Information Representation under Cognitive Constraints*

Wendy Ye Jin†, Yiqing Lü‡

August 28, 2026

Abstract

This paper studies how cognitive constraints shape the representation of information in economic decisions. Using retail mutual fund investment decisions, we show that investors do not always respond continuously to continuous price signals. In a Regression Discontinuity Design, a small price decline increases purchase probability by 13%. We propose that decision complexity induces investors to compress continuous price information into coarse, sign-based representations. Consistent with this mechanism, the discontinuity is larger when investors face larger consideration sets and less concentrated portfolios, and when products are more complex and investors have lower cognitive capacity. Importantly, browsing behavior shows that investors devote more effort to complex decisions, yet greater effort is associated with stronger categorical responses, suggesting that compression reflects binding cognitive constraints rather than simple inattention. On the sell side, we find a parallel discontinuity as cumulative gains cross zero. Controlled experiments show that increasing decision complexity causally shifts individuals toward direction-based representations, with estimated reliance on such cues ranging from 20–43% in the simpler condition and 37–64% in the more complex condition, depending on the payoff gap. Together, the evidence suggests that cognitive constraints affect not only how much information individuals process, but also how finely they represent it.

JEL code: D81, D83, G11, G41

Keywords: information representation, cognitive effort, decision complexity, cognitive constraints, investor behavior.

*We thank Yakov Amihud, Soo Hong Chew, Xavier Gabaix, Ryan Oprea, Jeffery Wurgler, and seminar participants at Experimental Finance Conference 2026, Science of Decision Making III conference, Capital University of Business and Economics Household Finance Workshop 2025, Shanghai Jiaotong University, Central University of Finance and Economics, Beijing Technology and Business University, and Wuhan University. Any errors are attributable solely to the authors. The experiment was approved by NYU Shanghai IRB (2025-088). The experimental design and main hypotheses were preregistered prior to data collection on AsPredicted.

†Institute of Social and Economic Research (ISER), University of Osaka, 6-6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan. Email: wendy.jin@iser.osaka-u.ac.jp.

‡New York University Shanghai. Email: yiqing.lu@nyu.edu.

1 Introduction

Economic models typically represent choices as functions of the information available to decision makers. When the underlying signal varies continuously, the resulting beliefs and choices are often modeled as varying continuously as well: small changes in a signal generate correspondingly small changes in behavior. This logic is pervasive in models of consumer demand, investment, and information aggregation. Yet real-world decisions are often made under cognitive constraints. When individuals must process many alternatives and multiple dimensions of information, preserving every incremental difference in the underlying signal may be costly or infeasible.

A growing literature on bounded rationality and costly cognition argues that individuals cope with such constraints by simplifying the way they represent information. Rather than preserving the full precision of continuous signals, they may compress information into a smaller number of decision-relevant states(Gabaix, 2014; Gabaix and Graeber, 2024; Enke and Graeber, 2023; Enke et al., 2024). More generally, when an optimal decision rule requires tracking too many states, individuals may adopt simpler representations or decision rules that sacrifice precision for tractability (Oprea, 2020). These theories suggest that cognitive constraints can affect not only the magnitude of responses to information, but also the functional form through which information translates into choice. Yet direct evidence on this representation stage remains scarce: while the information individuals observe is often measurable, how they encode that information before making a decision is generally not.

We study this question using granular data on the browsing and trading activity of more than 13,000 retail mutual fund investors on a Chinese trading platform. The setting offers two unusual advantages for studying information representation in a field environment. First, we observe investors' browsing activity before each trading decision, including the specific fund pages they visit, when they visit them, and how long they spend, providing a rare, albeit partial, measure of the cognitive effort devoted to a decision. Second, we observe the prices displayed when investors browse, allowing us to construct, for each investor $\times$ asset $\times$ day observation, the price most recently encountered by the investor. We can therefore identify the incremental price signal available to individual investors at the time of the decision rather than relying on market-wide price measures.

These features allow us to test whether investors preserve the precision of a continuous price signal or compress it into coarse categories. Consider a daily price change close to zero relative to the most recently observed price. An investor who represents the signal continuously should respond similarly to a slightly positive and a slightly negative change. By contrast, an investor who compresses the signal into coarse categories may treat these observations as belonging to different states—such as up and down—even when their magnitudes are arbitrarily close. Such categorical representation generates a distinctive empirical prediction: a continuous price signal should produce a discontinuous response at the boundary separating the categories.

Exploiting a Regression Discontinuity (RD) design with the change in the observed price as the running variable, we document a sharp discontinuity in purchase behavior at the zero crossing. When the current price falls just below the price most recently observed by the investor, the

probability of purchase increases by 13% relative to the sample mean. Although the price change is arbitrarily small at the threshold, crossing zero produces a discrete change in behavior—as if the continuous price signal has been transformed into a categorical distinction between "a price decline" and "a price increase." This pattern is difficult to reconcile with a smooth response to information under Bayesian updating. More broadly, the finding points to a departure from a continuous mapping between information and choice: rather than responding smoothly to the magnitude of the price signal, investors respond differently on the two sides of the zero threshold, consistent with a categorical rather than continuous representation of the signal.

We use the discontinuity at the zero threshold as an empirical anchor: if cognitive constraints induce investors to compress continuous price information into categorical representations, the discontinuity should become more pronounced when the demands on cognitive capacity are greater. This approach allows us to distinguish a cognitive representation mechanism from explanations based on beliefs or stable preferences (Oprea, 2024a).

A central question, then, is whether this discontinuous response reflects stable preferences or instead emerges when cognitive constraints make the underlying information difficult to represent. Following Gabaix and Graeber (2024), we examine three dimensions of cognitive difficulty: the complexity of the decision environment, the complexity of the product itself, and the cognitive capacity of the investor. A stable preference does not predict that the same preference-based response should systematically strengthen when the decision becomes more cognitively demanding. If cognitive constraints instead induce investors to compress continuous price information into categorical representations, the discontinuity should become more pronounced precisely when the demands on cognitive capacity are greater. We therefore use the zero-crossing discontinuity as an empirical anchor to distinguish cognitive representation from explanations based on beliefs or preferences (Oprea, 2024a).

We proxy the difficulty of the *decision environment* along two margins. The first is the size of the investor's consideration set, defined as the number of distinct assets browsed during the week before a trade, where a larger set increases the dimensionality of the choice problem. The second is portfolio concentration, measured by the Herfindahl index of the investor's holdings: less concentrated portfolios involve more interactions among positions and thus a more complex allocation problem. For *product complexity*, we construct two novel text-based indices from mutual fund prospectuses. The index captures product design and disclosure—fee structure and document complexity. Fee structure complexity reflects tiered fee schedules and the proliferation of share classes, while document complexity reflects prospectus length, structural density, and reliance on defined terms. Our text-based measures exhibit substantial within-category variation, allowing us to identify product-level heterogeneity that the standard equity/bond split would mask. We proxy investor *cognitive capacity* using educational attainment and occupation: higher education and employment as a finance professional are considered as being associated with a lower cost of processing complex financial information.

We find that the purchase discontinuity is significantly larger when investors face larger con-

sideration sets, hold less concentrated portfolios, and encounter products with more complex fee structures and documentation, as well as when investors have lower educational attainment. This systematic pattern across decision, product, and investor dimensions is difficult to reconcile with preference-based explanations and is consistent with the view that greater cognitive difficulty induces coarser representations of information (Enke et al., 2024).

The same categorical-response framework also applies to sell decisions. We document a parallel discontinuity: the propensity to sell jumps by 17% relative to baseline as cumulative gains on a held position cross from negative to positive territory, resembling the classic disposition effect (Odean, 1998). The discontinuity is stronger for investors with larger consideration sets and less concentrated portfolios, but shows little heterogeneity by product complexity or investor capacity. This asymmetry suggests that the sources of cognitive difficulty differ across buying and selling: while buying requires evaluating a new product against competing alternatives, selling an already-held asset involves less new product evaluation, shifting complexity toward the broader portfolio allocation problem. Thus, categorical representation appears to operate through different margins across the two decisions.

