between 14-18% more page views and 5.1 to 8.8 percentage points greater fund flows

[6] Other examples include: real estate listings (Chava and Yao, 2017), (mis)reporting of personal assets in loan applications above round numbers (Garmaise, 2015), analysts' rounding earnings per share forecasts to the nearest nickel (Herrmann and Thomas, 2005), excess stock buying and selling on and around round numbers (Bhattacharya, Holden, and Jacobsen, 2012), poor performance of investors' that submit a disproportionate amount of limit orders at round numbers (Kuo, Lin, and Zhao, 2015) and hedge funds much more likely to report returns just greater than zero versus just less than zero (Bollen and Pool, 2012).

[7] We examine the AUM screens because it is relatively straightforward to determine which side of an AUM threshold a fund falls. We are unable to recreate the exact percentile rank distribution or excess return over benchmark and, therefore, cannot use these characteristics to conduct similar analysis.

over the next four quarters compared to funds just below the threshold. These effects are economically very significant.

We take a number of steps to control for any differences between funds just above the \$500MM threshold (treated) and funds just below (control). First, we only examine funds within \$50M of the threshold in the OLS analysis. Second, we show the result is robust to including different combinations of fund style and time fixed effects. Third, we show the result is robust to matching funds based on past performance, past flows and within the same style-time bins using coarsened exact matching. Fourth, we use a regression discontinuity design and find a similar result. As further robustness, we examine three minimally used placebo thresholds (\$400M, \$600M and \$700M thresholds) and find no effect at these thresholds. These results provide comfort there is no systematic bias driving our results.

We next examine if fund managers respond to the call option-like payoff near the AUM threshold. We find that funds just below the \$500MM threshold earn approximately 20 basis points higher average returns per quarter compared to funds just above the threshold although the statistical significance is marginal. We do not find a similar return differential at the placebo thresholds. There are three potential drivers of the outperformance for funds just below the \$500MM threshold: additional risk-taking, additional effort or manipulation. We test for a differential in risk-taking by comparing the value-weighted average CAPM-beta, standard deviation, skewness and kurtosis of fund holdings across treatment and control funds. We do not find significant differences across the two groups. Either funds below the threshold do not take on greater risk, or they adjust holdings at quarter's end to mask any additional risk-taking, or they are shifting risk along another dimension. These tests do not rule out unobservable factors such as differences in effort expended or return manipulation.

Our paper complements the evidence provided by Jenkinson et al. (2016) on the drivers of investment consultants recommendations and, in turn, capital flows.[8] They examine

[8]Jenkinson et al. (2016) show consultants recommendations have a significant effect on flows. Although investment consultants recommendations affect flows, there are mixed results on the ability of investment

consultant survey responses about asset managers and find investment consultants rely more on "soft" factors than performance factors when selecting managers. We show, on the other hand, that consultants do use performance factors and assets under management (which is potentially related to both soft and hard factors) in their manager selection process - at least in the beginning stages of analysis. We do not observe the consultants final recommendations though. It is likely, based on Jenkinson et al. (2016), that after screening on "hard" factors, consultants rely more on "soft" factors in making their final selection. In other words, the consideration set decision is driven by "hard" factors, while the final "choice" decision is driven by much more costly to evaluate "soft" factors.

Our paper is closely related to the literature on the determinants of fund flows. There is already a significant amount of evidence in support of both rational and behavioral drivers of fund flows.[9] Most similar to our paper is the large literature documenting the effect of investors' limited attention on capital allocation.[10] Most of these studies require the authors to hypothesize a specific channel through which funds or stocks enter (or leave) investors attention set (like discussion in the media) and examine if flows, trading behavior or asset price dynamics are consistent with the proposed channel. Hence, in the existing studies, the investor's construction of a consideration set is rather passive or very limited, at least there is no direct evidence investors are actively constructing their consideration set. Our paper, in contrast, provides direct evidence on the active decisions ICs make to limit their consideration

consultants to predict future performance. Jenkinson et al. (2016) find consultant recommendations are not related to future performance. Goyal and Wahal (2008) find consultants add value for small plan sponsors, but are detrimental to the performance of large plan sponsors.

[9]There are a number of papers on the rational allocation of capital to funds. The seminal work by Berk and Green (2004) provides a rational framework where fund flows are determined in an equilibrium in which investors learn about the skill of fund managers. Pastor and Stambaugh (2012) and Kim (2017) proceed further to explain the size of the fund industry. Recent studies by Berk and Van Binsbergen (2016) and Barber, Huang, and Odean (2016) reverse-engineer the risk preference of investors by connecting the observed fund flow to the risk-adjusted abnormal returns.

[10]Guercio and Tkac (2008) examine the effect of Morningstar ratings on fund flows. A number of papers examine the role of media in affecting investor attention and capital allocation including Sirri and Tufano (1998), Tetlock (2007), Kaniel, Starks, and Vasudevan (2007), Barber and Odean (2008), Engelberg and Parsons (2011), Solomon, Soltes, and Sosyura (2014), Fang, Peress, and Zheng (2014), and Kaniel and Parham (2017). Additional studies on investor attention and capital allocation include: Da, Engelberg, and Gao (2011) and Ben-Rephael, Da, and Israelsen (2017, 2018), and Li (2018).

set. They start with the near-universe of relevant options and actively reduce the number of options. Overall our paper is the first to provide direct evidence of fund flows being partially driven by the CTC and cognitive reference numbers heuristics (to our knowledge).

2 Theory and Research Design

We illustrate the trade-offs facing investment consultants by building a simple model of fund choice, which is presented in Appendix B. We show the use of a CTC process with a cutoff-rule (i.e., eliminating funds below a certain threshold value) can be boundedly rational if the decision maker faces costs to evaluate each option. In our model, investors observe a costless, noisy signal of fund manager skill (e.g., past performance) and incur an evaluation cost to learn skill more precisely (e.g., learning about the "soft" factors of funds and fund managers). We are agnostic about the source of the evaluation cost, it could be mental costs associated with processing a complex information set or pecuniary costs related to hiring additional employees.

Under the mild assumption that ICs can infer higher fund manager skill from a higher signal, it turns out that a cutoff-rule eliminating funds with a signal value below a specific threshold is optimal. This result is strikingly consistent with the observed behavior of investment consultants. The funds remaining after applying the cutoff-rule form the consideration set. The investor then evaluates all funds in the consideration set and chooses the fund that provides maximum utility. The investor chooses an optimal threshold value such that the increase in expected utility from including the marginal fund just offsets the additional cost of evaluating the marginal fund. We extend the base model to include a utility "bonus" for selecting a cognitive reference number as the threshold value and show through simulation that the distribution of threshold values resembles patterns observed in the data, which will be shown below

2.1 Research Design

We empirically examine how investment consultants (ICs) construct a consideration set of funds and the effect their consideration set choices have on fund outcomes. We conduct two sets of analysis to this end. Our first set of analysis is straightforward: we examine the frequency different screening criteria are used. This allows us to assess the importance of different fund characteristics and how ICs choose their threshold values. Our second set of analysis allows us to assess the effect of screening behavior on fund outcomes. The use of a fund screen creates a sharp discontinuity in the probability a fund enters the investment consultant's attention set. This, in turn, can create a discontinuity in the probability a fund receives capital. If there is commonality in the threshold values used across investment consultants, this should create a discontinuity in the aggregate amount of attention and future capital a fund receives. Our examination of fund level outcomes tests for these effects.

In assessing the effect of screening behavior on fund outcomes, we focus our empirical tests on the assets under management (AUM) screens. As shown in Figure 1, most investment consultants use AUM and past returns as their selection criteria. Because the benchmark return used to calculate excess returns as well as the return evaluation period varies across screens, it is not easy to analyze funds based on returns. In contrast, AUM allows for a much cleaner comparison across funds. This is why we concentrate on AUM screens. First, we examine the correlation between a fund's probability of being eliminated by an assets under management screen and either future fund attention, as measured by page views, or future flows.

We run a regression of the form:

$$Y_{i,q+1 \text{ to } q+n} = \alpha + \beta \times P(Elimination \mid AUM_{i,q}) + \gamma \times X + f_i + t_q + \epsilon_{i,q+1}, \qquad (2.1)$$

where $Y_{i,q+1 \text{ to } q+n}$ is the future outcome of interest of fund i over quarters $q+1$ to $q+n$,

$P(Elimination \mid AUM_{i,q})$ is the probability fund i is eliminated given its assets under management at time q ($AUM_{i,q}$), X is a set of control variables including the logarithm of assets under management, f_i is a firm fixed effect, and t_q is a year-quarter (time) fixed effect. The probability of elimination is the probability a fund is eliminated conditional on a fund-level assets under management screen being used and the fund's current assets under management. We calculate the probability of elimination using the entire sample of fund-level assets under management screens. For example, if the fund has an AUM of \$10MM, we calculate the percentage of screens in our sample an AUM of \$10MM would fail to pass. The coefficient of interest is β with a $\beta < 0$ indicating a negative relationship between the probability of elimination and the outcome of interest.

The regression in Equation (2.1) may suffer from an omitted variable bias, in which funds that have a low probability of elimination are different from funds with a high probability of elimination along a number of dimensions. Our next set of tests addresses this concern by examining fund outcomes around a commonly used threshold. Our empirical strategy compares funds just above the common threshold to funds just below. We use two different regression specifications to estimate the effect the use of a common threshold value has on fund outcomes.

Our first regression specification is as follows:

$$Y_{i,q+1 \text{ to } q+n} = \alpha + \beta \times A_{i,q} + \epsilon_{i,q+1}, \tag{2.2}$$

where $Y_{i,q+1 \text{ to } q+n}$ is the future outcome of interest of fund i over quarters $q+1$ to $q+n$ and $A_{i,q}$ is a dummy variable equal to one if the fund is above the threshold of interest (e.g., when analyzing the \$500MM threshold, $A_{i,q} = 1$ if AUM$\geq$\$500MM).

The identifying assumption is that funds above and below the threshold are similar along all relevant dimensions except funds above meet the threshold criteria. In other words, there is no omitted variable that is correlated with $AboveThreshold_{i,q}$ that affects the outcome of interest. We only examine funds within a \$50M band of the threshold (e.g., \$450MM to \$550MM for the \$500MM threshold) to ensure we are comparing similar funds. This tight bandwidth should minimize concerns that the funds above the threshold are systematically different from funds just below.

We further control for any differences between the treatment ($A_{i,q} = 1$) and control ($A_{i,q} = 0$) funds in two ways. First, by including either fund style fixed effects or year-quarter fixed effects or both, or fund style × year-quarter fixed effects. Second, we further match treatment and control firms using the coarsened exact matching method proposed by Iacus, King, and Porro (2012). We match on three dimensions: fund style × year-quarter bins (exactly), past quarter return, and past quarter flow. We use Sturge's rule to coarsen the return and flow variables into bins for matching. After constructing the matched sample, we run a weighted least squares regression with weights determined according to Iacus et al. (2012). Treatment funds receive a weight of one and the control funds receive a weight of $\frac{1}{Z}$, where Z is the number of control funds matched to a specific treatment fund.

Our second specification uses a regression discontinuity design. We employ a local polynomial regression around the AUM threshold. The independent variable is the fund's assets under management and the cutoff is the AUM threshold.

$$Y_{i,q+1 \text{ to } q+n} = \alpha + \beta \times A_{i,q} + \sum_{p=1}^{z} \Big(\gamma_{0,p} \times (AUM_{i,q} - T)^p + \gamma_{1,p} \times A_{i,q} \times (T - AUM_{i,q})^p\Big) + \epsilon_{i,q+1}. \tag{2.3}$$

Recall that $A_{i,q}$ is equal to one if the fund size is above the threshold and zero otherwise. The local polynomials of $(AUM_{i,q} - T)^p$ for $p = 1, \cdots, z$ continuously converge to zero around the threshold of T. Hence, β reflects the discontinuity in the local effect of $AUM_{i,q}$ around the threshold of T on the variable Y of interest.[11] Once again, the identifying assumption

[11] Calonico, Cattaneo, and Titiunik (2014) propose a method to estimate and test RDD such as (2.3). We thank to the authors for providing the Stata code. The bandwidth and the order of polynomials are optimally

is that there are no systematic unobserved differences between the funds just above or just below the threshold.

3 Data and Sample

We obtained our data on fund performance and characteristics from eVestment. eVestment is a "data, analytics and research platform serving the global institutional investment community."[12] Institutional investors and investment consultants use the eVestment website and database to analyze funds and make investment decisions or recommendations. Both traditional and alternative investment funds self-report information on performance, assets under management, fund strategy, and a number of other fund characteristics to eVestment. eVestment takes a number of steps to ensure the accuracy of the data.[13] We focus our analysis on traditional U.S. equity and fixed income funds.

We address a number of potential biases present in investment fund data. eVestment does not drop funds from the database after they delist, which minimizes concerns of a survivorship bias.[14] To minimize concerns of backfill bias, we drop all observations occurring before the fund's initial reporting date. We are allowed to use only the data after eVestment transferred to a new database system in the second quarter of 2007 and, hence, we drop all fund observations before this date. Therefore, our sample starts in Q2 2007 and goes through Q4 2016.

chosen by the criteria proposed in Calonico et al.. The kernel function is triangular (Cheng, Fan, and Marron, 1997).

[12] `https://www.evestment.com`

[13] We also checked that the holdings data in eVestment is consistent with the holdings data in the CRSP Mutual Fund Database for a random selection of funds.

[14] Although survivorship bias is not an issue, we cannot eliminate an "extinction bias" in the data due to funds delisting. Funds delist for two main reasons: (1) they are no longer taking on more capital or (2) they are shutting down. Depending on the reason, this can lead to very different biases in the data (Getmansky, Lo, and Makarov (2004)). Any potential extinction bias in the data should not affect our results since we are not interested in the average behavior of funds and any bias should not be systematically correlated with the explanatory variables of interest. This is especially unlikely in our analysis of fund behavior around specific AUM thresholds.

In addition to the fund performance and characteristics data, we obtained proprietary data on the usage of the eVestment platform. In particular, users of the eVestment website leave interesting records when they use the eVestment platform to build datasets for analysis or to look up specific funds for further review. The data covers these two main activities. Specifically, the data includes: (1) fund page views each month (across all user types), and (2) the screens used by ICs when creating datasets for analysis. The page views data covers all traditional, U.S. equity and fixed income funds over the time period Q1 2008 to Q4 2016. In Figure 5, we plot the average page views each quarter over time. The average fund experiences around 20 page views per quarter during our sample. Within equity funds, we find large capitalization and small capitalization funds experience slightly more attention than all capitalization and mid-capitalization funds. We also find equity funds are viewed more than fixed income funds on average. There is an increase in usage in the first quarter of 2010 with average page views nearly doubling between the last quarter of 2009 and first quarter of 2010. We use time fixed effects in the majority of our analysis, so this time trend should not affect our results.

The screen data covers all return-related and assets under management-related screens from September 10, 2012 to November 2, 2017.[15] Screens on other criteria like manager tenure, fund location, fees, etc. are not included in the dataset. eVestment personnel claim that return and AUM screens are by far the most frequently used. A screen observation consists of the relevant aspect (e.g., fund assets under management), an operator (e.g., $\geq$), the threshold used (e.g., \$100M), the date of the screen, a fund universe (e.g., U.S. large cap value funds), and, for screens on excess returns over a benchmark, the benchmark index chosen. We do not observe if screens are linked through the same database query. For example, if an IC screens on both AUM and one year return in a query, we see each screen as a separate observation and cannot link them.

In Figure 6, we provide a screenshot of the eVestment webpage used to build a data set

[15] In the appendix, we plot the number of screens by year-quarter. There is no discernible time trend.

of funds for analysis. A critical feature of the eVestment platform is that users are relatively unconstrained in choosing the threshold value. The user must manually input or choose on a slider the threshold value. There is not a small set of drop down menu choices. This set-up allows us to interpret the clustering of thresholds as being driven by the users decision making process and not due to a feature of the eVestment website.

The main dependent variables in our sample are: fund page views over the next four quarters ($Views_{q+1-4}$), fund flow over the next four quarters ($Flow_{q+1-4}$), and next quarter return (r_{q+1}). The fund flow is calculated as the total dollar flow over four quarters from one quarter ahead ($q + 1$) to four quarters ahead ($q + 4$) divided by the initial AUM (i.e., AUM at the end of the quarter q). The main explanatory variables are a dummy variable equal to one if the fund's AUM at the end of quarter q is above \$500MM ($Above500$) and the fund's elimination rate conditional on its AUM at the end of quarter q ($ElimRate$). We calculate the elimination rate by dividing the number of AUM screens the fund would pass based on its AUM in quarter q by the total number of AUM screens in our screen sample. We use the full sample of AUM screens to calculate the elimination rate for each dollar amount of AUM.

We modify the sample and variables in two ways. First, we drop funds with AUM greater than \$2.5B because there is almost no variation in the probability a fund is eliminated by an AUM screen beyond \$2.5B. This minimizes concerns of outliers (in terms of AUM) affecting our results. Second, we Winsorize all flow and views variables at the 1% level, again to minimize concerns of outliers driving our results.

We provide summary statistics for the variables used in our regression analysis in Table 1. The summary statistics for the full sample are in Panel A. In Panel B (Panel C), we report the summary statistics for the funds within the \$450MM to \$500MM (\$500MM to \$550MM) AUM range. The three control variables are fund-level AUM (AUM_q), return in quarter q (Ret_q) and flow in quarter q ($Flow_q$). Examining the means of the dependent variables for funds just below the \$500MM threshold to those just above, we see funds just above the threshold receive more page views, greater flows and earn lower returns on average, which is in line with our predictions. We formalize the analysis and test for statistical significance across the treatment and control groups in Section 4.3. Importantly, the two main control variables' (Ret_q and $Flow_q$) means are similar and not statistically significantly different across the treatment (above \$500MM) and control (below \$500MM) groups. In some of our analysis, we match funds based on these variables as well as fund style × year-quarter bins to ensure the treatment and control samples are similar on these important dimensions. In the Appendix, we provide a histogram of funds by AUM near the \$500M threshold. We find no evidence there is a discontinuous change in the proportion of funds on either side of the \$500M threshold, which minimizes concerns that fund managers are somehow selecting to be in either the treatment or control sample.

In the full sample, the average fund is eliminated by 51% of the AUM screens. This is a significant reduction in candidate funds. Examining the funds around the \$500MM threshold, we see funds just below the \$500MM threshold are eliminated by 43.9% of screens, while funds just above are eliminated by only 29% of screens. There is almost no variation in the elimination rate on either side of the \$500MM threshold since ICs almost never use a threshold value between \$450MM and \$550MM that is not equal to \$500MM. This large discontinuity in the probability of elimination for otherwise similar funds allows us to estimate a causal impact of screening behavior on future fund outcomes.

4 Results

4.1 Investment Consultant Screening Behavior

We begin our analysis by presenting basic facts on investment consultant screening behavior.[16] In Figure 1, we document the frequency each screen-type is used. For return-

[16]In Appendix B, we propose a simple fund search model which rationalizes the observed screening behavior. In particular, our model predicts that investment consultants will screen funds using the characteristic(s) most

based screens, the user can screen on either fund return or fund excess return over a benchmark. The consultants are able to choose the benchmark used. The AUM screens are separated into four types: firm-level total AUM, firm-level institutional AUM, fund-level total AUM and fund-level institutional AUM. Return criteria are used more frequently than AUM criteria with fund returns the most commonly used criteria. Fund return screens account for 29% of the sample, excess return screens 27% of the sample and AUM screens 44% of the sample. The most commonly used AUM criteria are fund-level total AUM screens. Screens on fund-level total AUM are used almost twice as much as firm-level total AUM screens.

We find some heterogeneity in the direction of AUM screens. 90% of AUM screens eliminate funds below a certain AUM threshold, while the remainder eliminate funds above the threshold (results not presented). A subset of ICs appear to have a preference for smaller funds. These ICs may be looking for more flexible managers with more unique strategies. For example, it is becoming prevalent for public plans to allocate some portion of their portfolio to emerging managers.

Next, we examine the type of thresholds used to screen on returns. ICs can use a numerical value threshold (e.g., 0%) or a percentile rank within a fund universe. In Figure 7, we provide the frequencies that each type of threshold is used. For excess returns over the benchmark, a numerical value threshold is used more frequently than percentile rank. For raw returns, percentile rank is used much more frequently. Only 7.6% of the time are raw returns screened on numerical values. We find that over 97% of the time, funds with returns below the threshold value are eliminated (i.e., the $\geq$ or $>$ operator is used). These results confirm the notion that relative performance compared to a benchmark or your peers is much more important to ICs than raw performance.

Another interesting dimension is the time horizon investment consultants use to evaluate fund managers. In Figure 8, we plot the frequency each time horizon is used for screening

informative of fund manager skill, as stated in Proposition B.1. We can, therefore, infer from consultant screens the characteristics they believe are most informative of fund manager skill (within the set of characteristics that are relatively "costless" to observe).

on return performance. We find ICs are most likely to use medium-term performance to screen investments. The three year and five year horizons are used 30% and 32% of the time, respectively.[17] Very short time horizons are used frequently as well with one year and calendar year screens accounting for close to 15% of return screens. Longer time horizons are much less frequently used with time horizons greater than five years combining for fewer screens than the five year horizon alone. Although investment consultants usually make recommendations to "long-term" investors, their screening behavior indicates they care about short-term performance in their fund selection process.

Clearly, past returns and assets under management are important signals to ICs. There are a number of potential explanations for why ICs screen on these dimensions. Most likely, ICs believe past returns and current AUM are positively correlated with fund manager skill and, potentially, future performance. For example, the model by Berk and Green (2004) shows how investors' belief about fund manager skill will evolve with past performance and current AUM. Additionally, it is possible ICs use AUM to proxy for other important fund characteristics that are more difficult to observe (e.g., operational risk).

In the Appendix A, we provide a number of additional plots examining the seasonality in IC searches (as proxied by screening) and the performance "as of" date criteria. We find evidence of seasonality in both. IC search activity is fairly uniform throughout the year except there are spikes in screen activity in June and August. We find that fund information as of the end of the year (fourth quarter) is the most commonly used to screen funds with fourth quarter screens used approximately 20% more than the next most frequent quarter. Lastly, we find there is a lag between performance and IC search with a median of three months between the performance "as of" date and screen date. This lag combined with the time it takes to further analyze the data, make a final decision, and implement the decision indicates there should be a multi-quarter lag between the fund's reporting of information and its effect on fund attention and flows.

[17] A 3-year horizon is often considered a "market cycle" in the fund industry.

4.2 Threshold Values

In this section, we examine the distribution of threshold values used by investment consultants.[18] In Figure 2, we overlay the threshold frequencies for AUM screens on the probability a fund is eliminated conditional on its current AUM and conditional on an AUM screen being used. Two important observations stand out from this figure. First, ICs eliminate small funds at a significant rate. A fund with AUM of $10M has over an 80% chance of being eliminated. The elimination rate is decreasing in AUM until the $1B threshold, after which the probability of elimination levels out around 15%. Second, ICs frequently use round base 10 numbers as the threshold value. The $100M, $500MM and $1B threshold values are used 13%, 15% and 20% of the time, respectively, while values within $5M (e.g., $95M-$105M) of these thresholds are used zero times.

In Figures 3 and 4, we present similar plots for raw return percentile rank thresholds and excess return numerical value thresholds, respectively. Examining the raw return percentile rank thresholds, we see a similar pattern to the AUM thresholds with funds of low rank experiencing a very high elimination probability. Funds in the 5th percentile have an elimination probability near 100% and all funds below the 25th percentile have an elimination probability over 90%. There is significant clustering of threshold values at the 25th and 50th percentiles. At the 50th percentile, the elimination rate declines by over 50 percentage points. This creates a large discontinuity in the number of IC choice sets a fund enters into near this threshold. Surprisingly, screens on extremely good performance are not as common with zero screens at the 90th percentile and around 9% at the 95th percentile. This indicates that ICs do not necessarily chase after the top performers.

We find very similar patterns in the excess return numerical value thresholds. Funds with

[18]We provide a simple model in Appendix B. In the proposed model, if we allow investment consultants to have different evaluation costs, then we should observe significant variation in the threshold values used. Each investment consultant likely has a different evaluation cost and should, therefore, have a different optimal threshold value. In addition, if investment consultants use cognitive reference numbers in selecting threshold values, there can potentially be significant commonality in threshold values (as in Figure A.1). Our results are consistent with the predicted behavior of investment consultants using cognitive reference numbers.

excess returns less than zero are eliminated over 90% of the time. The elimination rate is decreasing in the excess return with large drops in the elimination rate at specific values. The 0% threshold is used a significant amount, accounting for close to 50% of the screens. There is also clustering at the 0.5%, 1%, 1.5% and 2% threshold values. Once again, values near the commonly used thresholds are rarely used.

The clustering of thresholds at specific values is unlikely to be the outcome of investment consultants selecting their optimal threshold value. Take the 50th percentile rank threshold as an example. In nearly every period, the 50th percentile rank is the most frequently used threshold. It is possible that in a certain period the 50th percentile rank was the optimal threshold for a number of ICs, but it is nearly impossible that the 50th percentile rank was the optimal threshold every period for a number of ICs. If an IC searches for the optimal threshold in each period, the IC will likely have a different optimal threshold value each time. The number of funds and the evaluation costs are constantly changing, which should change the optimal threshold value.

The clustering of threshold values is consistent with consultants selecting cognitive reference numbers as threshold values. It is important to highlight that our results do not imply that the consultants select random thresholds among cognitive reference numbers. Although ICs may be subject to the bias of congnitive references, their choice can be partially rational if they select cognitive reference numbers near the optimal threshold value (as in Proposition B.4).

4.3 Elimination and Fund Outcomes

Our goal for the next set of analysis is to examine the impact screen behavior has on fund outcomes. Specifically, we examine if funds with lower elimination rates experience greater attention, as measured by page views, and greater fund flows. We focus our analysis on AUM screens for this set of tests.[19]

We begin our analysis by regressing fund page views over the next four quarters on elimination rate according to Equation (2.1). The elimination rate is calculated using the fund's AUM at the end of the most recent quarter. A negative coefficient represents a decrease in page views as the elimination rate increases.

We present the results in Table 2. We find funds that are eliminated by AUM screens at a higher rate receive less attention over the next four quarters. The coefficient is between -30 to -37 across all specifications and is unaffected by controlling for time fixed effects, style fixed effects or time $\times$ style fixed effects. We cluster standard errors at the fund and year-quarter level and find the coefficient has a p-value< 0.01 in all specifications. The relationship between screen behavior and measured attention is significant. In the most stringent specification (Column (5)), a 10 percentage point increase in the elimination rate is associated with an average decline in page views of 3.6 views. For the median fund, this is an approximately 12.5% reduction in page views. These results highlight the strong correlation between elimination rates and attention after controlling for the effect of assets under management.

We next examine if elimination rates are associated with fund flows. We conduct similar tests with the percentage flow over the next four quarters as the dependent variable of interest. Results are presented in Table 3. We find a strong negative relationship between the elimination rate and future fund flows. The coefficient is between -0.57 and -0.61 and is significant at the 1% level in all specifications. A ten percentage points increase in the elimination rate is associated with 5.7 to 6.1 percentage points lower flows on average. Considering the large changes in elimination rate near certain threshold values, there are

[19] We are currently unable to examine the percentile rank or excess return thresholds due to an inability to precisely recreate these values. We cannot examine performance percentile ranks because we do not have the historical universe classifications for funds. We cannot use excess returns because we do not know eVestment's excess return calculation process (e.g., which fund is used, are gross/net return used, etc.). Additionally, this requires an assignment of fund to benchmark, which adds additional complexity. We are currently seeking the information required to conduct this analysis.

potentially significant different outcomes for otherwise similar funds right around these thresholds.

The previous tests examine the correlation between a fund's elimination rate and future attention and flows. There are potentially a number of omitted variables correlated with the elimination rate that are also related to attention and flows. In our next set of tests, we address these concerns by examining fund outcomes right around a commonly used AUM threshold, the \$500MM threshold. This test also provides an estimate of the impact the use cognitive reference numbers has on fund outcomes. We chose the \$500MM threshold because it is far enough away from the other commonly used thresholds that there should not be an overlapping effect (unlike the \$100M threshold), yet the funds are still small enough to not already be in a vast majority of ICs' consideration set (like the \$1B or \$2B thresholds).[20]

We first examine the effect of being above the \$500MM threshold on page views. We only include funds within \$50M of the threshold (\$450MM-\$550MM) in this analysis. The regression is specified according to Equation (2.2) with page views over the next four quarters regressed on a dummy variable equal to one if the fund's AUM is greater than \$500MM. A positive coefficient represents a positive effect from the sharp decline in the fund elimination rate (i.e., surviving more screens). The identifying assumption is that funds just below the \$500MM threshold are similar to funds just about the \$500MM threshold along other relevant dimensions.

Results are presented in Table 4. We find a significant effect of being above the \$500MM threshold on future page views with coefficient estimates between 8.4 and 11.2. This corresponds to a more than 20% increase in page views for the median fund. In columns (1)-(4), we control for combinations of fund style and year-quarter fixed effects. In column (4), we include style $\times$ year-quarter fixed effects. This specification removes any common

[20]In Appendix A, we provide the results for the \$1B threshold and find no effect likely because these funds have already entered most ICs consideration set. Additionally, although there is a stark decrease in the elimination rate as the \$1B threshold is crossed, funds just below the threshold still survive close to 80% of screens.

variation across funds of the same style over the same time period. After including these fixed effects, it is highly unlikely that there are omitted variables correlated with the above \$500MM dummy variable that are driving the results. Even so, we next use coarsened exact matching to ensure the sample of treatment (above \$500MM) and control funds (below \$500MM) are similar along relevant dimensions. In column (5), we exact match within style $\times$ year-quarter bins. In columns (6) and (7), we additionally match on last quarter's flow and last quarter's return, respectively. In column (8), we match on all three dimensions. Both continuous variables are coarsened using Sturge's rule. Across all specifications, the coefficient remains economically and statistically significant. Funds just above the commonly used \$500MM threshold receive significantly more attention in the future compared to similar funds just below the threshold.

In Table 5, we conduct similar analysis on future fund flows. We find a difference in fund flows of between 5.1 and 8.8 percentage points for funds just above the \$500MM threshold compared to funds just below the threshold. Examining the results in columns (1)-(8), we see the coefficient is just above 5 percent in all specifications and significant at the 10 percent level in all specifications, but one. The specification with style $\times$ year-quarter fixed effects estimates a coefficient of 5.1% with a p-value of 0.104.[21] Although the results become statistically weaker when we include a large number of fixed effects, the economic magnitude of the coefficient stays relatively stable. We next match funds in the treatment and control groups to ensure we are comparing similar funds across the treatment and control groups. The more stringent our matching method, the more economically and statistically significant the results. This provides confidence that our results are not driven by differences in funds across treatment and control groups. In column (8), we match on style $\times$ year-quarter, last quarter return and last quarter flow, and find a coefficient of 8.8% with a p-value of 0.01. There is a large and significant causal impact of the \$500MM threshold on future fund flows.

Do we see similar patterns at less commonly used thresholds? No. We conduct placebo tests

[21]We hope our readers do not suffer from a 10% cognitive reference number bias.

around the \$400M, \$600M and \$700M thresholds. We use the same regression specifications as in the previous analysis and find no robustly significant effects of being above these threshold values on page views or fund flows (results presented in the Appendix A). These results provide further comfort that the effect documented is due to screening behavior and not an omitted variable. The odds there is an omitted variable that is driving the \$500MM threshold result, that does not affect funds around the \$400M, \$600M or \$700M thresholds is extremely low.

Our final set of analysis uses a regression discontinuity design to estimate the effect of the \$500MM threshold on fund page views and fund flows. The regression specification is Equation (2.3). For these tests, we do not pre-specify a bandwidth, instead, the bandwidth is optimally chosen (see section 2 for further discussion). We present the coefficient estimates with the 99% confidence intervals for page views and fund flows in Figures 9 and 10, respectively. We also present the estimates for the \$400M, \$600M and \$700M threshold coefficients. Examining the page views result, we see the \$500MM threshold coefficient estimate is two to three times larger than the other thresholds and is the only significant coefficient estimate. The \$600M threshold estimate is actually negative and insignificant. Similarly, the placebo thresholds have a near zero effect on fund flows, while the \$500MM threshold has a significant coefficient estimate close to 10% of AUM. Taken together, the \$500MM threshold has a significant impact on fund attention and flows with no effect at the lesser used thresholds. IC screening behavior has a causal impact on fund outcomes. These results highlight the effect the use of a common cognitive reference number in investor decision making can have on fund outcomes.

Lastly, we examine the response of funds to the incentives created by the use of a common threshold. Based on the previous analysis, there is a significant increase in assets under management for funds that just cross the \$500MM threshold. Since fund fees are a percentage of AUM, the future fees funds collect should also experience a discontinuous jump at the \$500MM threshold on average. For fund managers, this creates a strong incentive to cross the threshold.