The mechanism is rooted in the interaction between decision complexity and attention effort under limited cognitive capacity. When a decision becomes more complex, investors have greater incentive to devote cognitive effort to processing the relevant information. Yet cognitive effort cannot increase without bound, so sufficiently complex decisions can exceed the capacity of the effort invested. In this region, investors may preserve tractability by compressing continuous information into coarse, decision-relevant categories.

Our browsing data provide direct evidence for this mechanism. We find that the discontinuity is more pronounced in decisions that attract greater browsing effort, indicating that categorical representation is not simply a consequence of failing to attend to small price movements. Rather, investors devote more resources to decisions that demand greater cognitive processing, yet still rely on categorical representations when those demands strain their cognitive capacity. More importantly, we directly link two components of the mechanism that have been difficult to measure in the field: attention effort and decision complexity. We find that product-level attention rises with the size of the investor's consideration set, but at a diminishing rate, producing a concave effort–complexity relationship. This pattern is consistent with a cognitive production function in which greater complexity elicits greater effort, while cognitive constraints increasingly limit the marginal amount of information that can be processed. Together, these results suggest that the discontinuity is not simply a failure to attend to small price movements, but an endogenous simplification of information when the demands of a complex decision approach the limits of cognitive capacity.

A limitation of the field evidence is that decision complexity is partially endogenous: investors choose what to browse, how many alternatives to evaluate, which products to hold, etc. To establish a causal link from complexity to categorical processing, we conduct a controlled experiment that exogenously varies decision complexity. Participants decide whether to invest in a focal asset or select an outside option, while one component of the focal asset's payoff is made visually salient but

normatively irrelevant for the optimal decision. We manipulate complexity by varying the number of outside options and the difficulty of comparing alternatives. Under full information processing, choices should be independent of the salient component's sign. Instead, investment decisions exhibit a pronounced discontinuity at the sign change, and this discontinuity grows systematically with experimentally assigned complexity. We discipline the interpretation through a structural model adapted from the discrete-choice framework of Gabaix and Graeber (2024, Section 5), which allows investors to place weight on both the magnitude and the sign of the salient signal. Estimated on the experimental data, the model implies that the share of decision weight investors place on the sign of the salient component rises from 20-43% in the simpler condition to 37-64% in the more complex one—a magnitude that mirrors the field heterogeneity.

We consider several alternative explanations, including reference-dependent preferences, salience-based mechanisms, interface-driven color cues, and preferences for simplicity. These mechanisms can potentially generate a response at the zero threshold, but they primarily operate through how investors value or attend to information. Our evidence instead points to a constraint on how information is represented. In particular, the discontinuity systematically increases with decision complexity and investor effort, while varying predictably with measures of cognitive capacity. These patterns are difficult to reconcile with mechanisms based solely on stable preferences or automatic salience, but are naturally predicted if greater cognitive demands induce investors to compress continuous signals into coarser representations. We discuss these alternative interpretations in detail and show that our mechanism provides the most unified account of the evidence.

Our field and experimental findings have broad implications for the structure of retail demand and investor welfare, while also offering a new perspective on a classic behavioral anomaly. First, our findings speak to the structure of retail demand for financial assets. A large literature shows that asset demand is not perfectly elastic and that demand elasticity has important implications for prices (Shleifer, 1986; Koijen and Yogo, 2019; Gabaix and Koijen, 2022). We identify a psychological microfoundation for a different departure from frictionless benchmarks: the mapping from information to demand need not be smooth. Instead, demand can change discretely at salient boundaries, with the magnitude of this non-smoothness increasing with the complexity of the decision environment. Cognitive constraints therefore affect not only the level of demand, but also its shape, generating localized and complexity-dependent departures from smooth demand.

Second, this non-smooth response has direct welfare implications. Trades induced by crossing the zero threshold do not improve investment performance: assets purchased just above and just below the threshold have statistically indistinguishable subsequent returns. These trades therefore generate transaction costs without improving portfolio allocation. We estimate the resulting platform-level cost at RMB 1–10 per investor per year in subscription fees alone, implying RMB 60–600 million annually across China's roughly 60 million retail mutual fund investors. Because the discontinuity is larger in more complex decision environments, complexity can generate not only cognitive costs but also economically meaningful trading costs.

Finally, the same mechanism offers a new perspective on the disposition effect. The standard

interpretation attributes gain–loss asymmetry to reference-dependent preferences (Kahneman and Tversky, 1979) or beliefs about mean reversion. Our evidence points to a complementary channel: investors may compress continuous gains and losses into categorical representations when the allocation problem is cognitively demanding. This interpretation generates testable predictions about when disposition-type behavior should be stronger or weaker—predictions that depend on the complexity of the decision rather than solely on stable preferences—and suggests that interventions targeting decision environments, such as interface design or product disclosure, may attenuate the behavior.

Our paper contributes to the emerging literature on cognitive constraints and decision-making under complexity. A central insight of this literature is that cognitive constraints attenuate the sensitivity of decisions to economically relevant information (Gabaix and Graeber, 2024; Enke and Graeber, 2023; Enke et al., 2024; Oprea, 2024b). We identify a distinct and important margin through which complexity affects economic decisions: the representation of information itself. Rather than simply responding less to a continuous signal, investors may compress it into a smaller set of decision-relevant states. This distinction generates a novel empirical prediction: compression can make the mapping from a continuous signal to choice discontinuous at the boundaries separating categories. We document this pattern in real-world investment decisions and show that the discontinuity becomes larger with decision complexity. Our evidence complements experimental studies of cognitive uncertainty, noisy information processing, and complexity-driven mistakes (Frydman and Jin, 2022; Enke and Shubatt, 2023; Charles et al., 2024; Oprea, 2024b; Banovetz and Oprea, 2023; Enke et al., 2024).

Our paper is also closely related to recent work showing that complexity increases reliance on coarse information or categories. Graeber et al. (2025), for example, show that investors respond disproportionately to category boundaries in earnings information when decisions are more complex. Our setting differs in an important respect: the underlying price signal is continuous and is not accompanied by an economically meaningful categorical label. The up/down distinction instead emerges in investors' responses at the zero threshold. We therefore provide evidence that cognitive constraints can endogenously transform a continuous economic signal into a categorical representation, rather than simply increasing reliance on coarse categories already supplied by the information environment.

Our granular data further contribute by allowing us to directly characterize the relationship between decision complexity, cognitive effort, and information representation. We observe investors' browsing activity before each trading decision, providing a rare, albeit partial, measure of cognitive effort. We document a concave relationship between decision complexity and effort: investors devote progressively more effort as decisions become more complex, but the marginal increase in effort declines. We further show that the categorical response is stronger among decisions receiving greater effort. Together, these findings provide field evidence not only that complexity consumes cognitive resources, but also that the way investors represent information changes as those resources become constrained.

Second, our paper contributes to the literature on reference-dependent trading behavior, including the disposition effect (Odean, 1998; Quispe-Torreblanca et al., 2025). Existing models explain gain–loss asymmetries primarily through reference-dependent preferences and beliefs. We identify a complementary channel: cognitive constraints can cause investors to compress continuous cumulative returns into categorical gain/loss states, generating a discrete response at the gain–loss boundary. The fact that this discontinuity varies systematically with decision complexity suggests that part of the disposition effect may reflect how investors represent information, rather than solely how they value gains and losses. We further show that complexity operates through different margins across buying and selling: product complexity and investor capacity are more relevant when evaluating new investment opportunities, whereas consideration-set and portfolio complexity remain important when deciding whether to sell an existing position.

Third, we contribute to the literature on financial product complexity by shifting attention from the supply of complexity to its cognitive consequences on the demand side. Célérier and Vallée (2017) show that financial institutions strategically design complex securities to cater to investor demand. We instead examine how complexity affects the investor's ability to represent and act on information. Our novel text-based measures of prospectus complexity allow us to distinguish product-level complexity from the complexity of the decision environment itself, showing that both can shape the way investors respond to information.

The paper proceeds as follows. Section 2 describes the data and sample construction. Section 3 presents the regression discontinuity design and validation. Section 4 establishes the buy-side discontinuity and its complexity heterogeneity. Section 5 extends the analysis to sell decisions and connects to the disposition effect. Section 6 examines the cognitive-effort mechanism and addresses alternative explanations. Section 7 reports the experimental and structural results. Section 8 discusses the implications and concludes.