We test if funds just below the \$500MM threshold earn higher average returns than funds just above the threshold by regressing fund return in the next quarter on the above \$500MM dummy variable. We present the results in Table 6. We find some evidence that funds just above the threshold earn lower returns than funds just below the threshold. The coefficient ranges between -0.001 and -0.009 and the statistical significance depends on the specification. Column (8) presents the most robust specification with matching on style $\times$ year-quarter fixed effects, past quarter return and past quarter flow. The coefficient is -0.002 (20 basis points) and has a p-value of 0.04. In the Appendix A, we examine the placebo thresholds and find no effect. We find similar results using the RDD specification (Table 11). Considering the difficulty in increasing returns while constrained by your investment mandate, it is not surprising these results are relatively muted. Any evidence of a return differential is relatively surprising.

In unreported results, we examine the source of the outperformance for funds just below the threshold. First, we examine if a differential in fees can explain the results. This is not the case. The results are similar using gross returns instead of net returns and we do not find evidence of fee differentials around the threshold. Next, we examine if funds just below the threshold take on greater risk. We construct measures of fund beta, return standard deviation, return skewness and return kurtosis using holdings data. We find no effect of the \$500MM threshold on these values. This is potentially due to measurement error in the risk measures since the holdings data is not as well populated as the other variables of interest or it could be due to managers window dressing their holdings so that any excess risk taking is not apparent to investors. Alternatively, fund managers just below the \$500MM threshold could be adding alpha to earn slightly higher returns or manipulating their performance. We cannot distinguish between these alternatives.

Overall, our results highlight the significant effect IC screening behavior has on fund outcomes. Funds that are eliminated at a higher rate based on threshold criteria experience less attention and lower future flows. There is some evidence funds respond to the incentive created by significant screening at commonly used thresholds by increasing their returns.

5 Discussion

The use of heuristics in the investment decision process is partially inconsistent with perfectly rational-agent models of investor behavior. This does not imply that the use of heuristics is a poor strategy for ICs to use. It is possible that the simple heuristics documented in this paper outperform or perform no worse than more complex decision making algorithms out-of-sample. In a number of other contexts, heuristics have actually been shown to outperform more complex strategies (Gigerenzer and Gaissmaier (2011)). Considering the inability of academic researchers to document predictability in mutual fund performance, it may be justifiable to use a simple rule for fund selection, especially if effort or complexity is costly. On the other hand, it is also possible these heuristics underperform or perform no better than an even simpler strategy of investing in low fee index funds. It is possible to assess the performance of various strategies for a given objective function, but we do not know the true objective function of ICs. Our aim with this paper is to provide evidence on the use of heuristics by a set of sophisticated and economically-important agents, and to show that heuristics affect the flow of capital in the economy. Analysis on the performance of the documented strategies is left to future work.

6 Conclusion

We provide direct evidence on how investors process information in making an investment decision. Specifically, we examine the investment decision processes of investment consultants that advise trillions of dollars of investment capital. We present a number of stylized facts about the formation of consideration sets of funds by investment consultants. By examining screen frequencies, we are able to document the fund characteristics consultants find most informative of fund quality. The most common screens are on fund-level AUM and 3-year and 5-year past returns. Examining the threshold values used for screens, we find significant commonality in the values chosen. ICs frequently use base 10 numbers for AUM threshold values, zero percent for excess return over a benchmark threshold values, and quartiles for return percentile rank threshold values.

We show significant correlations between the probability a fund is eliminated by a screen and future fund attention and fund flows. Examining fund outcomes around the commonly used $500MM AUM threshold, we provide a causal effect of screens on future fund attention and flows. These results highlight the significant impact the use of cognitive reference numbers by ICs can have on fund outcomes. Lastly, we find some evidence funds just below the $500MM threshold earn higher average returns consistent with fund managers increasing effort or risk or manipulating returns to cross the $500MM threshold.

Investment consultants advise plan sponsors with trillions of dollars of investment capital on their allocations. We believe we are the first to document the use of investment screens and consideration sets by investment consultants and show that they have a causal effect on fund flows. Their screening behavior is consistent with a consider-then-choose decision making heuristic and the commonality in threshold values aligns with a cognitive reference number heuristic. These heuristics are commonly used in various human decision making settings. We show they are also used in the process of allocating a significant portion of global wealth.

The behavior documented in this paper is likely taking place in other financial decision making contexts where investors are incapable of processing all potentially relevant pieces of information. With the emergence of detailed information on individuals' decisions in the big data era, we expect this line of research, which directly examines the processes and steps individuals' use to make investment decisions, to become more prevalent.

References

Barber, Brad M, Xing Huang, and Terrance Odean, 2016, Which factors matter to investors? evidence from mutual fund flows, *Review of Financial Studies* 29, 2600–2642.

Barber, Brad M, and Terrance Odean, 2008, All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors, *Review of Financial Studies* 21, 785–818.

Ben-Rephael, Azi, Bruce I Carlin, Zhi Da, and Ryan D Israelsen, 2018, Demand for Information and Asset Pricing, *NBER Working Paper* .

Ben-Rephael, Azi, Zhi Da, and Ryan D Israelsen, 2017, It depends on where you search: Institutional investor attention and underreaction to news, *Review of Financial Studies* 30, 3009–3047.

Berk, J B, and R C Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* .

Berk, J B, and J H Van Binsbergen, 2016, Assessing asset pricing models using revealed preference, *Journal of Financial Economics* 119, 1–23.

Bhattacharya, Utpal, Craig W Holden, and Stacey Jacobsen, 2012, Penny Wise, Dollar Foolish: Buy-Sell Imbalances On and Around Round Numbers, *Management Science* 58, 413–431.

Blackwell, David, 1951, Comparison of experiments, Technical report, Working Paper, Howard University.

Blackwell, David, 1953, Equivalent comparisons of experiments, *The annals of mathematical statistics* 265–272.

Bollen, Nicolas P B, and Veronika K Pool, 2012, Suspicious Patterns in Hedge Fund Returns and the Risk of Fraud, *Review of Financial Studies* 25, 2673–2702.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, 2014, Robust nonparametric confidence intervals for regression-discontinuity designs, *Econometrica* 82, 2295–2326.

Caplin, Andrew, Mark Dean, and John Leahy, forthcoming, Rational inattention, optimal consideration sets and stochastic choice, *Review of Economic Studies* .

Chava, Sudheer, and Vincent W Yao, 2017, Cognitive reference points, the left-digit effect, and clustering in housing markets .

Cheng, Ming-Yen, Jianqing Fan, and James S Marron, 1997, On automatic boundary corrections, *The Annals of Statistics* 1691–1708.

Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2011, In Search of Attention, *The Journal of Finance* 66, 1461–1499.

Engelberg, Joseph E, and Christopher A Parsons, 2011, The Causal Impact of Media in Financial Markets, *The Journal of Finance* 66, 67–97.

Fang, Lily H, Joel Peress, and Lu Zheng, 2014, Does Media Coverage of Stocks Affect Mutual Funds' Trading and Performance?, *Review of Financial Studies* 27, 3441–3466.

Garmaise, Mark J, 2015, Borrower Misreporting and Loan Performance, *The Journal of Finance* 70, 449–484.

Getmansky, M, A.W. Lo, and I Makarov, 2004, An econometric model of serial correlation and illiquidity in hedge fund returns, *Journal of Financial Economics* .

Gigerenzer, Gerd, and Wolfgang Gaissmaier, 2011, Heuristic decision making, *Annual review of psychology* 62, 451–482.

Goyal, Amit, and Sunil Wahal, 2008, The Selection and Termination of Investment Management Firms by Plan Sponsors, *The Journal of Finance* 63, 1805–1847.

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

Hauser, John R, M Ding, and Steven P Gaskin, 2009, Non-compensatory (and compensatory) models of consideration-set decisions, in *Proceedings of the Sawtooth Software Conference*, Delray Beach, Fl.

Hauser, John R, and Birger Wernerfelt, 1990, An evaluation cost model of consideration sets, *The Journal of Consumer Research* 16, 393–408.

Herrmann, Don, and Wayne B Thomas, 2005, Rounding of analyst forecasts, *The Accounting Review* 80, 805–823.

Hoyer, Wayne D, 1984, An examination of consumer decision making for a common repeat purchase product, *Journal of Consumer Research* 11, 822–829.

Iacus, Stefano M, Gary King, and Fiuseppe Porro, 2012, Causal inference without balance checking: Coarsened exact matching, *Political Analysis* 1–24.

Jenkinson, Tim, Howard Jones, and Jose Vicente Martinez, 2016, Picking Winners? Investment Consultants' Recommendations of Fund Managers, *The Journal of Finance* 71, 1851–1882.

Kaniel, Ron, and Robert Parham, 2017, WSJ Category Kings – The impact of media attention on consumer and mutual fund investment decisions, *Journal of Financial Economics* 123, 337–356.

Kaniel, Ron, Laura Starks, and Vasudha Vasudevan, 2007, Headlines and Bottom Lines: Attention and Learning Effects from Media Coverage of Mutual Funds, *Working Paper* .

Kim, Jeong-Ho, 2017, Why has active asset management grown?: Learning about heterogeneity in skills 2221–2256.

Kuo, Wei-Yu, Tse-Chun Lin, and Jing Zhao, 2015, Cognitive limitation and investment performance: Evidence from limit order clustering, *The Review of Financial Studies* 28, 838–875.

Li, Jiacui, 2018, Sophisticated inattention and investor underreaction .

Merton, Robert C, 1987, A simple model of capital market equilibrium with incomplete information, *The Journal of Finance* 42, 483–510.

Pastor, L, and R F Stambaugh, 2012, On the size of the active management industry, *Journal of Political Economy* 120, 740–781.

Pensions and Investments, 2017, Consultant Directory, Pensions and Investments.

Pope, Devin, and Uri Simonsohn, 2010, Round Numbers as Goals, *Psychological Science* 22, 71–79.

Roberts, John H, and James M Lattin, 1991, Development and testing of a model of consideration set composition, *Journal of Marketing Research* 429–440.

Rosch, Eleanor, 1975, Cognitive reference points, *Cognitive Psychology* 7, 532–547.

Schindler, Robert M, and Patrick N Kirby, 1997, Patterns of Rightmost Digits Used in Advertised Prices: Implications for Nine-Ending Effects, *Journal of Consumer Research* 24, 192–201.

Sims, Christopher A, 2003, Implications of rational inattention, *Journal of monetary Economics* 50, 665–690.

Sirri, Erik R, and Peter Tufano, 1998, Costly Search and Mutual Fund Flows, *The Journal of Finance* 53, 1589–1622.

Solomon, David H, Eugene Soltes, and Denis Sosyura, 2014, Winners in the spotlight: Media coverage of fund holdings as a driver of flows, *Journal of Financial Economics* 113, 53–72.

Tetlock, Paul C, 2007, Giving Content to Investor Sentiment: The Role of Media in the Stock Market, *The Journal of Finance* 62, 1139–1168.

Van Nieuwerburgh, Stijn, and Laura Veldkamp, 2009, Information immobility and the home bias puzzle, *The Journal of Finance* 64, 1187–1215.

Van Nieuwerburgh, Stijn, and Laura Veldkamp, 2010, Information acquisition and under-diversification, *Review of Economic Studies* 77, 779–805.

7 Tables & Figures

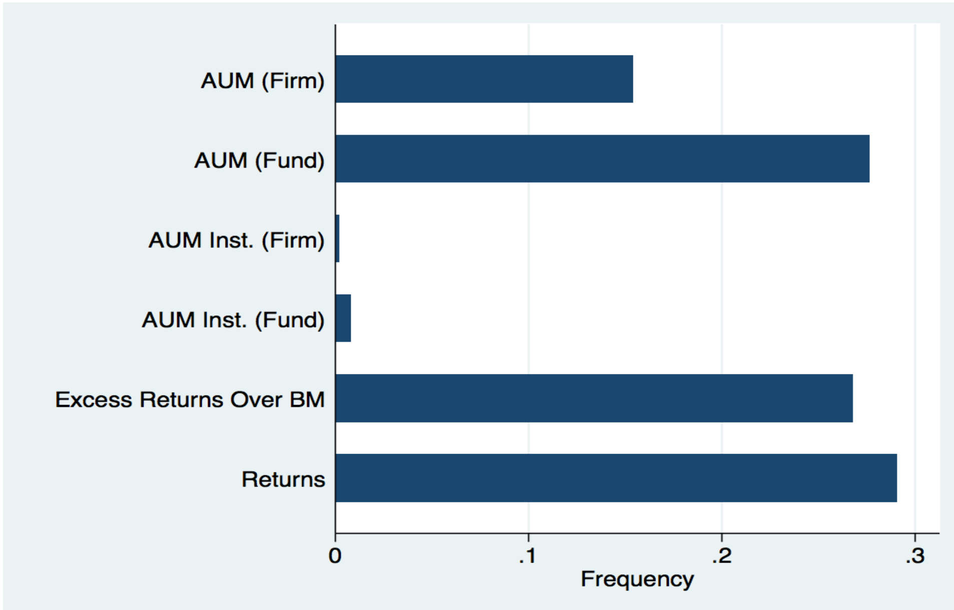


Figure 1: Screen Frequency by Type
This figure plots the frequency of each type of screen

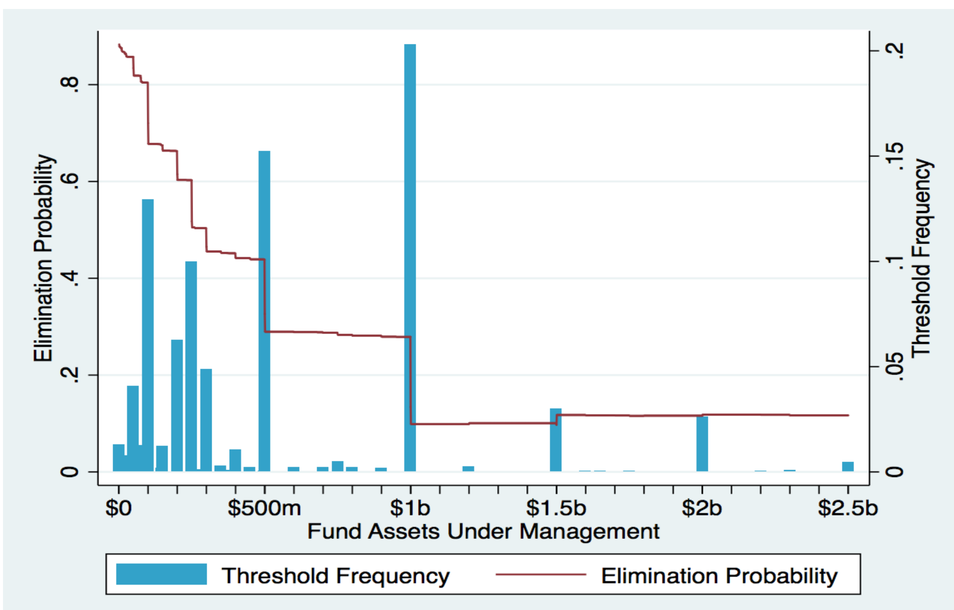


Figure 2: Fund-Level AUM Thresholds
This figure plots the frequency of fund-level assets under management thresholds and the probability of elimination conditional on a fund-level assets under management screen being used

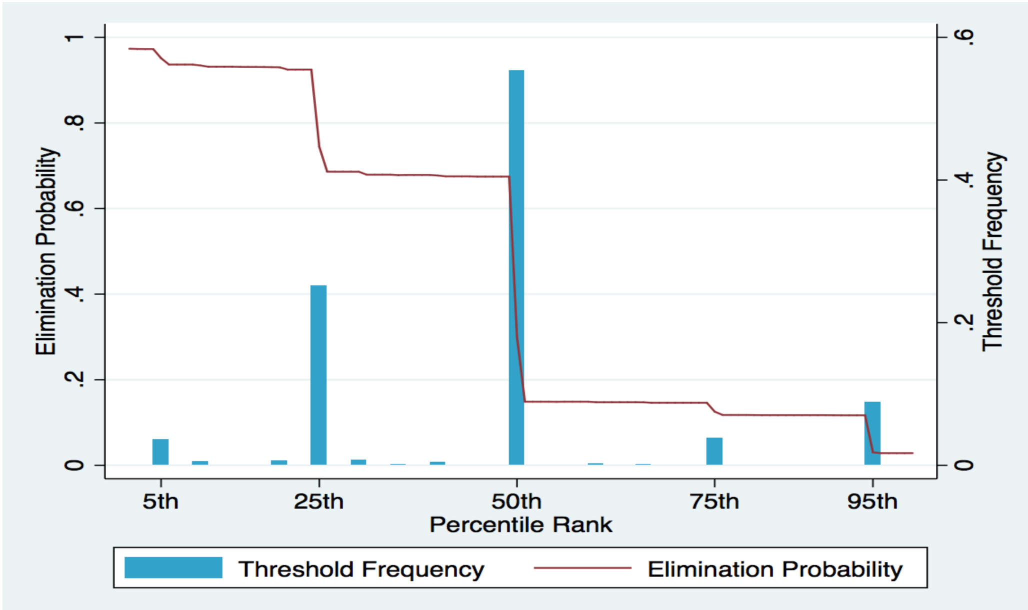


Figure 3: Percentile Rank Thresholds
This figure plots the frequency of return percentile rank thresholds and the probability of elimination conditional on a percentile rank threshold being used.

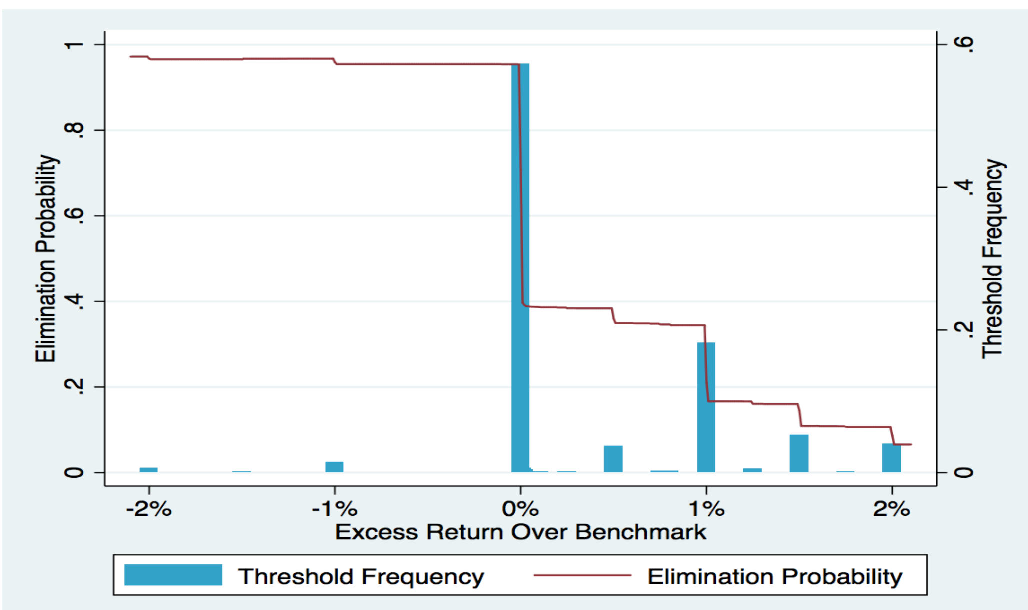


Figure 4: Excess Return Value Thresholds
This figure plots the frequency of numerical value excess return thresholds and the probability of elimination conditional on an excess return numerical threshold being used. Range: -2.1% to 2.1%.

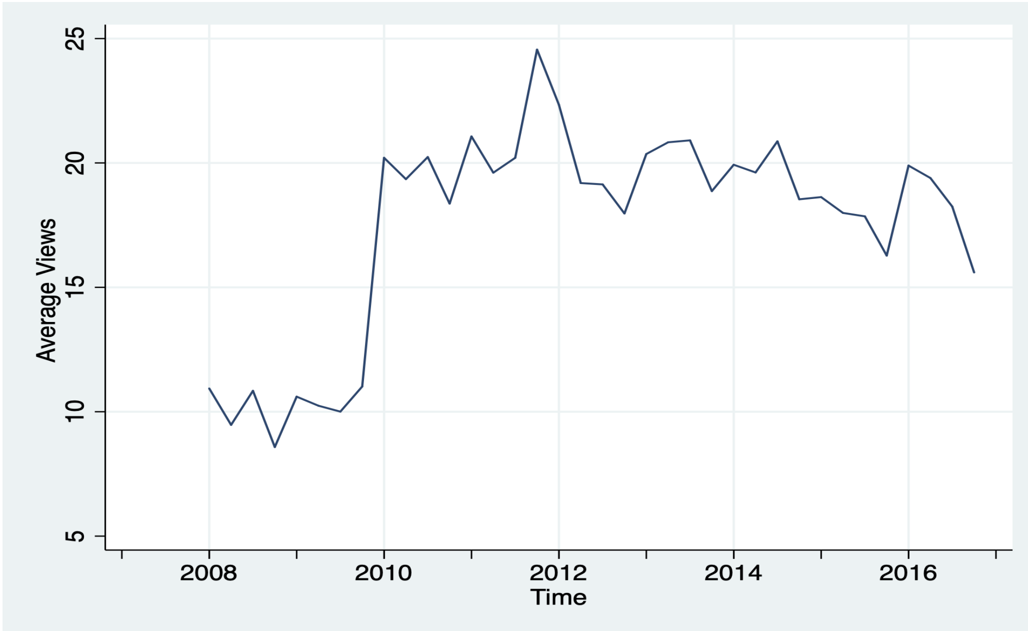


Figure 5: Average Views Over Time
This figure plots the average number of views per fund each quarter

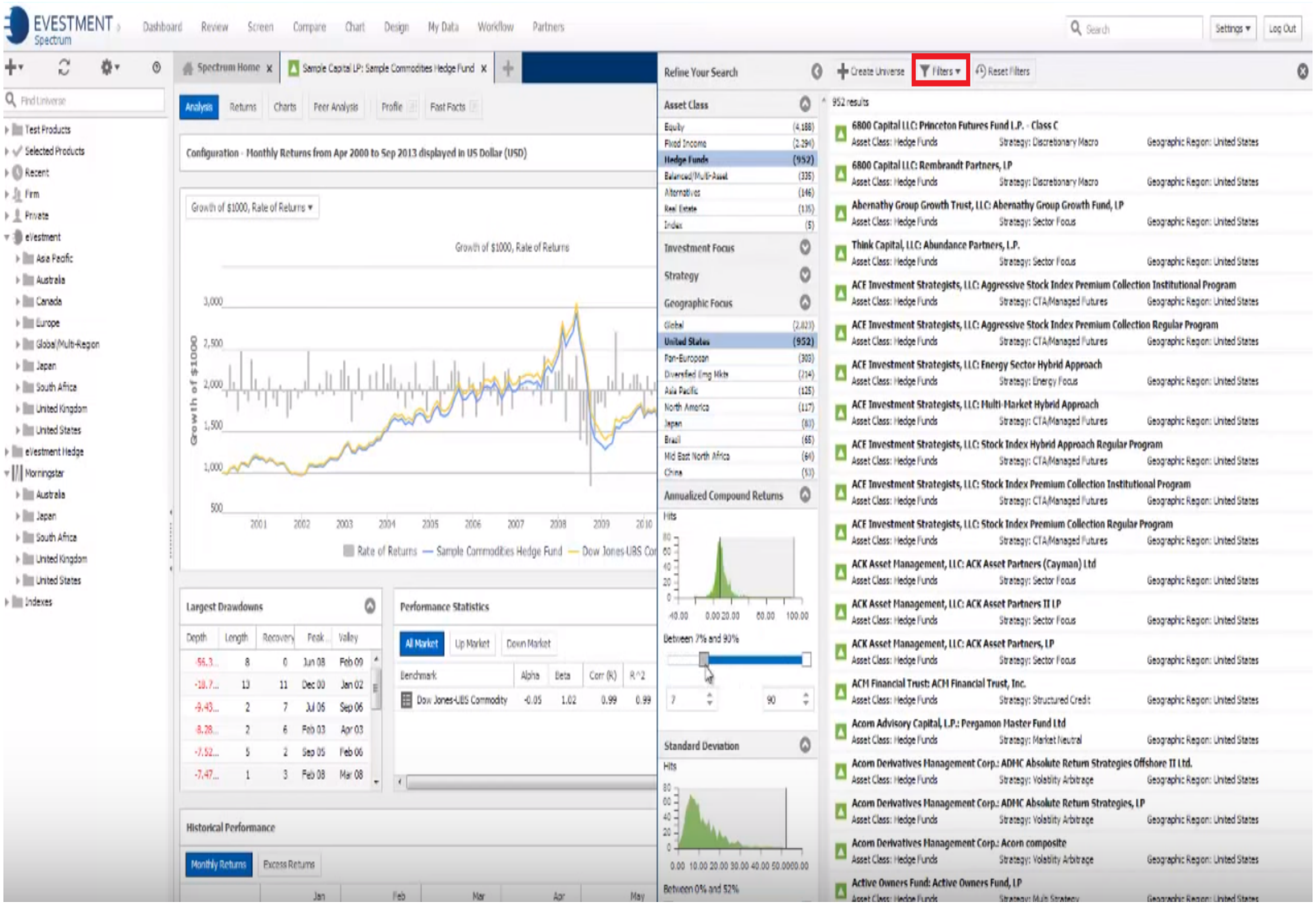


Figure 6: Screenshot of the eVestment platform

This figure presents a screenshot of the eVestment website.

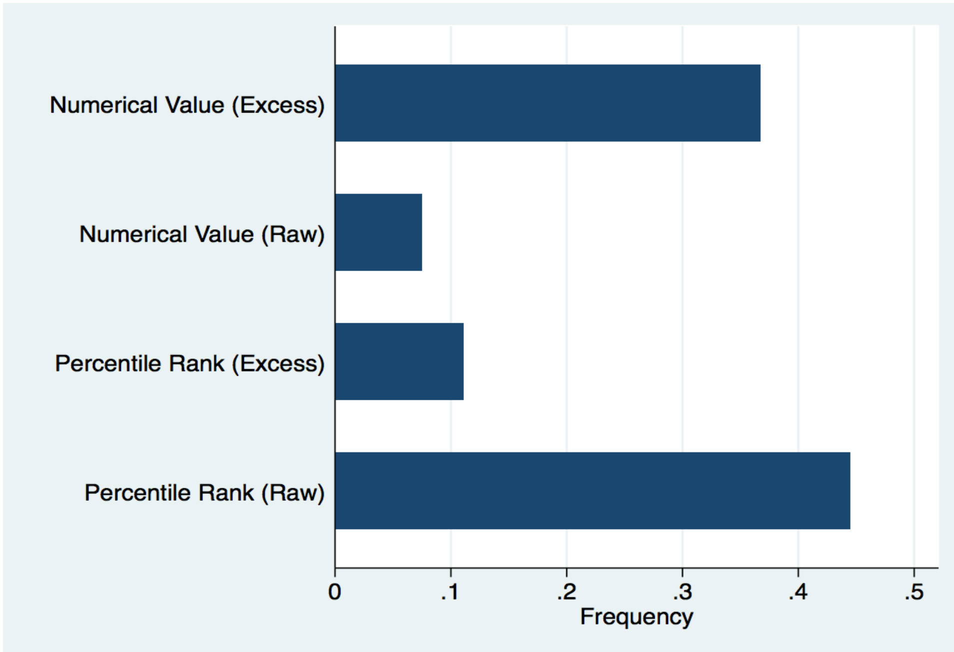


Figure 7: Return Threshold Types
This figure plots the frequency of return threshold types.

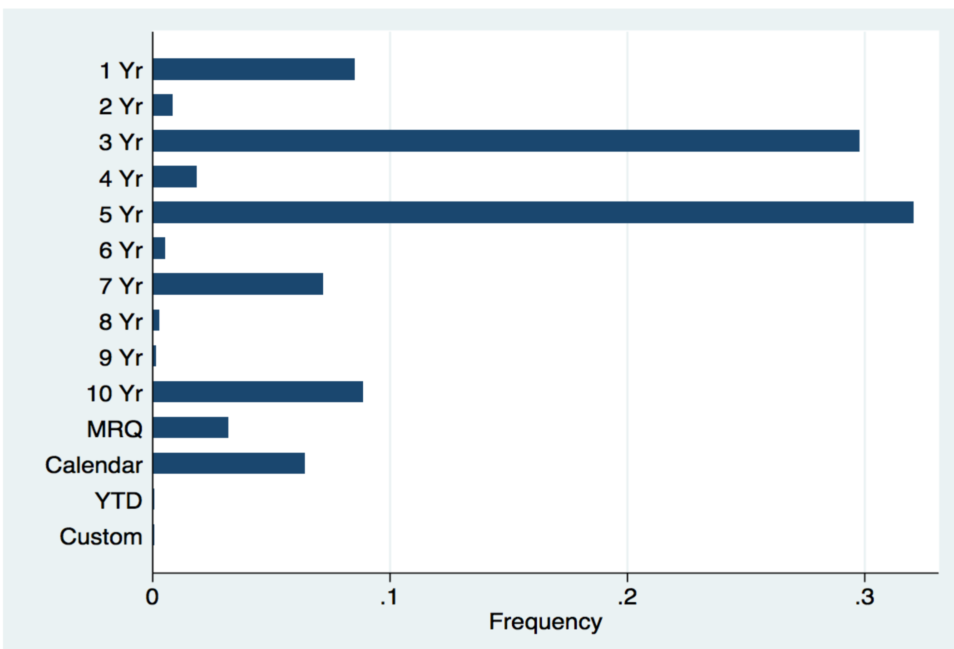


Figure 8: Return Time Horizon
This figure plots the frequency of the time horizons used for return screens.

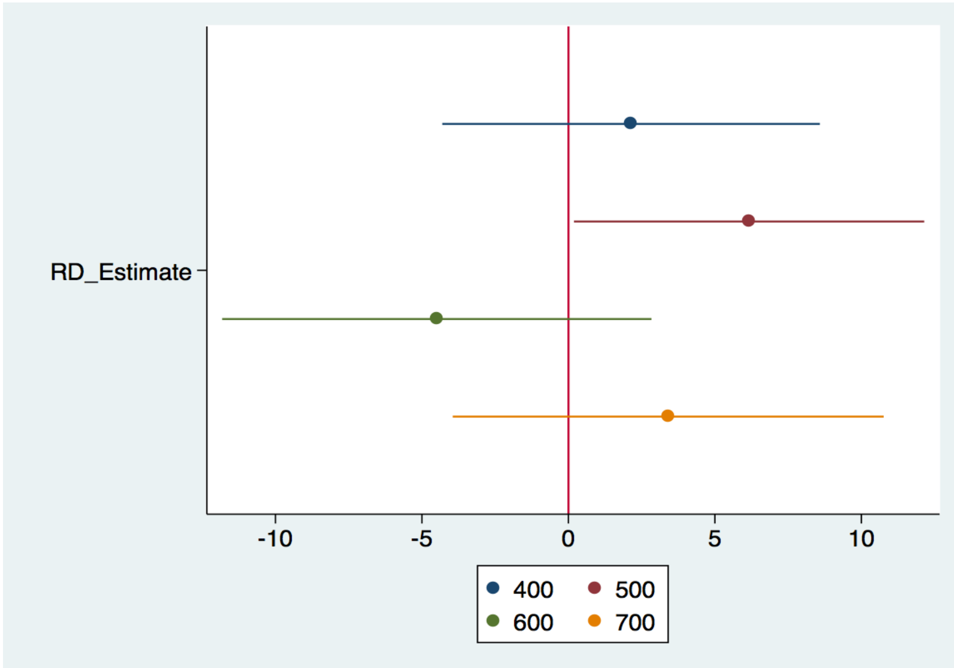


Figure 9: Page Views RDD Estimates
This figure plots the coefficient estimates capturing the effect of being above specific AUM thresholds on future fund attention. The dependent variable is page views over the next four quarters. The regression specification is a regression-discontinuity design following Equation (2.3). Coefficient estimates and the 99% confidence intervals are plotted for the \$500MM threshold and three placebo thresholds: \$400MM, \$600MM, and \$700MM.