2 Data and Sample Construction

Our proprietary data is from a professional third-party financial institution in China, spanning from December 1, 2021 to October 31, 2022 with over 13,000 investors. The institution holds an independent fund sales license, and its trading app facilitates transactions in both public and private funds while also offering retail financial services. Our data consists of only public not private fund investors. We have access to individual investors' browsing records on the institution's trading app, which capture the start and end time of each instance in which an investor views a fund's information page. Additionally, the dataset contains details on investors' fund holdings and trading activities on the platform, including transaction timestamps, trade types, and trade values. Lastly, we observe investors' demographic characteristics, such as gender, age, education level, and profession.

We study whether investors choose to trade a specific asset on a given day. This requires a decision-level sample containing both realized trades and counterfactual opportunities to trade.

We therefore first identify potential trading dates for each investor based on their activity within the app. We define a *potential trading date* as any day on which an investor interacts with the app, either by browsing assets or executing a trade. For each potential purchase date, we focus on assets that the investor browsed on the preceding day. This restriction captures assets that were recently under consideration, provides a more meaningful counterfactual for an actual trading decision, and allows us to identify the price most recently observed by the investor. It also improves comparability across assets by conditioning on recent investor engagement with the asset.

Our setting further facilitates the identification of the price signal available at the time of the decision. We focus on retail trading of over-the-counter (OTC) mutual funds in China, whose prices are typically updated only once per trading day, after the underlying securities markets close. Unlike exchange-traded securities, OTC mutual funds do not have an intraday market price. Consequently, on any potential trading day, the most recent price available to investors is the fund's price from the preceding trading day. We therefore match each investor–fund–day observation with the fund's price from the prior trading day and define the price signal as the change from the price most recently observed by the investor. This institutional feature allows us to precisely measure the continuous price information available at the time of each potential trading decision.

Our constructed sample consists of 11,713 investment accounts, each of which has at least one potential trading date within the sampling period. Summary statistics is provided in Table 1. The average age of account holders is 46.2 years (median: 40), and 44% of them are female. The average account balance is RMB 198,208 (median: RMB 51,769.9). On average, each account holds 8.4 assets (median: 5) and executes 7.3 trades per month (median: 4).

[Insert Table 1 here]

As mentioned earlier, each observation in this sample represents a combination of investor $\times$ potential trading date $\times$ potential tradable asset, totaling 605,727 observations. For these observations, the average price change between the trading day and the previous day is 0 (median: 0), with a standard deviation of 0.043. Among them, approximately 2.7% of cases result in an actual buy transaction. The average transaction amount for these purchases is RMB 4,818.1.

Additionally, we observe investors' browsing activity both on and before the potential trading day. Specifically, we define the set of assets viewed within the week leading up to and including the potential trading day as the investor's consideration set. We find that, on average, an investor's consideration set contains 48.9 assets (median: 22). The average number of views per asset is 7.3, with an average viewing duration of 91.5 seconds. At the set level, investors view all assets in the set 142.6 times, with a total browsing length of approximately 1,608 seconds.

In terms of complexity measures, the average portfolio HHI in this sample is 0.22 (median: 0.14); the average fee-structure complexity score is 4.69 (median: 5.38); and the average document-complexity score is 2.90 (median: 3.57).

Sample construction on the sell side parallels the buy-side approach with a key adaptation, motivated by the fact that the original purchase price serves as the natural reference point for

reviewing a held position. We use an investor's cumulative gains (measured in thousand RMB) since purchase as the forcing variable. Because the reference does not depend on attention devoted to the price browsed the day before the trade, we define potential sell trades if investors browsed the asset in the three weeks preceding a potential sell date. This window captures holdings that were recently under consideration while allowing for a less immediate relationship between browsing activity and the subsequent sell decision.

Each observation in the sell sample represents a combination of an investor, an asset, and a potential sell day, totaling 2,245,474 observations. For each potential sell date, we do not require assets that the investor browsed on the preceding day. The average cumulative gain since purchase is RMB 208 (median: −RMB 5), with a standard deviation of RMB 538. Approximately 0.56% of cases result in an actual sell transaction, with an average transaction amount of RMB 11,535.1 (median: RMB 1,202.1) for these sales. We define the sell-side consideration set over the week leading up to and including the potential trading day, reflecting that investors may revisit a held position multiple times before deciding to sell. On average, this set contains 30.7 assets (median: 15). The average number of views per asset is 3.5 (median: 1), with an average viewing duration of 40.0 seconds. At the set level, investors view all assets in the set 58.8 times on average, with a total browsing length of approximately 572 seconds.

In terms of complexity measures for sell trades, the average portfolio HHI is 0.15 (median: 0.11). the average fee-structure complexity score is 4.70 (median: 5.38); and the average document-complexity score is 2.90 (median: 3.53).

3 Empirical Methodology

Consider on date t, an investor i observes an asset j's price changes from the previous day's browsed price $p_{j,t-1}$ to $p_{j,t}$. For a rational investor who possesses unlimited cognitive capacity, she is able to update her belief ($\tilde{V}(\cdot)$) of asset value $V_{j,t}$ following Bayes' rule. Assuming that the asset's price changes by an arbitrarily small amount of $\Delta p_{j,t}$ (i.e., from $p_{j,t-1}$ to $p_{j,t}$) without other contemporaneous changes, the investor should change her belief about the asset value by an amount from $\tilde{V}(V_{j,t} \mid p_{j,t-1})$ to $\tilde{V}(V_{j,t} \mid p_{j,t})$, proportional to the change in likelihood also by a sufficiently small amount following a continuous likelihood function ($L(p_{j,t} \mid V_{j,t}) - L(p_{j,t-1} \mid V_{j,t})$). In other words, a rational response to observing a small price change should lead to a negligible adjustment in perceived asset value and, consequently, the probability of making a purchase.

3.1 A Regression Discontinuity Design

Cognitive constraints may prevent investors from preserving the full precision of continuous price information, particularly in complex decisions, leading them to compress signals into coarse, decision-relevant categories. Such compression implies that the mapping from a continuous price signal to investment decisions need not be continuous: crossing a categorical boundary can generate a discrete change in behavior even when the underlying signal changes arbitrarily little.

This motivates our Regression Discontinuity (RD) design, which uses the zero crossing of the investor's most recently observed price change as the boundary between two such categories. RD designs can be used to measure treatment effects when a treatment is based on whether an underlying continuous forcing variable crosses a threshold. Under the condition that there is no other source of discontinuity, the treatment effect induces a discontinuity in the outcome of interest at the threshold. Thus, the outcomes on the two sides of the threshold are unequal, and the difference between these two directional limits measures the treatment effect. A necessary condition for the validity of the RD design is that the forcing variable itself is continuous at the threshold (Hahn et al., 2001).

In our setting, we construct a forcing variable which is the deviation of the asset's price today from its previous-day price browsed by the investor. When such a deviation is small, we consider it as almost random, especially in light of a typical retail mutual fund investor who has no market impact. Consider a case in which we wish to test whether on date t, an investor i who observes a drop in an asset j's price $p_{j,t}$ from the previous day's browsed price $p_{j,t-1}$ has a significantly different probability of purchasing the asset from observing a price increase. Let $y_{i,j,t}$ refer to the purchase response (or not) for the investor i, and let $\Delta p_{j,t}$ be the one-day price change. We select bandwidth θ for the maximum absolute value of $\Delta p_{j,t}$. The bandwidth can be an arbitrarily small value, and different values can be used in the robustness checks. The RD design compares observations for which $0 < \Delta p_{j,t} < \theta$ with those for which $-\theta < \Delta p_{j,t} < 0$. Thus, we retain observations for which $|\Delta p_{j,t}| < \theta$. Finally, we estimate the effect of observing a price change on purchase behavior by utilizing a subsample in which the price changes are within a specified bandwidth. In robustness tests, we also use one-day return as the running variable.