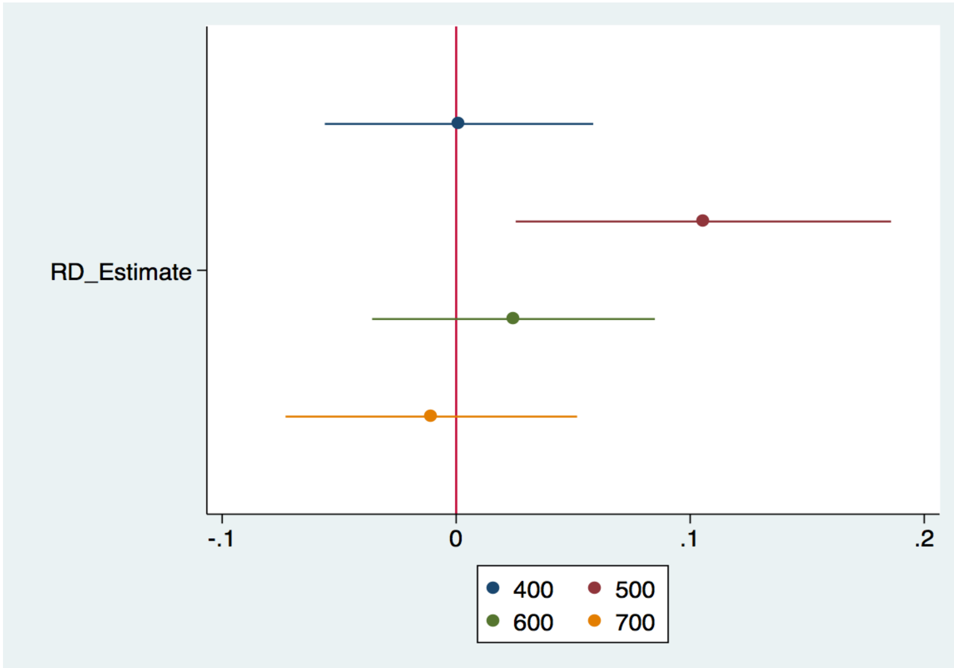


Figure 10: Flows RDD Estimates
This figure plots the coefficient estimates capturing the effect of being above specific AUM thresholds on future fund flows. The dependent variable is percentage fund flow over the next four quarters. The regression specification is a regression-discontinuity design following Equation (2.3). Coefficient estimates and the 99% confidence intervals are plotted for the $500MM threshold and three placebo thresholds: $400MM, $600MM, and $700MM.

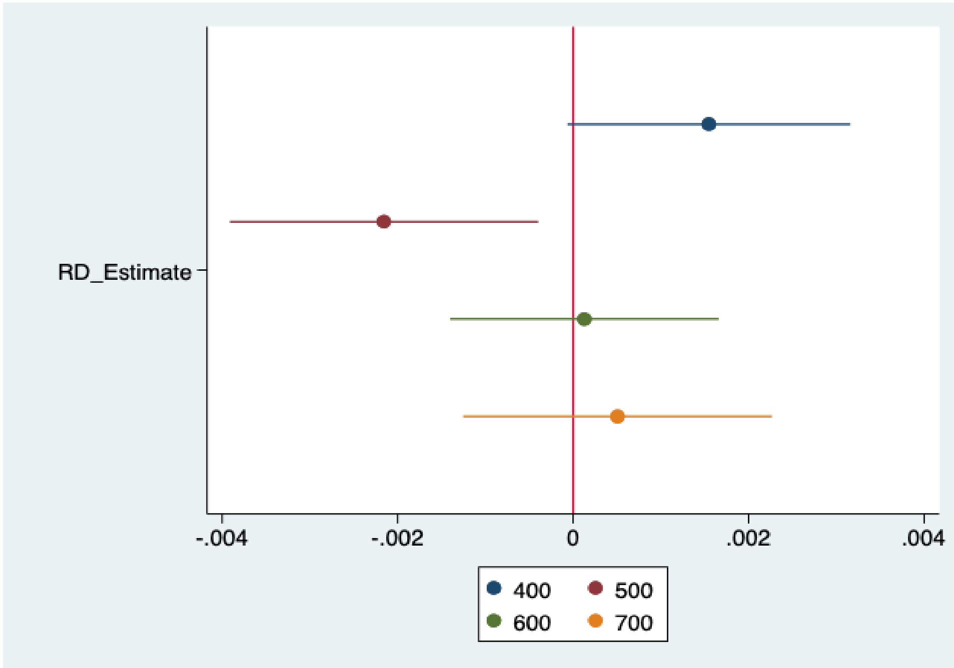


Figure 11: Returns RDD Estimates
This figure plots the coefficient estimates capturing the effect of being above specific AUM thresholds on future fund returns. The dependent variable is fund returns over the next quarter net of the average return for the fund's style group. The regression specification is a regression-discontinuity design following Equation (2.3). Coefficient estimates and the 99% confidence intervals are plotted for the \$500MM threshold and three placebo thresholds: \$400MM, \$600MM, and \$700MM.

Table 1: Summary Statistics

This table presents the sample summary statistics. $Views_{q+1-4}$ is the fund's total number of page views from quarter $q+1$ to $q+4$, $Flow_{q+1-4}$ is the fund's flow from quarter $q+1$ to $q+4$ as a percentage of assets under management in quarter q, Ret_{q+1} is the fund's return in quarter $q+1$, AUM_q is the fund's assets under management (in millions) at the end of quarter q, *ElimRate* is the probability a fund is eliminated by an AUM screen conditional on its assets under management in quarter q, Ret_q is the fund's return in quarter q, $Flow_q$ is the fund's flow in quarter q as a percentage of its assets under management in quarter $q-1$. Panel A presents summary statistics for the full sample. Panel B (Panel C) presents summary statistics for all funds with assets under management between \$450MM and \$500MM (\$500MM and \$550MM).

Panel A: Full Sample

Variable	Obs	Mean	Std. Dev.	P10	P50	P90
$Views_{q+1-4}$	48,723	61.991	81.149	5	32	158
$Flow_{q+1-4}$	60,238	.176	1.038	-.374	-.026	.644
Ret_{q+1}	78,539	.023	.073	-.06	.02	.113
AUM_q	82,866	518.087	620.826	7.16	248.037	1,546.793
ElimRate	78,841	.511	.281	.116	.504	.867
Ret_q	82,866	.023	.073	-.06	.02	.112
$Flow_q$	73,708	.003	.229	-.123	-.005	.128

Panel B: Funds With AUM \$450MM-\$500MM

Variable	Obs	Mean	Std. Dev.	P10	P50	P90
$Views_{q+1-4}$	1,198	60.598	72.114	6	35	144
$Flow_{q+1-4}$	1,510	.025	.604	-.383	-.06	.424
Ret_{q+1}	1,815	.025	.07	-.05	.021	.11
AUM_q	1,909	474.148	14.796	453.949	474.02	494.65
ElimRate	1,909	.439	0	.439	.439	.439
Ret_q	1,909	.026	.07	-.052	.022	.115
$Flow_q$	1,780	.016	.192	-.102	-.006	.13

Panel C: Funds With AUM \$500MM-\$550MM

Variable	Obs	Mean	Std. Dev.	P10	P50	P90
$Views_{q+1-4}$	1,103	68.766	86.904	8	40	169
$Flow_{q+1-4}$	1,362	.083	.728	-.377	-.032	.585
Ret_{q+1}	1,686	.017	.07	-.067	.018	.101
AUM_q	1,759	524.842	14.48	504.96	525	545
ElimRate	1,759	.29	.003	.29	.29	.29
Ret_q	1,759	.024	.065	-.044	.021	.107
$Flow_q$	1,630	.012	.181	-.107	-.004	.137

Table 2: Page Views and Elimination Rate

This table presents regression results examining the relationship between fund elimination rates and future page views. The dependent variable is the fund's page views over next four quarters. *ElimRate* is the probability a fund is eliminated by an AUM screen conditional on its assets under management in quarter q. $Log(AUM_q)$ is the logarithm of the fund's assets under management (in millions) at the end of quarter q. Ret_q is the fund's past quarter return. Standard errors are double clustered at the fund and year-quarter level.

	(1)	(2)	(3)	(4)	(5)
	Views	Views	Views	Views	Views
ElimRate	-33.628***	-30.792***	-32.106***	-36.209***	-36.331***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Log(AUM_q)$	12.977***	12.189***	12.286***	11.547***	11.635***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Ret_q					60.018***
					(0.00)
Observations	45946	46901	45946	45782	45782
r2	0.839	0.842	0.841	0.853	0.853
Absorbed FE	fund, style	fund, yq	fund, style, yq	fund, style×yq	fund, style×yq
Clustered by	fund, yq	fund, yq	fund, yq	fund, yq	fund, yq

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Flows and Elimination Rate

This table presents regression results examining the relationship between fund elimination rates and future flows. The dependent variable is the fund's percentage flow over next four quarters. *ElimRate* is the probability a fund is eliminated by an AUM screen conditional on its assets under management in quarter q. $Log(AUM_q)$ is the logarithm of the fund's assets under management (in millions) at the end of quarter q. Ret_q is the fund's past quarter return. Standard errors are double clustered at the fund and year-quarter level.

	(1)	(2)	(3)	(4)	(5)
	Flow	Flow	Flow	Flow	Flow
ElimRate	-0.611***	-0.577***	-0.573***	-0.570***	-0.572***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Log(AUM_q)$	-0.670***	-0.697***	-0.690***	-0.699***	-0.697***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Ret_q					1.214***
					(0.00)
Observations	57782	59029	57782	57621	57621
r2	0.444	0.447	0.448	0.468	0.469
Absorbed FE	fund, style	fund, yq	fund, style, yq	fund, style×yq	fund, style×yq
Clustered by	fund, yq	fund, yq	fund, yq	fund, yq	fund, yq

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Page Views

Y=Page views over next four quarters. Columns (1)-(4) are OLS regressions with various sets of fixed effects. In columns (5)-(8), treatment and control are matched on style x quarter buckets (column (5)), style x quarter buckets and last quarter's flow (column (6)), style x quarter buckets and last quarter's return (column (7)), style x quarter buckets, last quarter's flow and last quarter's return (column (8)). The style x quarter buckets are matched exactly and the Sturge's rule is used to coarsen last quarter's flow and last quarter's return. Standard errors are clustered at the year-quarter level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Views	Views	Views	Views	Views	Views	Views	Views
Above 500 Threshold	8.423***	8.482***	9.233***	10.291***	9.835**	9.912***	11.217***	10.231***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)	(0.01)	(0.01)	(0.00)
Observations	2235	2301	2235	1986	1925	1384	1707	1174
r2	0.208	0.023	0.224	0.311	0.003	0.005	0.004	0.005
Absorbed FE	style	yq	style, yq	style×yq				
Matching variables:					style×yq	style×yq, $Flow_q$	style×yq, Ret_q	style×yq, $Flow_q$, Ret_q
Clustered by	yq	yq	yq	yq	yq	yq	yq	yq

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Flows

Y=Flow over next four quarters. Columns (1)-(4) are OLS regressions with various sets of fixed effects. In columns (5)-(8), treatment and control are matched on style x quarter buckets (column (5)), style x quarter buckets and last quarter's flow (column (6)), style x quarter buckets and last quarter's return (column (7)), style x quarter buckets, last quarter's flow and last quarter's return (column (8)). The style x quarter buckets are matched exactly and the Sturge's rule is used to coarsen last quarter's flow and last quarter's return. Standard errors are clustered at the year-quarter level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Flow	Flow	Flow	Flow	Flow	Flow	Flow	Flow
Above 500 Threshold	0.065**	0.060*	0.066**	0.051	0.066**	0.070**	0.073**	0.088**
	(0.04)	(0.06)	(0.04)	(0.10)	(0.05)	(0.02)	(0.02)	(0.01)
Observations	2809	2872	2809	2575	2400	1730	2104	1457
r2	0.077	0.021	0.097	0.197	0.003	0.004	0.004	0.007
Absorbed FE	style	yq	style, yq	style×yq				
Matching variables:					style×yq	style×yq, $Flow_q$	style×yq, Ret_q	style×yq, $Flow_q$, Ret_q
Clustered by	yq	yq	yq	yq	yq	yq	yq	yq

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Returns

Y=Return over next quarter. Columns (1)-(4) are OLS regressions with various sets of fixed effects. In columns (5)-(8), treatment and control are matched on style x quarter buckets (column (5)), style x quarter buckets and last quarter's flow (column (6)), style x quarter buckets and last quarter's return (column (7)), style x quarter buckets, last quarter's flow and last quarter's return (column (8)). The style x quarter buckets are matched exactly and the Sturge's rule is used to coarsen last quarter's flow and last quarter's return. Standard errors are clustered at the year-quarter level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}
Above 500 Threshold	-0.009***	-0.005**	-0.005**	-0.001	-0.002*	-0.002**	-0.001	-0.002**
	(0.00)	(0.01)	(0.01)	(0.17)	(0.08)	(0.04)	(0.18)	(0.04)
Observations	3422	3501	3422	3144	2946	2129	2585	1782
r2	0.021	0.576	0.604	0.920	0.000	0.000	0.000	0.000
Absorbed FE	style	yq	style, yq	style×yq				
Matching variables:					style×yq	style×yq, $Flow_q$	style×yq, Ret_q	style×yq, $Flow_q$, Ret_q
Clustered by	yq	yq	yq	yq	yq	yq	yq	yq

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A Appendix. Tables and Figures

Table A.1: Page Views

Y=Page views over next four quarters. Columns (1)-(4) are OLS regressions with various sets of fixed effects. In columns (5)-(8), treatment and control are matched on style x quarter buckets (column (5)), style x quarter buckets and last quarter's flow (column (6)), style x quarter buckets and last quarter's return (column (7)), style x quarter buckets, last quarter's flow and last quarter's return (column (8)). The style x quarter buckets are matched exactly and the Sturge's rule is used to coarsen last quarter's flow and last quarter's return. Standard errors are clustered at the year-quarter level.

	(1) Views	(2) Views	(3) Views	(4) Views	(5) Views	(6) Views	(7) Views	(8) Views
Above 400 Threshold	2.698	4.567	3.462	0.615	0.521	3.965	-0.960	0.946
	(0.34)	(0.10)	(0.22)	(0.87)	(0.88)	(0.23)	(0.78)	(0.78)
Observations	3019	3091	3019	2738	2674	2032	2296	1599
r2	0.237	0.014	0.246	0.322	0.000	0.001	0.000	0.000
Absorbed FE	style	yq	style, yq	style×yq				
Matching variables:					style×yq	style×yq, $Flow_q$	style×yq, Ret_q	style×yq, $Flow_q$, Ret_q
Clustered by	yq	yq	yq	yq	yq	yq	yq	yq

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1) Views	(2) Views	(3) Views	(4) Views	(5) Views	(6) Views	(7) Views	(8) Views
Above 600 Threshold	0.711	-0.531	0.440	-2.635	-2.604	-4.882	-4.719	-9.324**
	(0.76)	(0.82)	(0.84)	(0.44)	(0.39)	(0.21)	(0.30)	(0.05)
Observations	1946	1998	1946	1679	1610	1131	1335	897
r2	0.217	0.025	0.234	0.350	0.000	0.001	0.001	0.003
Absorbed FE	style	yq	style, yq	style×yq				
Matching variables:					style×yq	style×yq, $Flow_q$	style×yq, Ret_q	style×yq, $Flow_q$, Ret_q
Clustered by	yq	yq	yq	yq	yq	yq	yq	yq

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Page Views continued...

	(1) Views	(2) Views	(3) Views	(4) Views	(5) Views	(6) Views	(7) Views	(8) Views
Above 700 Threshold	10.103**	9.321*	11.012**	8.316	8.044	0.259	10.153	-0.365
	(0.03)	(0.05)	(0.02)	(0.14)	(0.20)	(0.97)	(0.12)	(0.96)
Observations	1649	1686	1649	1419	1337	910	1145	728
r2	0.235	0.027	0.254	0.366	0.002	0.000	0.003	0.000
Absorbed FE	style	yq	style, yq	style×yq				
Matching variables:					style×yq	style×yq, $Flow_q$	style×yq, Ret_q	style×yq, $Flow_q$, Ret_q
Clustered by	yq	yq	yq	yq	yq	yq	yq	yq

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1) Views	(2) Views	(3) Views	(4) Views	(5) Views	(6) Views	(7) Views	(8) Views
Above 1000 Threshold	2.940	1.665	2.117	3.801	0.704	-3.375	-0.133	-1.173
	(0.61)	(0.81)	(0.72)	(0.68)	(0.92)	(0.69)	(0.99)	(0.91)
Observations	1107	1132	1104	837	773	549	607	403
r2	0.338	0.031	0.356	0.449	0.000	0.000	0.000	0.000
Absorbed FE	style	yq	style, yq	style×yq				
Matching variables:					style×yq	style×yq, $Flow_q$	style×yq, Ret_q	style×yq, $Flow_q$, Ret_q
Clustered by	yq	yq	yq	yq	yq	yq	yq	yq

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Flows

Y=Flow over next four quarters. Columns (1)-(4) are OLS regressions with various sets of fixed effects. In columns (5)-(8), treatment and control are matched on style x quarter buckets (column (5)), style x quarter buckets and last quarter's flow (column (6)), style x quarter buckets and last quarter's return (column (7)), style x quarter buckets, last quarter's flow and last quarter's return (column (8)). The style x quarter buckets are matched exactly and the Sturge's rule is used to coarsen last quarter's flow and last quarter's return. Standard errors are clustered at the year-quarter level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Flow	Flow	Flow	Flow	Flow	Flow	Flow	Flow
Above 400 Threshold	-0.010	-0.012	-0.009	-0.008	-0.009	-0.007	-0.018	-0.021
	(0.61)	(0.51)	(0.60)	(0.68)	(0.65)	(0.73)	(0.38)	(0.34)
Observations	3724	3795	3724	3410	3290	2481	2812	1931
r2	0.051	0.009	0.059	0.166	0.000	0.000	0.000	0.000
Absorbed FE	style	yq	style, yq	style×yq				
Matching variables:					style×yq	style×yq, $Flow_q$	style×yq, Ret_q	style×yq, $Flow_q$, Ret_q
Clustered by	yq	yq	yq	yq	yq	yq	yq	yq

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Flow	Flow	Flow	Flow	Flow	Flow	Flow	Flow
Above 600 Threshold	0.000	0.003	-0.003	-0.010	-0.001	0.005	-0.011	0.005
	(0.98)	(0.89)	(0.89)	(0.74)	(0.98)	(0.85)	(0.72)	(0.88)
Observations	2342	2393	2342	2078	1934	1360	1590	1073
r2	0.024	0.012	0.040	0.200	0.000	0.000	0.000	0.000
Absorbed FE	style	yq	style, yq	style×yq				
Matching variables:					style×yq	style×yq, $Flow_q$	style×yq, Ret_q	style×yq, $Flow_q$, Ret_q
Clustered by	yq	yq	yq	yq	yq	yq	yq	yq

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Flow Views continued...

	(1) Flow	(2) Flow	(3) Flow	(4) Flow	(5) Flow	(6) Flow	(7) Flow	(8) Flow
Above 700 Threshold	0.001	0.004	0.008	-0.010	0.014	-0.021	0.010	-0.010
	(0.96)	(0.81)	(0.64)	(0.63)	(0.56)	(0.51)	(0.70)	(0.72)
Observations	1950	1990	1950	1707	1559	1048	1323	828
r2	0.054	0.022	0.076	0.236	0.000	0.000	0.000	0.000
Absorbed FE	style	yq	style, yq	style×yq				
Matching variables:					style×yq	style×yq, $Flow_q$	style×yq, Ret_q	style×yq, $Flow_q$, Ret_q
Clustered by	yq	yq	yq	yq	yq	yq	yq	yq

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1) Flow	(2) Flow	(3) Flow	(4) Flow	(5) Flow	(6) Flow	(7) Flow	(8) Flow
Above 1000 Threshold	-0.022	-0.032	-0.025	0.019	0.008	0.065	0.005	0.075
	(0.35)	(0.18)	(0.28)	(0.55)	(0.81)	(0.17)	(0.89)	(0.17)
Observations	1314	1343	1311	1046	907	637	705	462
r2	0.087	0.030	0.109	0.349	0.000	0.004	0.000	0.004
Absorbed FE	style	yq	style, yq	style×yq				
Matching variables:					style×yq	style×yq, $Flow_q$	style×yq, Ret_q	style×yq, $Flow_q$, Ret_q
Clustered by	yq	yq	yq	yq	yq	yq	yq	yq

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Returns

Y=Return over next quarter. Columns (1)-(4) are OLS regressions with various sets of fixed effects. In columns (5)-(8), treatment and control are matched on style x quarter buckets (column (5)), style x quarter buckets and last quarter's flow (column (6)), style x quarter buckets and last quarter's return (column (7)), style x quarter buckets, last quarter's flow and last quarter's return (column (8)). The style x quarter buckets are matched exactly and the Sturge's rule is used to coarsen last quarter's flow and last quarter's return. Standard errors are clustered at the year-quarter level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}
Above 400 Threshold	-0.002	-0.000	-0.001	-0.000	-0.001	0.000	-0.001	-0.001
	(0.38)	(0.81)	(0.40)	(0.87)	(0.41)	(0.89)	(0.42)	(0.37)
Observations	4592	4687	4592	4261	4066	3069	3484	2393
r2	0.030	0.511	0.546	0.901	0.000	0.000	0.000	0.000
Absorbed FE	style	yq	style, yq	style×yq				
Matching variables:					style×yq	style×yq, $Flow_q$	style×yq, Ret_q	style×yq, $Flow_q$, Ret_q
Clustered by	yq	yq	yq	yq	yq	yq	yq	yq

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}
Above 600 Threshold	0.005*	0.001	0.000	-0.000	0.001	0.001	0.001	0.002
	(0.09)	(0.43)	(0.86)	(0.87)	(0.70)	(0.35)	(0.55)	(0.31)
Observations	2853	2932	2853	2548	2371	1669	1968	1329
r2	0.023	0.559	0.589	0.920	0.000	0.000	0.000	0.000
Absorbed FE	style	yq	style, yq	style×yq				
Matching variables:					style×yq	style×yq, $Flow_q$	style×yq, Ret_q	style×yq, $Flow_q$, Ret_q
Clustered by	yq	yq	yq	yq	yq	yq	yq	yq

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}	Ret_{q+1}
Above 700 Threshold	-0.001	-0.002	-0.001	0.002	0.002	0.001	0.002	0.002
	(0.84)	(0.34)	(0.59)	(0.20)	(0.31)	(0.66)	(0.11)	(0.34)
Observations	2424	2472	2424	2156	1943	1305	1651	1038
r2	0.031	0.553	0.588	0.908	0.000	0.000	0.000	0.000
Absorbed FE	style	yq	style, yq	style×yq				
Matching variables:					style×yq	style×yq, $Flow_q$	style×yq, Ret_q	style×yq, $Flow_q$, Ret_q
Clustered by	yq	yq	yq	yq	yq	yq	yq	yq

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

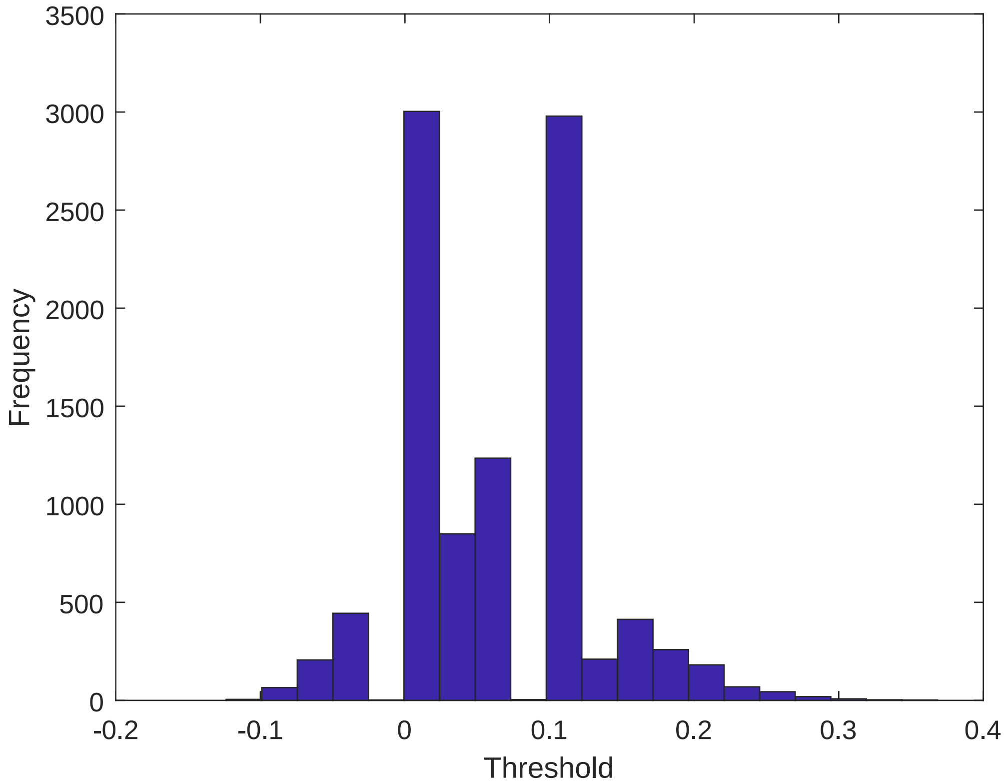
Table A.4: Risk Measures

This table presents results examining differential risk-taking around the $500MM threshold. The dependent variables are various measures of fund risk. In column (1), the dependent variable is the average variance of the fund's holdings. Columns (2)-(8) are similar except with different risk measures (skewness, kurtosis, market capitalization, book-to-market, the market beta, size beta, and HML beta, respectively). Treatment and control are matched on style x quarter buckets, last quarter's flow and last quarter's return. The style x quarter buckets are matched exactly and the Sturge's rule is used to coarsen last quarter's flow and last quarter's return. Standard errors are clustered at the year-quarter level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	variance_score	skewness_score	kurtosis_score	size_score	btm_score	beta_mkt	beta_smb	beta_hml
Above 500 Threshold	-0.001	0.004	-0.001	0.001	0.001	0.002	0.021*	-0.003
	(0.90)	(0.46)	(0.85)	(0.93)	(0.78)	(0.88)	(0.10)	(0.83)
Observations	1085	1085	1085	1085	1085	1085	1085	1085
r2	0.000	0.001	0.000	0.000	0.000	0.000	0.001	0.000
Clustered by	yq	yq	yq	yq	yq	yq	yq	yq

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure A.1: Simulated Histogram



This figure plots the simulated frequency of cutoff thresholds from 10,000 repetitions. The investor observes $s_i = \alpha_i + \varepsilon_i$ where $\alpha_i, \varepsilon_i \sim N(0, 0.2^2)$. We set $A = 1$, $K = 10^{-8}$ and $L = 10^{-8}$ and assume that $Ref_1 = 0$ and $Ref_2 = 0.1$.

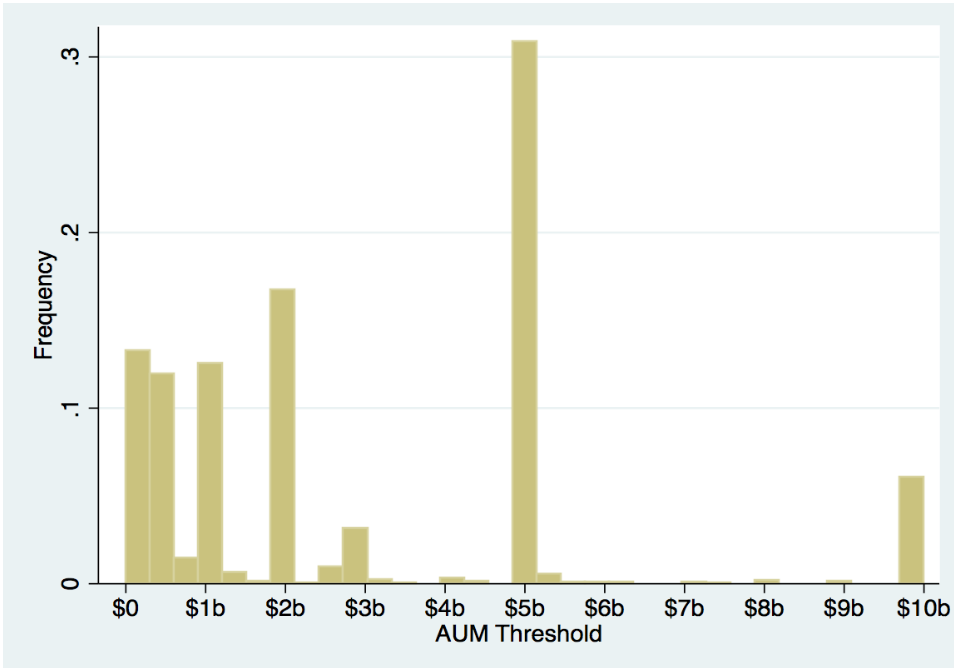


Figure A.2: Firm Level AUM Thresholds
This figure plots the frequency of firm level assets under management thresholds over the range $0 to $10B

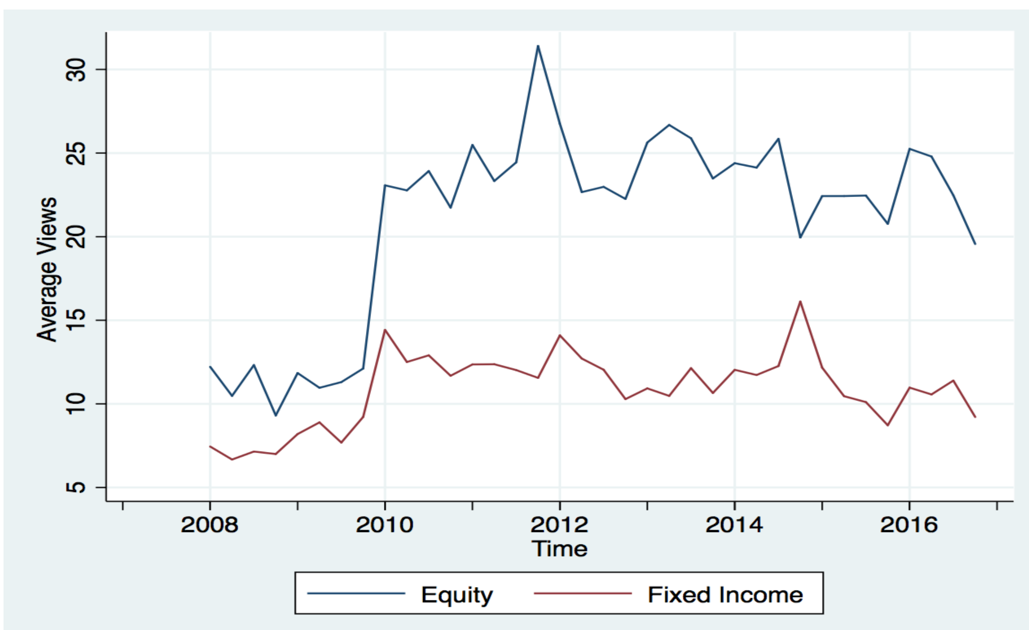


Figure A.3: Average Views Over Time
This figure plots the average number of views per fund each quarter. Views for equity and fixed income funds are plotted separately.

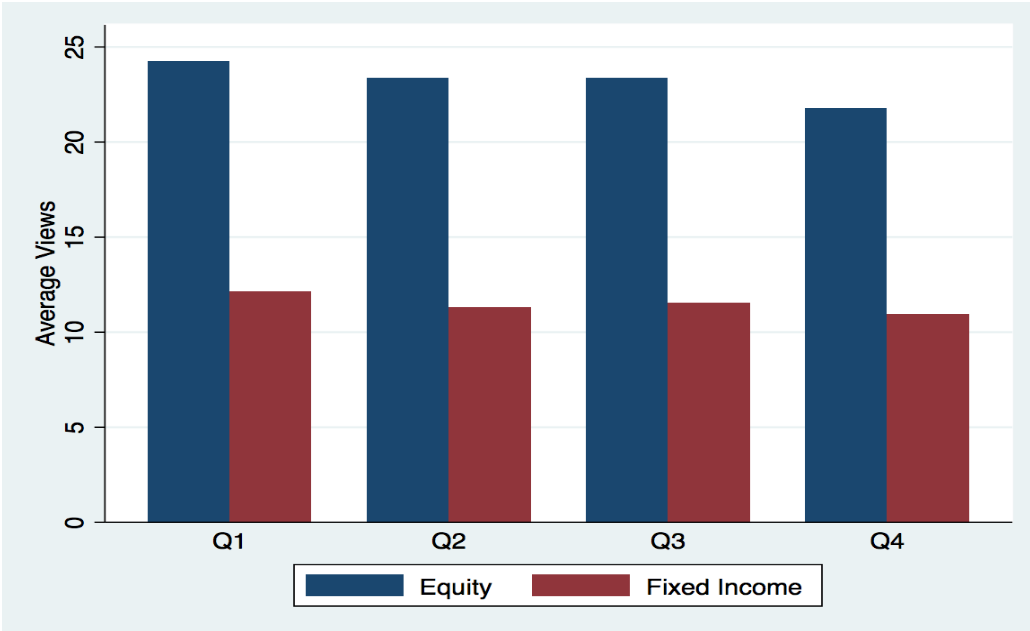


Figure A.4: Average Views Per Quarter
This figure plots the average number of views per fund by quarter. Views for equity and fixed income funds are plotted separately.