We follow Lee and Lemieux (2010) to approximate the underlying relationship between investors' purchase response $y_{i,j,t}$ (the dependent variable) and one-day price change $\Delta p_{j,t}$ with the inclusion of four polynomials of the first and second order that are different for each side. The treatment effect is represented by coefficient β in Equation (1) below.

$$y_{i,j,t} = \alpha + \beta \mathbb{1}\{\Delta p_{j,t} \geq 0\} + P(\Delta p_{j,t}, r) + P(\Delta p_{j,t}, l) + \lambda_i + \gamma_j + \nu_m + \varepsilon_{i,j,t}. \tag{1}$$

In Equation (1), β is the coefficient of interest. The term $\mathbb{1}\{\Delta p_{j,t} \geq 0\}$ denotes an indicator function equal to one if the observed price change is nonnegative and zero otherwise. The functions $P(\Delta p_{j,t}, r)$ and $P(\Delta p_{j,t}, l)$ include first- and second-order polynomials of the price changes for the observations on the right-hand side of the price changes (i.e., the positive price changes), and $P(\Delta p_{j,t}, l)$ includes two polynomials of the first and second order of the price changes for observations on the left-hand side of the price changes (i.e., the negative price changes). The polynomials aim to capture continuous relationship between $\Delta p_{j,t}$ and $y_{i,j,t}$ and account for the effects of other determinants of purchase behavior that vary smoothly with the price change (Imbens and Lemieux, 2008; Lee and Lemieux, 2010). λ_i, γ_j, and ν_m refer to investor, product and yearmonth fixed effects, respectively. We cluster the standard errors at the investor level.

Because the RD design is identified only at the discontinuity, in the empirical estimation of equa-

tion (1), the bandwidth θ should be chosen to balance the goal of staying as local to the zero cutoff as possible with that of ensuring that there are enough data to yield informative estimates. We follow the approach of (Calonico et al., 2014) to calculate the optimal bandwidth that remarkably improves on alternative nonparametric estimators in consistency and robustness. For robustness, we also employ a practical implementation of RD within various pre-specified bandwidths. Our results are robust across these estimation approaches using different bandwidths.

3.2 Validation

There are two important conditions for the validity of the RD design. First, the sole source of discontinuity should be the treatment (in our case, price changes). In other words, other observable and unobservable characteristics of the investors and products the left and to the right of the cutoff should be similar. We compare a set of covariates between observations to the left and right of the cutoff, namely, the annual standard deviation of the asset price, the investor's browsing frequency and length of the asset one week prior to the potential trading date, as well as the investor's average holding value of the month. We present the results in Table A.1, which show no significant jump in these variables at the cutoff.

[Insert Table A.1 here]

The second condition for the RD design's validity is a lack of manipulation or self-selection near the cutoff. In other words, potential outcomes by treatment status should be continuous functions of the running variable as that running variable crosses the cutoff. Otherwise, it would raise the concern that agents can endogenously manipulate their observed running variable values to opt themselves into or out of treatment. In our analysis, we focus on retail mutual fund investors who have relatively limited capital for investment. As a result, they should be considered as having no price impact and acting solely as price takers in the mutual fund market. In addition, we provide in Figure 1, a visual representation of the continuity of the density of the running variable around the cutoff of zero. Specifically, we implement manipulation testing procedures of one-day price changes of traded assets using the second-order local polynomial density estimators with a triangular kernel function proposed in Cattaneo et al. (2019). Black curve represents the point estimates and grey area indicates 95% confidence intervals. Our analysis shows a P-value of 0.86 around the cutoff, implying smoothness in the density function. The result is robust to using one-day return as the running variable.

[Insert Figure 1 here]

4 Results: Buy Decisions

We begin our empirical analysis with buy decisions based on the RD design. As noted in the previous section, under Bayesian updating, a small variation in asset price implies a negligible change in

asset fundamental value and, consequently, purchase probability, so any observed discontinuity constitutes a clean deviation from rationality. We then show how the discontinuity varies with the complexity of financial decisions.

4.1 Main Results Based on the RD Design

Table 2 presents the relationship between the likelihood of an investor purchasing an asset and the price change of that asset within a single day. *Buy* is a dummy variable that equals one if the asset is bought and zero if it is not. *Pos Price Chg* is a dummy variable that equals one if the change in asset price since the previous day is greater than or equal to zero. Investor, product, and year-month fixed effects are included in all specifications. We control polynomials, different for each side, of order two in columns (1)–(3), order one in columns (4)–(5), and no polynomials in columns (6)–(7). We show the RD regressions at the optimal and other bandwidths: column (1) uses the optimal bandwidth of 0.011 selected under a uniform kernel, while column (2) uses a wider bandwidth of 0.015 selected under a triangular kernel for robustness. Columns (4) and (6) adopt bandwidths of 0.008 and 0.002, optimal under the respective polynomial functions. Columns (3), (5), and (7) use a global support of 0.5 for an examination of external validity.

[Insert Table 2 here]

In Table 2, we find that investors' purchase decisions exhibit a sharp discontinuity at the zero cutoff. When an asset's price at the trading date has just dropped below the price the investor browsed one day earlier, the probability of purchasing that asset jumps discretely upward. Specifically, When the optimal bandwidth is used in our preferred specification (column (1)), we find that the probability of purchasing the asset increases by 13% (0.3%/2.3%) relative to the mean of 2.3% when the asset price drops compared to the previous day's browsed price. This baseline finding is robust to both local and global bandwidths as shown in columns (2) and (3), and the magnitude is comparable across different specifications employing different polynomial structures. We interpret the 13% jump as the deviation from the rational level of purchase probability.

Figure 2 provides a visual representation of the discontinuity. Observations are sorted, based on the price change, into bins of size 0.02, with five bins on each side of zero shown on the x-axis. Points on the left (right) side of zero represent observations from negative (positive) price changes. The graph plots the average purchase likelihood on the y-axis, where each point represents observations within their respective bin (with 95% confidence intervals). We use polynomials of order two, different for each side. Consistent with the findings in Table 2, we observe a notable visual discontinuity in the probability of purchasing a mutual fund around the zero cutoff.

[Insert Figure 2 here]

The baseline pattern is robust to alternative sample and running variable construction. As shown in Panel A of Appendix Table A.2, restricting counterfactuals to dates on which the investor

traded at least one other asset—thereby excluding instances in which the investor was entirely inactive on the potential trading day—yields a similar discontinuity. In addition, using the return rate rather than the level of price changes as the running variable produces comparable results (Panel B).

In sum, we find a robust pattern that investors' response to a tiny price change deviates from the rational level. The observed jump can be interpreted as a form of behavioral attenuation toward a sign-based default (Enke and Graeber, 2023; Gabaix and Graeber, 2024): because processing fine-grained continuous price signals is cognitively demanding, investors substitute the sign of the price change for its magnitude, producing a discrete jump at zero even when the price movement is infinitesimally small. If complexity indeed affects how investors perceive and respond to price signals, the discontinuity should be stronger in decision environments where processing continuous price signals is more cognitively demanding. We turn to testing this prediction next.

4.2 Complexity Heterogeneity

If the zero-crossing jump reflects investors' cognitive simplification induced by decision complexity, it should be more pronounced in environments where processing the continuous price signal is more cognitively demanding. Following Gabaix and Graeber (2024), we measure the complexity of a trade decision in three aspects: the dimension of the decision (the asset choice problem), the complexity of analyzing each dimension, and the cognitive capacity of the decision maker. We divide the sample based on whether the complexity proxy is above or below the sample median. We then compare the local jumps between the two subsamples.Table 3 reports the heterogeneity results.

Dimensions of Asset Choice Problems

The higher the dimension of an asset choice problem, the more complex the decision is. We proxy this dimension in three ways. First, we conjecture that the more assets an investor considers, the more complex the asset choice problem. We measure the size of investors' consideration set by the number of assets the investor browsed during the week prior to the trade.

Second, when making an investment decision, an investor must consider how the asset relates to the existing assets in her portfolio (e.g., covariances). The more complex the composition of current holdings, the more complex the decision to purchase an asset becomes. We use the Herfindahl-Hirschman Index (HHI) of an investor's portfolio to measure this complexity, where a lower HHI indicates that the investor must take into account a greater number of existing assets.

[Insert Table 3 here]

We construct a dummy variable, *Complex*, which equals one if the consideration set size is above the sample median (a large set), or if the HHI of the investor's portfolio is below the median (low portfolio HHI). We then interact the dummy variable of complexity with the price change dummy,

and report the result in Columns (1)–(2) of Table 3. Consistent with the complexity hypothesis, we find that the coefficient on the interaction term is significantly negative across the two columns. In other words, the discontinuos response is more pronounced for asset choice problems that are more cognitively demanding, i.e., featuring larger consideration sets and lower portfolio HHI. Note that the largely insignificant coefficients on *Pos. Price Chg* suggest that zero-crossing jump is less noticeable in the subsample of less complex decisions.