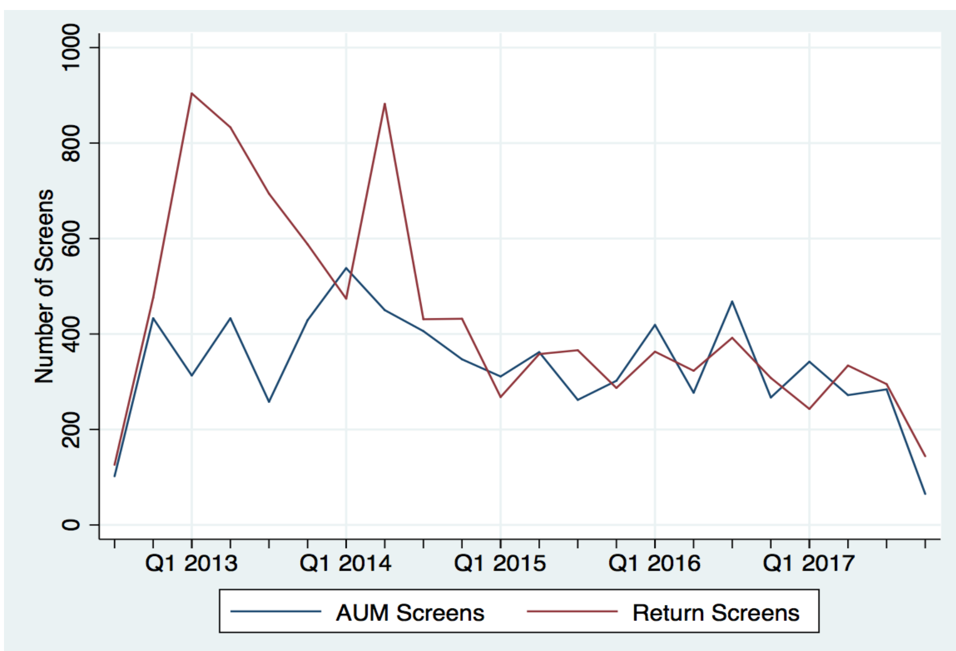


Figure A.5: Screen Frequency Over Time
This figure plots the frequency of screens over time

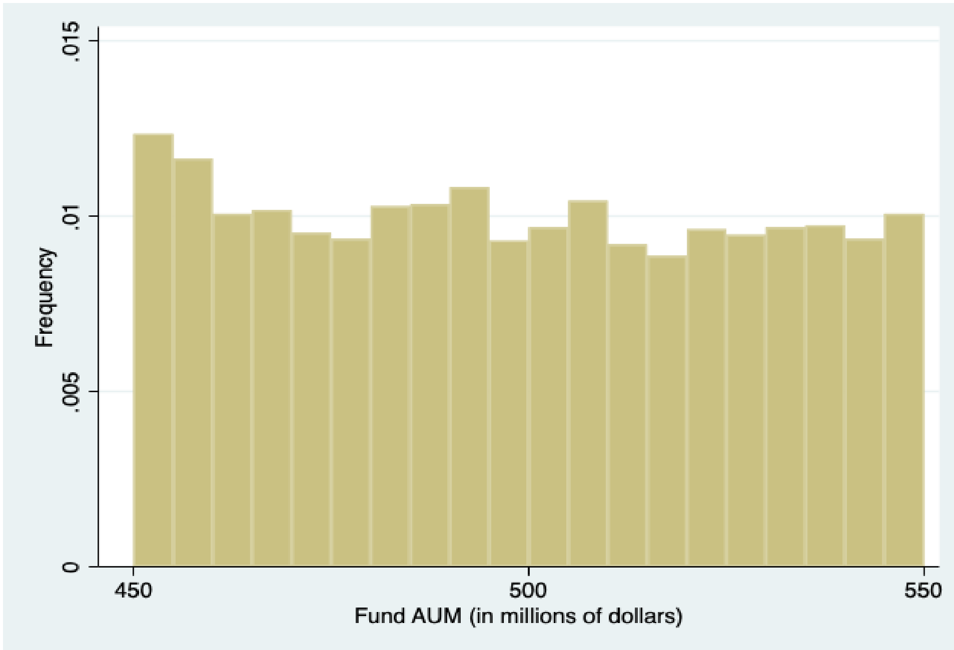


Figure A.6: Histogram of Fund Assets Under Management Near $500MM
This figure plots the empirical frequency of funds assets under management for funds with assets under management between $450MM and $550MM.

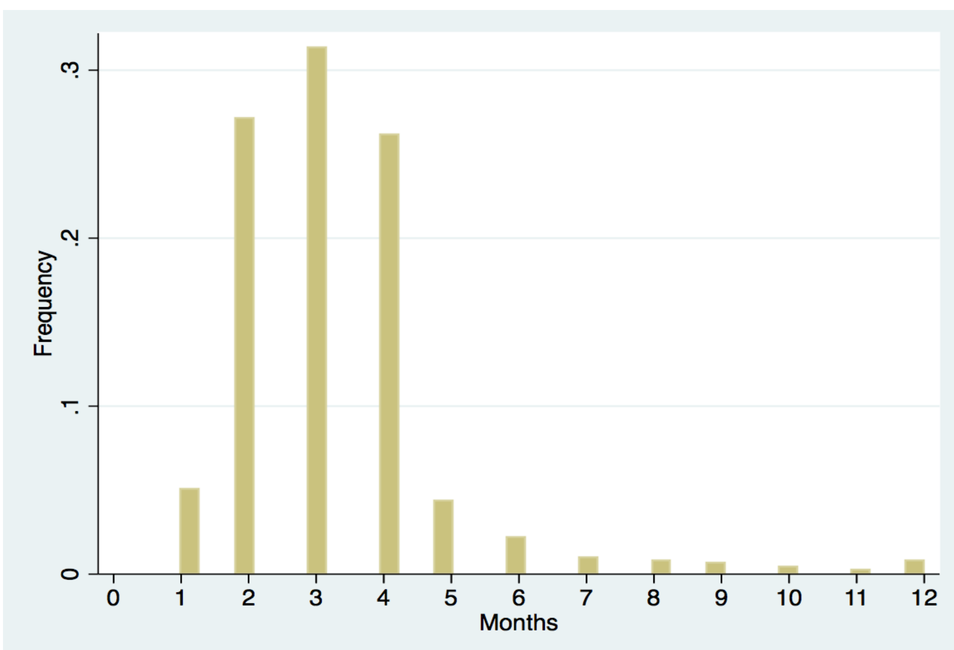


Figure A.7: Time Between Screen Date and Performance "as of" Date
This figure plots the time (in months) between the date of the screen and the performance "as of" date.

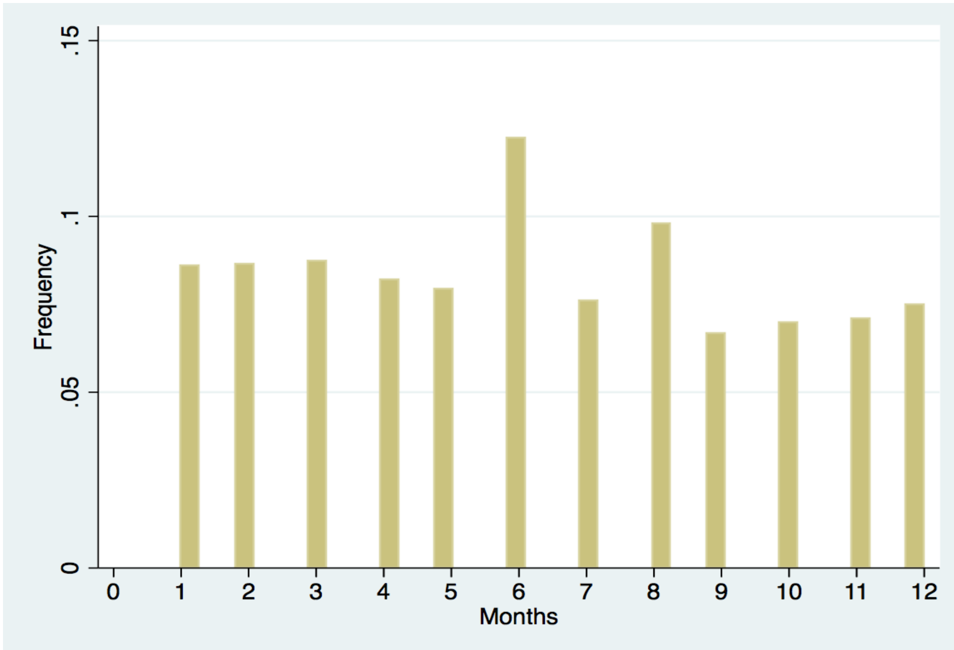


Figure A.8: Screen Frequency By Month
This figure plots the frequency of screens by month over the 2013-2016 period.

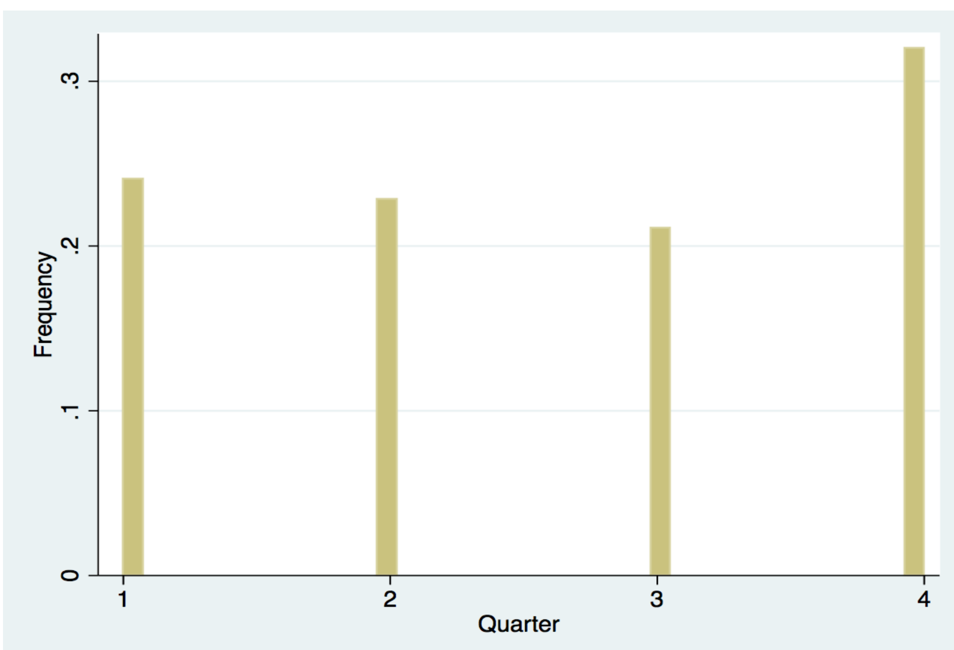


Figure A.9: Screen Frequency By Performance "as of" Quarter
This figure plots the frequency of screens by performance "as of" quarter over the 2013-2016 period.

B Theoretical Motivation

We present a simple model of fund selection in which the investor incurs a cost to evaluate fund manager skill. The model builds on the evaluation cost model of consumer choice in Hauser and Wernerfelt (1990). The purpose of the model is to illustrate some of the major trade-offs faced by investors in the fund selection process. Although we frame the discussion in terms of an investor selecting a fund for investment, the model potentially applies to a number of other settings in which an agent is selecting an item from a choice set and faces evaluation costs. All proofs are in the following section.

In our model economy, there exist I funds available to an investor, indexed by $i = 1, \cdots, I$. In the first stage of the fund selection process, the investor chooses the consideration set C, which is a subset of I funds, to maximize her utility, given by

$$U(C) = A \cdot \mathbb{E}\left[\max_{i \in C} \{\alpha_i\}\right] - n(C) \cdot K, \tag{B.1}$$

where A is the amount of assets to be invested in the chosen fund, $n(C)$ is the number of funds in the consideration set C and K is the cost incurred by the investor to evaluate each fund in the consideration set. For each fund in consideration set C, the investor incurs evaluation cost K to learn α_i, the skill of fund i. After evaluating all funds in the consideration set, she picks the fund with maximum skill.

Before she constructs a consideration set C, she observes a public signal s_i for each fund $i = 1, \cdots, I$. The public signal s_i is associated with the skill of fund manager i, α_i. In particular, the pair of (s_i, α_i) are drawn independently across funds from a common continuous distribution of $F(s, \alpha)$. Because we consider a continuous distribution of signals, the probability of tied signals is zero. Hence, we analyze the case where each signal is distinct.[22] Without loss of generality, we have $s_i > s_{i+1}$ for $i = 1, \cdots, I-1$. We assume that

[22]The case where signals tie can be handled by randomizing among funds with the same signal. The extended results are available upon request.

the public signals and alphas satisfy the conventional monotone likelihood ratio property.

Assumption 1. *Let $f(s,\alpha)$ denote the joint density function of s and α. The joint distribution of the public signal of s and the fund manager's skill α satisfies that when $s_H > s_L$ and $\alpha_H > \alpha_L$, it holds that*

$$\frac{f(s_H,\alpha_H)}{f(s_H,\alpha_L)} > \frac{f(s_L,\alpha_H)}{f(s_L,\alpha_L)}.$$

This assumption assures that the signals are informative. In particular, note that the inequality in Assumption 1 can be rearranged as

$$\frac{f(s_H,\alpha_H)}{f(s_H,\alpha_H)+f(s_H,\alpha_L)} > \frac{f(s_L,\alpha_H)}{f(s_L,\alpha_H)+f(s_L,\alpha_L)}.$$

Hence, Assumption 1 states that when an investor observes two signals of s_H and s_L such that $s_H > s_L$, the high signal of s_H implies it is more likely the fund manager has the high skill, α_H, relative to the low skill of α_L, than when the low signal of s_L is observed.[23] For the rest of this section, we take Assumption 1 as given.

The following property follows from Assumption 1.

Lemma B.1. *When $s_H > s_L$, it holds that*

$$\Pr(\alpha < \overline{\alpha}|s_L) > \Pr(\alpha < \overline{\alpha}|s_H).$$

The above lemma states that for a fund manager with a stronger signal of s_H, it is more likely that their skill is above a fixed level of $\overline{\alpha}$ than a fund manager with a weaker signal of s.

Before we proceed further, we provide some defense for the assumed information structure. In our model, there exists only one public signal of fund skill. However, this restriction can be partially justified by Blackwell (1951, 1953). The theorem by Blackwell states that when

[23]This type of assumption is widely used in literature and it is well known that this property holds for various families of distributions. The list of families with this property includes Exponential, Binomial, Poisson, and Normal.

an investor has an option to choose an information system generated by different signals, the investor always prefers an informative signal in the following sense.

Proposition B.1. *Consider a garbled version of signal s such that $g_i = f(s_i) + \varepsilon_{i,g}$, where $\varepsilon_{i,g}$ is any zero mean random variable independent of s_i for $i = 1, \cdots, I$. Then, when both signals of s and g are available, an investor does not use the information of the garbled signal g in constructing a consideration set.*

Resorting to the above proposition, the signal s can be interpreted to be chosen by an investor because of its informativeness of fund manager skill.

Next, we show that the optimal choice of consideration set can be simplified into a cutoff rule - selecting all funds with signals above a certain cutoff level and dropping all other funds. Define $C_s = \{j | s_i \leq s\}$, a consideration set constructed by a cutoff-rule with the threshold of s.

Proposition B.2. *There always exists an optimal consideration set C_s which maximizes $U(C)$ given by (B.1).*

The intuition of the above proposition follows. Let $\widetilde{C}$ denote any optimal consideration set. Note that $n\left(\widetilde{C}\right)$ is the number of funds in $\widetilde{C}$ and that $s_{n(\widetilde{C})}$ is the $n\left(\widetilde{C}\right)$ -th highest signal. Then, consider an alternative consideration set $C_{s_{n(\widetilde{C})}} = \left\{i | s_i \leq s_{n(\widetilde{C})}\right\}$. Noting that Lemma B.1 implies that $\Pr\left(\max_{i\in\widetilde{C}}\{\alpha_i\} \leq \overline{\alpha}\right) = \Pi_{i\in\widetilde{C}} \Pr(\alpha_i \leq \overline{\alpha})$ is larger than $\Pr\left(\max_{i\in C_{s_{n(\widetilde{C})}}}\{\alpha_i\} \leq \overline{\alpha}\right) = \Pi_{i\in C_{s_{n(\widetilde{C})}}} \Pr(\alpha_i \leq \overline{\alpha})$, we find that $\max_{i\in C_{s_{n(\widetilde{C})}}}\{\alpha_i\}$ first order stochastically dominates (FOSD) $\max_{i\in\widetilde{C}}\{\alpha_i\}$. Hence, from well known properties of FOSD, it follows that $\mathbb{E}\left[\max_{i\in\widetilde{C}}\{\alpha_i\}\right] \leq \mathbb{E}\left[\max_{i\in C_{s_{n(\widetilde{C})}}}\{\alpha_i\}\right]$, which, in conjunction with the number of funds in $\widetilde{C}$ being equal to the number of funds in $C_{s_{n(\widetilde{C})}}$, shows that the alternative consideration set $C_{s_{n(\widetilde{C})}}$ is as good as $\widetilde{C}$. In other words, a consideration set rule that sorts funds and uses a cut-off will give an expected payoff as good or better than any other consideration set with the same number of funds.