Product Complexity

We next turn to product-level complexity. Rather than relying on coarse fund-category proxies that mask substantial within-category variation, we construct data-driven measures of product-level complexity from mutual fund prospectuses along two dimensions: product design, captured by *Fee Structure*, and information disclosure, captured by *Document Complexity*. *Fee Structure* complexity reflects tiered fee schedules and the proliferation of share classes; *Document Complexity* reflects prospectus length, structural density, and reliance on defined terms.

Fee Structure complexity combines three sub-components into a single index rescaled to a 0–10 scale. Fee type count awards one point for each of seven distinct fee types disclosed in the prospectus—management, custodian, sales service, performance, index licence, subscription, and redemption fees—for a maximum of 7 points. Share class count awards up to 3 points based on the number of share classes explicitly paired with a fee rate within the fee section, which avoids inflating the count from cross-references to other funds. Tiered fee conditions award up to 3 points for structural features indicating conditional or graduated pricing, including holding-period tiers in the redemption schedule, an amount-based auto-upgrade mechanism, a performance fee, and automatic share-class conversion. The three sub-scores are summed and rescaled by 13 to a common 0–10 index, capped at 10.

Document Complexity is constructed from four dimensions of a fund's prospectus: document length, defined terms, section count, and table density. Document length is measured by the number of characters (excluding whitespace), defined terms are identified from quoted terms in the definitions or glossary section, section count is the number of top-level numbered sections, and table density is measured by the number of lines containing percentages, currency amounts, or four-digit numerical values. Each component is normalized by dividing by its sample 90th percentile, scaling the result to a 0–10 range, and truncating values above 10. The overall document complexity index is then constructed by taking the simple average of these normalized components.

Specifically, we parse the full prospectus documents and map unstructured text into structured variables using automated text analysis. We parse the full prospectus documents and map unstructured text into structured variables using three approaches[1]. First, we use keyword dictionaries to identify fee types. Second, we apply pattern-based rules to detect structured elements such as

[1] After classifying each document, 269,171 are identified as full prospectuses and scored; 54,201 are prospectus summaries (with the word "summary" appearing within the first 2,000 characters), 52,966 are too short to score (typically redirect notices or with under 500 characters), and 1,436 are amendment announcements. Only full prospectuses receive a CI score.

fee schedules and section patterns. Third, we compute document-level statistics—such as holding-period tiers, length, number of sections, and density of numerical content—to capture the overall complexity of information disclosure. The complexity index exhibits substantial cross-sectional variation: *Fee structure* has a standard deviation of 1.58, with the 5th and 95th percentiles at 3.08 and 6.92, respectively. *Document Complexity* has a standard deviation of 2.62, with the 5th and 95th percentiles at 0.23 and 6.45, respectively.

Columns (3)—(4) Table 3 report the results for the comparison between the high and low complexity subsamples. The discontinuous response jump is significantly more pronounced for more complex funds. This result suggests that the deviation from continuous responses is stronger when the traded product itself is more cognitively demanding to evaluate.

Cognitive Capacity

Finally, we turn to investor-level cognitive capacity. We conjecture that investors with greater cognitive capacity or a lower cost of effort—whether through higher educational attainment or through financial expertise—are less prone to resort to simplified, categorical decision rules based on the sign of price changes when analyzing asset choices.

We proxy cognitive capacity in two ways: by education (college degree or above versus below) and by profession (finance industry versus others). Because the profession data contain substantial missing values, we focus on the magnitude rather than statistical significance when interpreting the profession-based split. Columns (5)–(6) of Table 3 report the results. The education split delivers a clean contrast: the zero-crossing jump is more evident in the low-education subsample or in professions outside the finance industry. These patterns indicate that the discontinuity in purchase probability is stronger when investors are less equipped to process complex price information.

Taken together with the decision-environment and product-complexity findings above, all three heterogeneity channels point to the same conclusion: the discontinuous response is more pronounced when the cognitive demands of the decision are greater.

5 The Disposition Effect

The disposition effect—the tendency of investors to sell winning positions and hold losing ones—is one of the most robust patterns documented in behavioral finance (Shefrin and Statman, 1985; Odean, 1998). Traditional explanations attribute this asymmetry to reference-dependent preferences or to mean-reverting beliefs about asset returns. We revisit the disposition effect through the lens of decision complexity: if investors reduce continuous cumulative gains to a binary signal—positive versus negative—when deciding whether to sell, the zero-crossing discontinuity documented for buy decisions should reappear on the sell side. This section offers a novel lens on the disposition effect: a zero-crossing discontinuity that mirrors the buy-side pattern and, as we show in Section 5.2, scales with decision complexity.

As noted in Section 2, we use an investor's cumulative gain (measured in thousand RMB) since

purchase as the forcing variable. Because the reference does not depend on attention devoted to the price browsed the day before the trade, we define potential sell trades if investors browsed the asset in the three weeks preceding a potential trading date.

5.1 Main Results

Table 4 reports the sell-side RD estimates. *Sell* is a dummy variable that equals one if the asset is sold on a given trading date and zero otherwise. *Pos. Cum. Perf.* is a dummy variable that equals one if the cumulative gain since purchase is greater than or equal to zero. Column (1) uses the optimal bandwidth of 0.813 selected under a uniform kernel, while columns (2) and (3) report estimates at a wider local bandwidth of o.954 selected under a triangular kernel and a global bandwidth of 10 for robustness. Investor, product, and year-month fixed effects are included in all specifications, and standard errors are clustered at the investor level.

[Insert Table 4 here]

We show that sell decisions also exhibit a sharp discontinuity at the zero cumulative-gain cutoff. Table 4 and Figure 3 establish this result both statistically and visually. When an investor's cumulative gain on a held position has just crossed zero from a small loss to a small gain, the probability of selling that position jumps discretely upward. Specifically, at the optimal bandwidth (column (1)), we find that the probability of selling the asset increases by approximately 15% (0.0008/0.0052) relative to the mean sell rate of 0.52% when the cumulative gain on the position turns just positive. This relative magnitude is comparable to 13% documented for buy decisions in Section 4.1. The pattern is robust across different specifications.

Figure 3 provides a visual representation of the discontinuity. Observations are sorted by cumulative gain into bins of size 1, with five bins on each side of zero shown on the x-axis. The graph plots the average sell likelihood on the y-axis, with second-order polynomials fitted separately on each side of the cutoff. Consistent with Table 4, we observe a clear visual discontinuity in selling probability at the zero threshold.

[Insert Figure 3 here]

The pattern is robust to alternative sample constructions. Restricting the counterfactual to dates on which the investor sold at least one other position in her portfolio yields an even larger discontinuity (Appendix Table A.4), lending additional support to the core finding.

Taken together, these results establish that the discontinuous response to price signals is not confined to purchase decisions: investors' willingness to sell a held position also jumps discretely as the cumulative gain crosses zero, mirroring the categorical (sign-based) processing of price signals on the buy side. While this jump may reflect reference-dependent preferences, we are primarily interested in whether complexity-induced simplification of price signals plays a part. We test next whether the discontinuity in sell decisions also exhibits the same complexity heterogeneity as its buy-side counterpart.

5.2 Complexity Heterogeneity

In this subsection, we test whether the complexity dimensions under which the buy-side discontinuity was more pronounced also shape its sell-side counterpart. We apply the three proxies developed in Section 4.2—the decision environment, product-level complexity, and investor cognitive capacity—to the sell-side RD specification. Table 5 reports the results.

[Insert Table 5 here]

Columns (1)–(3) report results for the three decision-environment proxies, Column (4) the prospectus-based Complexity Index, and (5)–(6) investor cognitive capacity. The consideration-set and portfolio-concentration splits mirror the buy-side patterns: the zero-crossing jump is larger for investors with larger consideration sets and for those with less concentrated portfolios. By contrast, heterogeneity along the product-complexity and investor-capacity dimensions is much more muted. This pattern suggests that sell-side complexity operates through different margins than buy-side complexity. Because investors already hold these products at the time of the sell decision, product complexity and investors' cognitive capacity are likely to be less binding constraints than when evaluating new investment opportunities.

6 Cognitive Constraints and Information Compression

In this section, we examine more directly the proposed mechanism of compressed mental representation. Our central hypothesis is that decision complexity strains investors' limited cognitive resources, leading them to rely more heavily on coarse, direction-based representations of continuous price signals. Section 6.1 tests this mechanism by examining how cognitive effort varies with decision complexity and whether greater effort is associated with stronger categorical responses, as predicted when additional effort is insufficient to fully overcome cognitive constraints. Section 6.2 then considers two alternative explanations. We first assess whether the discontinuity reflects the reference-dependent preferences emphasized by prospect theory. We next examine whether complexity heterogeneity instead reflects persistent differences in investors' cognitive ability or sophistication. Together, these analyses distinguish cognitive compression from preference-based and investor-level explanations.

6.1 Greater Effort on More Complex Decisions

The cognitive production function of Gabaix (2014) makes a specific prediction about the effort–discontinuity relationship: when a decision is complex, investors exert more effort engaging with it, but complexity can place demands that exceed the limits of effort given capacity constraint, leading investors to rely on simplified, direction-based signals nonetheless. Under this mechanism, the discontinuous response to price signals could be *larger* among investors who are most actively engaged with the decision. A competing interpretation—under which the discontinuity simply

reflects low-effort inattention to small price movements—predicts the opposite sign. We test both predictions by splitting the sample along three effort proxies.

We use browsing behavior to construct two proxies for the effort an investor devotes to a given decision. They are total frequency and total duration of an investor's browsing activity on a given asset during the week before a potential trading day. For each investor, we define a dummy *Effort* indicating whether the frequency (or duration) falls above the investor's own median. Columns (1)–(2) of Table 6 report the buy-side results. For both columns, we find that significant coefficients on the interaction term between *Pos. Price. Chg* and *Effort*. Our results suggest that investors recognize that a decision is important and complex and devote more cognitive resources to it, but the additional effort is insufficient to fully process the complexity, so they still rely on a coarse representation.

The same pattern holds for sell decisions. Columns (3) and (4) of Table 6 report the differences in sell-side RD estimates between decisions involving high and low attention. We find that the effort–discontinuity gap is also statistically significant under the two effort measures.

[Insert Table 6 here]

The positive effort–discontinuity relationship on both the buy and sell sides is consistent with the categorical representation mechanism: effort is deployed where complexity is greater, and sign-based processing is more pronounced among investors who are actively engaged with the decision. This places the source of the discontinuity in the cognitive demand of complex decisions on engaged investors, rather than in a lack of effort. As a result, the competing low-effort inattention account predicts the opposite sign and is not supported by the data.

A key implication of the proposed mechanism is the relationship between two latent variables that have previously been difficult to measure in the empirical literature: attention effort and decision complexity. If the mechanism is correct, attention effort should initially increase with decision complexity, but the increase should become less pronounced as complexity rises because of cognitive constraints.

We test this prediction in Figure 4. We proxy attention effort by product-level view frequency (Panel A) and view duration (Panel B) as used in Table 6, while decision complexity is measured by the size of the consideration set. We measure attention effort at the product level rather than the account level to avoid a potential mechanical relationship, as total account-level attention necessarily increases with the size of the consideration set. As the consideration set expands, evaluating a given product becomes more difficult because investors must compare it with an increasingly large set of alternatives. If our mechanism holds, investors should allocate more attention to each asset they consider as decision complexity increases.

We first residualize each attention measure by regressing it on investor, product, year, and month fixed effects, consistent with the main analysis. We then partition observations into 100 equally populated bins based on consideration set size. Figure 4 plots the mean residual attention measure within each bin, together with a fitted quadratic regression.

Across both measures of attention effort, we observe a pronounced concave relationship between product-level attention and decision complexity. Attention effort rises as the consideration set expands, but the increase becomes progressively smaller at higher levels of complexity. Although attention eventually declines for consideration sets exceeding roughly 100 funds, such observations are rare, and most of the sample lies in the increasing but flattening portion of the curve. Our finding implies that categorical processing is potentially an endogenous response to cognitive scarcity, not simply a preference or an anomaly.

[Insert Table 7 here]

Table 7 formally quantifies this relationship. We regress the two attention measures on the consideration set size (measured in hundreds of funds) and its quadratic term. Columns (1), (2), (5), and (6) use the optimal bandwidth of 0.011 around the cutoff, whereas columns (3), (4), (7), and (8) use a broader bandwidth of 0.5. Across all specifications, the coefficient on View set is positive and statistically significant. Once the quadratic term is included, the linear coefficient remains positive, while the quadratic coefficient is negative and statistically significant in all specifications. Overall, the estimates provide strong evidence of a concave relationship between product-level attention effort and decision complexity, consistent with the prediction that cognitive constraints reduce the marginal increase in attention as decisions become more complex.

Our finding that attention rises with decision complexity but at a diminishing rate connects to two recent literatures formalizing how complexity shapes cognition. Gabaix and Graeber (2024) model the mind as a cognitive production function in which performance is increasing and concave in cognitive labor and saturates once effort reaches the task's complexity; the concave, saturating relationship we document between view frequency and duration and consideration set size is a field analogue of this functional form, obtained from incentivized trading behavior rather than an elicited laboratory task. Oprea (2020) shows, in a setting stripped of risk, ambiguity, and framing, that both errors and implementation time—an incentivized proxy for effort closely related to our attention measures—rise with an automata-theoretic measure of rule complexity, suggesting the complexity-effort relationship we document is unlikely to be an artifact of financial framing or interface design.

6.2 Alternative Explanations

The alternative accounts considered above—canonical prospect theory, salience-based reweighting of payoffs, color-coded salience, and preference for simplicity—share a common structure: each locates the source of the zero-crossing discontinuity in how investors *value* or *attend to* given information, rather than in a constraint on their capacity to *process* it. Because they share this structure, we test them jointly against an expanding set of facts rather than in isolation, in the spirit of a specification race. We organize the comparison around four regularities established above: (i) a genuine level discontinuity in purchase and sell probability at the zero-crossing threshold, rather than a smooth kink (Table 2); (ii) heterogeneity in the size of that discontinuity across proxies

for decision complexity—consideration set size, portfolio concentration, product complexity, and investors= cognitive ability (Sections 4.2 and 5.2); (iii) a positive relationship between cognitive effort and the size of the discontinuity, documented using browsing behavior (Section 6.1). Each alternative survives longer than the last, but none survives all three.

Canonical prospect theory (Kahneman and Tversky, 1979) fails at the first hurdle. A kink in the value function implies continuity of choice probability at the reference point: as the price change approaches zero from either side, purchase probability should converge to the same limit. Our estimated level jump, stable across bandwidths (Table 2, columns (1)–(3)), rejects this prediction qualitatively, before heterogeneity is even considered.

Salience theory applied to payoffs (Bordalo et al., 2012) also struggles to clear the first hurdle, though for a different reason. The salience of an outcome is ordered by its contrast with a comparison context—typically the average payoff across a choice set—and this contrast function is continuous and *minimized* exactly where the outcome equals the context, i.e., at the investor's reference price. As the price change shrinks toward zero from either side, the contrast, and hence the salience-driven distortion of decision weights, shrinks toward its minimum on both sides. The natural prediction is therefore continuity in a neighborhood of the cutoff and *growing* divergence for large gains and losses, not a level jump concentrated at a vanishing bandwidth—the opposite of what an RD design with a shrinking neighborhood is built to detect. Generating a genuine jump at zero from this mechanism requires an auxiliary assumption that the *sign* of the deviation, not merely its magnitude, governs the salience ranking, which reintroduces a categorical, reference-based asymmetry rather than deriving one from contrast alone.

Color-coded categorization requires separating two readings that are easy to conflate, only one of which is a genuine rival to our account. Trading interfaces typically encode gains and losses in discrete colors, so the platform itself assigns a binary category label—positive-colored above the reference price, negative-colored below it—defined exactly at zero regardless of magnitude. This categorical assignment clears the first hurdle by construction: it generates a genuine discontinuity at zero irrespective of the size of the price change.