Next, we proceed to determine the optimal threshold for the cutoff rule on a given informative signal. The following lemma states that the marginal increase in the expected level of maximum skill decreases as an investor sequentially adds funds into her consideration set.

Lemma B.2. *It holds that*

$$\mathbb{E}\left[\max_{i\in C_{s_j}}\{\alpha_i\}\right]-\mathbb{E}\left[\max_{i\in C_{s_{j-1}}}\{\alpha_i\}\right]>\mathbb{E}\left[\max_{i\in C_{s_{j+1}}}\{\alpha_i\}\right]-\mathbb{E}\left[\max_{i\in C_{s_j}}\{\alpha_i\}\right]$$

for any $j=2,\cdots,I-1$.

From the above lemma on the decreasing marginal benefit and the assumption of the constant marginal evaluation cost K for each fund, we obtain the following proposition which pins down an optimal threshold.

Proposition B.3. *The consideration set of* $C_{s_{j^*}}=\{i|s_i\leq s_{j^*}\}$ *is optimal if the following conditions are met:*

$$\mathbb{E}\left[\max_{i\in C_{s_{j^*}}}\{\alpha_i\}\right]-\mathbb{E}\left[\max_{i\in C_{s_{j^*-1}}}\{\alpha_i\}\right]\geq K \text{ if } 2\leq j^*\leq I \text{ and}$$
$$\mathbb{E}\left[\max_{i\in C_{s_{j^*+1}}}\{\alpha_i\}\right]-\mathbb{E}\left[\max_{i\in C_{s_{j^*}}}\{\alpha_i\}\right]< K \text{ if } 1\leq j^*\leq I-1.$$

The following corollary summarizes the relation between the evaluation costs and an optimal level of cutoff signal.

Corollary B.1. *The optimal consideration set* $C_{s_{j^*}}$ *in Proposition B.2 satisfies the followings:*
(i) when K *is sufficiently small,* $C_{s_{j^*}}=C_{s_I}$*,*
(ii) when K *increases,* j^* *weakly decreases.*

The result (i) of the above proposition is interpreted as follows. When K is very small, an investment consultant would be mostly concerned about $\mathbb{E}\left[\max_{i\in C}\{\alpha_i\}\right]$, which increases as the consideration set C expands. Hence, she considers all funds. The result (ii) shows that as the evaluation cost increases, the investment consultant starts dropping funds with low signals one by one.

B.1 Cognitive Reference Number Bias

Thus far, we have examined the optimal consideration set construction when investors are subject to evaluation costs and verified that a consideration set made by a cutoff rule constitutes an optimal consideration set. We now restrict our attention to the consideration set $C_s = \{i|s_i \leq s\}$ and introduce a cognitive reference number bias to the choice of threshold of s for C_s. We assume the investor has H reference numbers, $Ref_1 = -\infty < \cdots < Ref_H = \infty$, which are indexed by $h = 1, \cdots, H$. The investor has a preference for these reference numbers and she receives a mental reward of L by choosing a reference number as the threshold value. Under this setup, the objective utility of (B.1) is modified as follows:

$$U_{Ref}\left(C_s\right) = A \cdot \mathbb{E}\left[\max_{i \in C_s}\left\{\alpha_i\right\}\right] - n\left(C_s\right) \cdot K + L \sum_{h=1}^{H} \mathbf{1}\left(Ref_h = s\right). \tag{B.2}$$

We are interested in how the optimal threshold decision in Proposition B.3 changes with the introduction of a mental reward for choosing a reference number. The next proposition shows how to find the optimal threshold with a reference number bias.

Proposition B.4. *The optimal threshold s which maximizes $U\left(C_s\right)$ given by (B.2) is either the solution s_{j^*} in Proposition B.2 or the reference numbers Ref_h or Ref_{h+1} such that $Ref_h \leq s_{j^*} \leq Ref_{h+1}$.*

The intuition of the above proposition is straightforward. From Lemma B.2, $A \cdot \mathbb{E}\left[\max_{i \in C_{s_j}}\left\{\alpha_i\right\}\right] - n\left(C_{s_j}\right) \cdot K$ is concave in j, and hence, reference numbers of Ref_h or Ref_{h+1} such that $Ref_h \leq s_j \leq Ref_{h+1}$ are always better than other non-adjacent references. Hence, it suffices to check the solution in Proposition B.2 and adjacent references.

B.2 Simulations and Testable Hypotheses

Next, we simulate the model and characterize the distribution of threshold values. We consider an investor who solves (B.2) by choosing the consideration set among 1,000 candidate funds. The investor observes $s_i = \alpha_i + \varepsilon_i$ where $\alpha_i, \varepsilon_i \sim N\left(0, 0.2^2\right)$. We set $A = 1,\ K = 10^{-8}$ and $L = 10^{-8}$ and assume the investor has reference numbers of $Ref_1 = 0$ and $Ref_2 = 0.1$. Figure A.1 shows the realized histogram of thresholds from 10,000 repetitions. We see the threshold values are clustered at the cognitive reference numbers of $Ref_1 = 0$ and $Ref_2 = 0.1$.

Finally, we close this section by establishing the following testable implications: (1) when investors are subject to evaluation costs, they will construct a consideration set to be evaluated further, (2) in constructing a consideration set, they will drop funds below a certain threshold (Proposition B.2) in a dimension informative of fund manager skill (Proposition B.1), and (3) if investors are subject to a cognitive reference number bias, then the observed thresholds will be clustered at the cognitive reference numbers (Figure A.1).

C Proofs

Proof of Lemma B.1 From Assumption 1, we have that

$$f\left(s_H, \overline{\alpha}\right) f\left(s_L, \alpha\right) > f\left(s_L, \overline{\alpha}\right) f\left(s_H, \alpha\right)$$

for $\alpha < \overline{\alpha}$, which implies that

$$\begin{aligned} \int_{-\infty}^{\overline{\alpha}} f\left(s_H, \overline{\alpha}\right) f\left(s_L, \alpha\right) d\alpha &> \int_{-\infty}^{\overline{\alpha}} f\left(s_L, \overline{\alpha}\right) f\left(s_H, \alpha\right) d\alpha \\ \frac{f\left(s_H, \overline{\alpha}\right)}{f\left(s_L, \overline{\alpha}\right)} &> \frac{\int_{-\infty}^{\overline{\alpha}} f\left(s_H, \alpha\right) d\alpha}{\int_{-\infty}^{\overline{\alpha}} f\left(s_L, \alpha\right) d\alpha} \end{aligned} \tag{C.1}$$

Also, Assumption 1 gives that

$$f(s_H,\overline{\alpha})\,f(s_L,\alpha) > f(s_L,\overline{\alpha})\,f(s_H,\alpha)$$

for $\overline{\alpha} < \alpha$, which implies that

$$\begin{aligned}
\int_{\overline{\alpha}}^{\infty} f(s_H,\alpha)\,f(s_L,\overline{\alpha})\,d\alpha &> \int_{\overline{\alpha}}^{\infty} f(s_L,\alpha)\,f(s_H,\overline{\alpha})\,d\alpha \\
f(s_L,\overline{\alpha})\int_{\overline{\alpha}}^{\infty} f(s_H,\alpha)\,d\alpha &> f(s_H,\overline{\alpha})\int_{\overline{\alpha}}^{\infty} f(s_L,\alpha)\,d\alpha \\
\frac{\int_{\overline{\alpha}}^{\infty} f(s_H,\alpha)\,d\alpha}{\int_{\overline{\alpha}}^{\infty} f(s_L,\alpha)\,d\alpha} &> \frac{f(s_H,\overline{\alpha})}{f(s_L,\overline{\alpha})}.
\end{aligned} \tag{C.2}$$

Hence, combining (C.1) and (C.2) yields that

$$\frac{\int_{\overline{\alpha}}^{\infty} f(s_H,\alpha)\,d\alpha}{\int_{\overline{\alpha}}^{\infty} f(s_L,\alpha)\,d\alpha} > \frac{\int_{-\infty}^{\overline{\alpha}} f(s_H,\alpha)\,d\alpha}{\int_{-\infty}^{\overline{\alpha}} f(s_L,\alpha)\,d\alpha},$$

which implies

$$\Pr(\alpha < \overline{\alpha}|s_L) \geq \Pr(\alpha < \overline{\alpha}|s_H).$$

This completes the proof of the lemma. $\square$

Lemma C.1. *Consider two random variables of X and Y. Let F_X and F_Y denote the cdf of X and Y, respectively. Then, it holds that*

$$\mathbb{E}[X] - \mathbb{E}[Y] = \int_{-\infty}^{\infty} (F_Y(v) - F_X(v))\,dv.$$

Proof From integration by parts, we have that

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} v dF_X(v) = [vF_X(v)]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} F_X(v)\,dv. \tag{C.3}$$

Similarly, it holds that

$$\mathbb{E}\left[Y\right] = \left[vF_Y\left(v\right)\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} F_Y\left(v\right) dv. \tag{C.4}$$

From (C.3), (C.4) and the properties of $F_X\left(\infty\right) = F_Y\left(\infty\right) = 1$ and $F_X\left(-\infty\right) = F_Y\left(-\infty\right) = 0$, it hold that

$$\mathbb{E}\left[X\right] - \mathbb{E}\left[Y\right] = \int_{-\infty}^{\infty} \left(F_Y\left(v\right) - F_X\left(v\right)\right) dv \leq 0,$$

where the last inequality is from the assumption of $\Pr\left(X \leq k\right) \geq \Pr\left(Y \leq k\right)$ for any k. This completes the proof of the lemma. □

Lemma C.2. *Consider two random variables of X and Y such that $\Pr\left(X \leq k\right) \geq \Pr\left(Y \leq k\right)$ for any k. Then, it holds that $\mathbb{E}\left[X\right] \leq \mathbb{E}\left[Y\right]$.*

Proof From Lemma and the assumption of $\Pr\left(X \leq k\right) \geq \Pr\left(Y \leq k\right)$ for any k, it follows that

$$\mathbb{E}\left[X\right] - \mathbb{E}\left[Y\right] = \int_{-\infty}^{\infty} \left(F_Y\left(v\right) - F_X\left(v\right)\right) dv \leq 0.$$

This completes the proof of the lemma. □

Proof of Proposition B.1 This is a corollary of Theorem 2 in Blackwell (1951). □

Proof of Proposition B.2 Let $\widetilde{C}$ denote an optimal consideration set. Construct another consideration set $C_{s_{n(\widetilde{C})}} = \left\{i | s_i < s_{n\left(\widetilde{C}\right)}\right\}$. Because $s_i > s_{i+1}$, Lemma B.1 implies that

$$\Pr\left(\max_{i \in \widetilde{C}} \left\{\alpha_i\right\} \leq \overline{\alpha}\right) = \Pi_{i \in \widetilde{C}} \Pr\left(\alpha_i \leq \overline{\alpha}\right) \geq \Pi_{i \in C_{s_{n(\widetilde{C})}}} \Pr\left(\alpha_i \leq \overline{\alpha}\right) = \Pr\left(\max_{i \in C_{s_{n(\widetilde{C})}}} \left\{\alpha_i\right\} \leq \overline{\alpha}\right),$$

which, in conjunction with Lemma C.2, yields that

$$\mathbb{E}\left[\max_{i \in \widetilde{C}} \left\{\alpha_i\right\}\right] \leq \mathbb{E}\left[\max_{i \in C_{s_{n(\widetilde{C})}}} \left\{\alpha_i\right\}\right].$$

From the above inequality, $n\left(\widetilde{C}\right) = n\left(C_{s_{n(\widetilde{C})}}\right)$ and the definition of $U(C)$ given by (B.1), we have that

$$U\left(\widetilde{C}\right) \leq U\left(C_{s_{n(\widetilde{C})}}\right)$$

Because $\widetilde{C}$ is optimal, $U\left(\widetilde{C}\right) \geq U\left(C_{s_{n(\widetilde{C})}}\right)$. Hence, $U\left(\widetilde{C}\right) = U\left(C_{s_{n(\widetilde{C})}}\right)$, showing that $C_{s_{n(\widetilde{C})}}$ is also optimal. This completes the proof of the proposition. □

Proof of Lemma B.2 Fix j. Let F_{j-1}, F_j, F_{j+1}, G_j and G_{j+1} denote the cdf of $\max_{i \in C_{s_{j-1}}} \{\alpha_i\}$, $\max_{i \in C_{s_j}} \{\alpha_i\}$, and $\max_{i \in C_{s_{j+1}}} \{\alpha_i\}$, α_j and α_{j+1}, respectively. From Lemma C.1, we have that

$$\mathbb{E}\left[\max_{i \in C_j} \{\alpha_i\}\right] - \mathbb{E}\left[\max_{i \in C_{j-1}} \{\alpha_i\}\right] = \int_{-\infty}^{\infty} (F_{j-1} - F_j)\, dv$$

and that

$$\mathbb{E}\left[\max_{i \in C_{j+1}} \{\alpha_i\}\right] - \mathbb{E}\left[\max_{i \in C_j} \{\alpha_i\}\right] = \int_{-\infty}^{\infty} (F_j - F_{j+1})\, dv.$$

Hence, to prove the lemma, it suffices to show

$$F_{j-1} - F_j \geq F_j - F_{j+1}.$$

The above inequality holds because

$$\begin{aligned} F_j - F_{j+1} &= G_j F_{j-1} - G_{j+1} F_j \\ &\leq G_j F_{j-1} - G_j F_j = G_j (F_{j-1} - F_j) \\ &\leq F_{j-1} - F_j, \end{aligned}$$

where the second inequality is from Lemma B.1 and the last inequality is from $G_j \leq 1$ and $F_{j-1} - F_j \geq 0$. This completes the proof of the lemma. □

Proof of Proposition B.3 From Lemma B.2, we know that $U(C_j) - U(C_{j-1})$ is decreasing in j. Hence, the optimal j is found when $U(C_{j+1}) - U(C_j)$ becomes negative at the first moment. This completes the proof. $\square$

Proof of Corollary B.1 (i) Set $K = 0$. Because $\mathbb{E}\left[\max_{i \in C_{s_{I-1}}} \{\alpha_i\}\right] - \mathbb{E}\left[\max_{i \in C_{s_I}} \{\alpha_i\}\right] < 0$, it holds that $j^* = I$ from Proposition B.3. Since the utility is continuous in K, it still holds that $j^* = I$ when K is sufficiently small. (ii) Fix K, j^* as the solution which satisfies the conditions of Proposition B.3. Assume that the new evaluation cost is $K + \varepsilon$ with $\varepsilon > 0$. Then, it is clear that the two conditions cannot be satisifed when j^* is replaced with any $j^{**} > j^*$ from Lemma B.2. $\square$

Proof of Proposition B.1 Let s_j be the solution of From Proposition B.3. Then, it holds that

$$U_{Ref}\left(C_{s_j}\right) \geq U_{Ref}\left(C_s\right) \text{ for any } s \text{ such that } \sum_{h=1}^{H} \mathbf{1}\left(Ref_h = s\right) = 0.$$

Next, fix references of Ref_h or Ref_{h+1} such that $Ref_h \leq s_j \leq Ref_{h+1}$. Since $A \cdot \mathbb{E}\left[\max_{i \in C_{s_j}} \{\alpha_i\}\right] - n\left(C_{s_j}\right) \cdot K$ is in j, it holds that

$$U_{Ref}\left(C_{Ref_h}\right) \geq U_{Ref}\left(C_{Ref_{h'}}\right) \text{ for } h' < h \tag{C.5}$$

and that

$$U_{Ref}\left(C_{Ref_{h+1}}\right) \geq U_{Ref}\left(C_{Ref_{h'+1}}\right) \text{ for } h' > h. \tag{C.6}$$

Combining (C), (C.5) and (C.6) yields that

$$\max\left\{U_{Ref}\left(C_{s_j}\right), U_{Ref}\left(C_{Ref_h}\right), U_{Ref}\left(C_{Ref_{h+1}}\right)\right\} \geq U_{Ref}\left(C_s\right) \text{ for any } s,$$

which completes the proof of the proposition. $\square$