Under the first reading, *color-coded salience*, the color captures the investor's attention automatically, independent of whether processing the underlying number would have been costly. Nothing in this reading connects the strength of the response to the complexity of the decision or the investor's ability to resolve it; an automatic, resource-independent trigger has no reason to be more or less potent in a larger consideration set or for a less numerate investor. This version therefore already fails to reach result (ii): unlike the other accounts considered here, it does not need the effort result to be ruled out, because it has no channel to generate complexity heterogeneity in the first place.

Under the second reading, *color-based compression*, the color is not capturing attention against the investor's will; it is the channel through which an already capacity-constrained investor's own coarsening becomes observable. On this reading the investor cannot resolve the continuous magnitude of a price change regardless of how it is displayed, and the interface's color convention hap-

pens to supply, at zero cost, exactly the coarse, sign-based representation that a limited-capacity investor would arrive at anyway. Read this way, color-based compression can accommodate the discontinuity, its complexity heterogeneity, its ability heterogeneity, and—because the constraint binds regardless of effort—its positive relationship with effort as well. This is not a coincidence: color-based compression is not a rival explanation for the zero-crossing discontinuity, it is a specific institutional account of *why* the coarsening we document takes a directional, gain/loss form rather than some other partition of the outcome space, in a setting where trading interfaces universally render price changes as a binary directional cue. We do not attempt to distinguish it from our own account here; we return to it in Section 6 as a plausible microfoundation for the specific form the capacity constraint takes in a digital trading environment.

Another related theory is the preference for simplicity of Puri (2025), which models aversion to outcome-rich alternatives as lower valuation—or, equivalently, a higher required risk premium—for lotteries with more outcomes. The theory concerns the valuation of simple versus complex objects, whereas our mechanism concerns the representation of information used to evaluate those objects. One could potentially extend the preference-for-simplicity framework to allow individuals to value simpler representations of signals. A preference for simplicity is better understood as a stable taste over alternatives than as a constraint on the resources available to process them, and therefore does not directly predict stronger simplicity-based responses with greater attention effort. By contrast, with cognitive constraints, coarse representation may instead arise endogenously from the cognitive costs of encoding complex signals. Distinguishing these interpretations is beyond our empirical setting, but both point to a complementary margin of simplicity—not the simplicity of the choice object, but the simplicity of the representation through which it is evaluated.

7 Experiment

Our field analysis in previous sections establishes a sharp zero-crossing discontinuity in trading decisions and shows that it scales systematically with several proxies for decision complexity. Three features of the field setting, however, limit the inferences one can draw from observational data alone. First, investors endogenously choose which assets to browse and when to trade; our attention and complexity proxies could therefore be correlated with unobserved beliefs and preferences that no set of fixed effects can fully absorb. Second, the information set on which each trade is based is only partially observed: we see what investors browsed but not how they integrated that information into a decision. Third, because the value of a mutual fund is uncertain and investor-specific, the field setting does not provide a well-defined decision problem or a transparent normative benchmark against which to assess the magnitude of deviations from rational behavior.

To overcome these limitations, we design a controlled experiment that addresses all three concerns simultaneously. We exogenously assign subjects to decision environments that differ only in their complexity; we fully control the payoff-relevant information presented on each trial; and the task is constructed so that the rational benchmark is a flat 50% investment rate, independent

of the salient payoff component. This design allows us to test whether complexity itself—holding information, beliefs, and preferences fixed—causes individuals to rely more heavily on simplified, categorical decision rules.

Section 7.1 describes the design and implementation, Section 7.2 presents reduced-form evidence that mirrors the discontinuous behavior documented in the field, and Section 7.3 reports structural estimates that quantify how reliance on the payoff's sign grows with complexity.

7.1 Design and Implementation

Experimental Design

The Decision Problem In each trial, a participant decides whether to invest in a focal asset or to take the best outside option. The return from investing in the focal asset is determined by two deterministic numbers, A and B, with total value $A + B$. Introducing stochastic payoffs would confound sign-based complexity effects with risk preferences. Oprea (2024b) shows that these two channels are difficult to separate in risky settings—behavior traditionally attributed to risk preferences largely reflects complexity-driven cognitive errors. Our deterministic design eliminates this confound at the source. Component A is the salient component of the payoff—analogous to a daily price change in the field—and takes one of eight values, $A \in \{\pm 15, \pm 35, \pm 65, \pm 90\}$. Positive values of A are displayed in green and negative values in red. The second component B, analogous to asset value, is a two-digit positive integer drawn randomly across trials and provides a baseline level of value common to all assets on the screen.

Each focal asset is paired with a set of outside options that follow the same payoff rule, and the value of the best outside option equals the focal asset's value minus a deviation term $\Delta \in \{\pm 10, \pm 20, \pm 30\}$. A positive Δ means that the focal asset dominates the best outside option, while a negative Δ means that the outside option dominates. A subject's decision in a given round is objectively correct if she invests in the focal asset when $\Delta > 0$ and selects the outside option when $\Delta < 0$, and incentives are tied to this correctness criterion.

Complexity Conditions We vary decision complexity along two dimensions. The first is the size of the outside-option set. In the *Small Set Condition*, the focal asset is compared against a single outside option. In the *Large Set Condition*, the focal asset is compared against five outside options, of which the best is identical to the single alternative in the Small Set Condition; the other four are strictly dominated distractors constructed to be inferior to both the focal asset and the best outside option. These distractors increase cognitive load without altering the rational decision rule: under full information processing, the optimal choice is identical across the two conditions, and the Small and Large Set differ only in the complexity of the decision environment.

The second dimension of complexity is the payoff gap Δ between the focal asset and the best outside option. Smaller values of $|\Delta|$ make the comparison more cognitively demanding, as subjects must process payoff information more precisely to identify the dominant alternative.

Rational Benchmark The task is designed so that the rational decision rule is straightforward. A participant with unlimited cognitive capacity should invest in the focal asset if and only if Δ is positive, regardless of A. Because Δ is evenly distributed around zero, a fully rational decision-maker would invest in exactly half of the cases, yielding an average investment rate of 50% that is independent of A. Conditional on $\Delta > 0$, the rational investment rate is 100%; conditional on $\Delta < 0$, it is 0%. Figure A.1 plots these rational benchmarks.

[Insert Figure A.1 here]

Hypotheses The rational benchmark above relies on a participant's ability to fully integrate all payoff-relevant information. We conjecture that this ability deteriorates as the complexity of the decision environment increases, leading individuals to place disproportionate weight on information that is easier to process—in particular, on the salient component A and on its sign—even when doing so is normatively irrelevant. These considerations motivate three pre-registered hypotheses.

H1. Under the Large Set Condition, investment decisions are more sensitive to the level of the salient payoff component A.

H2. Under the Large Set Condition, investment decisions rely more heavily on the sign of A, indicating increased categorical processing.

H3. These effects are stronger when the payoff gap $|\Delta|$ between the focal asset and the best outside option is small.

Implementation

Each participant completes 48 independent decision problems corresponding to all combinations of the eight focal-asset values of A and the six outside-option values of Δ. Each decision screen presents six assets—one focal asset, the best outside option, and four distractors in the Large Set Condition, or a single outside option in the Small Set Condition. The order of problems and the on-screen position of the alternatives are randomized across participants to prevent learning or positional biases, and no feedback is provided after each decision to avoid reinforcement effects or dynamic updating. Before the main task, subjects complete a small number of practice trials to ensure they understand the payoff rule and the decision interface. The time limit is 20 seconds per decision, with a short break after 24 trials.[2]

We recruit 100 subjects through Prolific, with 50 subjects randomly assigned to each condition in a between-subject design. The experiment takes 12–18 minutes to complete. Participants receive a base payment of \$3.00 plus a performance-based bonus of up to \$2.00, determined by the outcome of a randomly selected decision problem from those they completed. The experimental design and main hypotheses were preregistered on AsPredicted prior to data collection, and the study was approved by the NYU Shanghai IRB (2025-088).

[2]On average, participants spend about 5–10 seconds per decision, and the entire experiment lasts approximately 15 minutes, including instructions and practice rounds.

7.2 Reduced-Form Results

We find that subjects' behavior deviates sharply from the rational benchmark, and the deviation is more pronounced under higher complexity along exactly the margins predicted by **H1**–**H3**. Figure 5 presents the average investment rate in the two conditions, first as a function of the salient component A (Panel A) and then broken down by the payoff gap $|\Delta|$ (Panel B). Table 8 then formalizes these patterns in a set of linear probability regressions.

[Insert Figure 5 here]

Two features of Panel A reject the rational benchmark. First, the investment rate is not flat at 50%; it increases monotonically in A in both conditions, indicating that subjects place weight on the salient component even though it is normatively irrelevant. Second, the investment rate exhibits a visible jump at $A = 0$ in both conditions, and the jump is visibly larger in the Large Set Condition than in the Small Set Condition. The slope of the invest-rate schedule is also steeper in the Large Set Condition on both sides of the cutoff. Taken together, these patterns are consistent with **H1** (greater sensitivity to the level of A under higher complexity) and **H2** (greater reliance on the sign of A under higher complexity).

Panel B shows that the discontinuous response is most pronounced precisely when the comparison between the focal asset and the best outside option is hardest. When $|\Delta| = 10$, the investment rate in the Large Set Condition jumps by more than 50% as A crosses zero; the jump becomes less pronounced for $|\Delta| = 20$ and $|\Delta| = 30$. This monotone gradient in $|\Delta|$ supports **H3**.

Table 8 formalizes these observations using linear probability models. Columns (1)–(4) focus on sensitivity to the level of A. Across the full sample and within each payoff-gap subsample ($|\Delta| = 10, 20, 30$), the interaction between A and the Large Set indicator is positive and statistically significant at the 1% level, with coefficient 0.002 that is broadly stable across $|\Delta|$ bins. The Large Set Condition therefore induces a roughly two-fold increase in the sensitivity of investment decisions to the level of A, consistent with **H1**.

[Insert Table 8 here]

Columns (5)–(8) add an indicator A_{pos} for whether A is positive, which captures the categorical (sign-based) component of the decision. The coefficient on A_{pos} is positive and significant in every specification: in the full sample, the investment rate jumps by 17.2 percentage points as A crosses zero (column (5)). The interaction between A_{pos} and the Large Set indicator is positive but not statistically significant in the full sample (column (5)), yet it is large and significant at the 5% level in the $|\Delta| = 10$ subsample (column (6), coefficient 0.205). That is, the additional categorical processing induced by the Large Set Condition is concentrated in the hardest subset of comparisons, exactly the pattern predicted jointly by **H2** and **H3**.

In sum, the experimental evidence provides a clean causal counterpart to the field results. Exogenously increasing the complexity of the decision environment makes both sensitivity to the

level of the salient component A and reliance on the sign of A more pronounced, with the latter concentrated where the comparison is most cognitively demanding.

7.3 Structural Estimation

Setup We embed the experimental decision in the discrete-choice framework of Gabaix and Graeber (2024, Section 5). The investor selects a probability distribution $a = (a_0, a_1)$ over the two actions—invest in the focal asset ($k = 1$) or the outside option ($k = 0$)—to maximize

$$u(a) \;=\; \sum_{k\in\{0,1\}} a_k\, v^k \;-\; \beta\, D(a, a^d),$$

where v^k is the perceived value of action k, $a^d = (a_0^d, a_1^d)$ is a default distribution capturing the investor's pre-deliberation tendency over the two actions, and $D(a, a^d) = \sum_k a_k \log(a_k/a_k^d)$ is the Kullback–Leibler divergence between a and a^d. The parameter $\beta > 0$ governs the cost of moving away from the default: $\beta \to 0$ corresponds to full value maximization, while $\beta \to \infty$ leaves the investor at the default. The first-order condition for a yields the optimal choice probability

$$a_k^* \;=\; \frac{a_k^d \exp\big(v^k/\beta\big)}{\sum_j a_j^d \exp\big(v^j/\beta\big)}, \tag{2}$$

which in log-odds form becomes

$$\log \frac{a_1^*}{a_0^*} \;=\; \log \frac{a_1^d}{a_0^d} \;+\; \frac{v^1 - v^0}{\beta}. \tag{3}$$

The choice log-odds decompose additively into a default tilt and a value contribution. We now specify the default a^d and the value channel v^k in turn.

Default We model the investor's default distribution a^d as depending on the sign of the salient component A. Specifically, we assume

$$\log \frac{a_1^d}{a_0^d} \;=\; \frac{\theta\,\overline{|A|}}{\beta}\,\mathrm{sgn}(A), \tag{4}$$

where $\theta > 0$ measures the strength of the sign-based default $\mathrm{sgn}(A)$ in units of the *mean* salient magnitude $\overline{|A|}$ ($\overline{|A|} \approx 50$ by experimental design), and β is the cost of deviating from the default introduced above.[3] Combined with the constraint $a_1^d + a_0^d = 1$, Equation (4) yields

$$a_1^d \;=\; \frac{\exp\big(\theta\,\overline{|A|}\,\mathrm{sgn}(A)/\beta\big)}{1 + \exp\big(\theta\,\overline{|A|}\,\mathrm{sgn}(A)/\beta\big)}, \qquad a_0^d \;=\; 1 - a_1^d.$$

[3] Choice data identify only the combination $\theta\,\overline{|A|}/\beta$; we adopt this parameterization so that θ is dimensionless and directly interpretable as the size of the sign-based jump relative to the typical salient magnitude.

The default is uniform at $A = 0$ and tilts toward investing for $A > 0$ and away for $A < 0$ by the same amount in either direction. It depends on A only through its sign—not through $|A|$—formalizing the categorical (sign-based) processing established in the empirical sections.

Value Channel We specify the perceived value difference between the focal asset and the best outside option as

$$v^1 - v^0 \;=\; \Delta \;+\; \kappa\, A, \tag{5}$$

where Δ is the true payoff gap defined in Section 7.1 and κA captures the investor's deviation from face-value processing of the salient component. $\kappa = 0$ corresponds to face-value processing, in which A contributes to perceived value at exactly its monetary scale; $\kappa > 0$ corresponds to over-weighting and $\kappa < 0$ to under-weighting. This term plays the role of the superficial first-glance value in Gabaix and Graeber (2024), capturing the visible level of A that enters the investor's value comparison prior to deliberation.

Choice Probability Substituting the default specification (Equation (4)) and the value channel (Equation (5)) into the log-odds form of the decision rule (Equation (3)) yields

$$\log \frac{a_1^*}{a_0^*} \;=\; \frac{\theta\,\overline{|A|}\,\mathrm{sgn}(A)}{\beta} \;+\; \frac{\Delta + \kappa\, A}{\beta} \;=\; \frac{1}{\beta}\Big[\,\theta\,\overline{|A|}\,\mathrm{sgn}(A) \;+ \Delta + \kappa\, A\Big]. \tag{6}$$

The model implies a binary logit in which the linear predictor decomposes additively into three components: the sign-based default term $\theta\,\overline{|A|}\,\mathrm{sgn}(A)$, the rational benchmark Δ, and an over-weighting term κA on the salient component.

Estimation We estimate the parameters (β, κ, θ) by maximum likelihood, separately for the Small Set and Large Set Conditions, using the binary logit implied by Equation (6).

We estimate two specifications. Model 1 is an as-if rational benchmark that imposes $\kappa = 0$ (face-value processing of A) and $\theta = 0$ (no sign-based default), so that A has no special role beyond its contribution to the rational benchmark Δ. Model 2 relaxes both restrictions, allowing κ and θ to be estimated freely.

Results Table 9 reports the estimates. Model 2 strictly dominates the as-if rational benchmark (Model 1) in both conditions: the log-likelihood improves by 60 points in the Small-Set Condition and by more than 200 points in the Large-Set Condition, with corresponding drops in AIC and BIC well in excess of conventional selection thresholds.

[Insert Table 9 here]

The three structural parameters move in directions consistent with the framework of Gabaix and Graeber (2024). First, the cost of deviating from the default, β, roughly doubles across conditions, rising from approximately 14 in the Small-Set Condition to approximately 30 in the Large-Set

Condition in Model 2. This is the central comparative static predicted by the framework: as the environment becomes more cognitively demanding, the investor's KL cost of moving away from the default increases, pulling behavior closer to the default and farther from the value-maximizing benchmark.

Second, the sign-based default specified in Equation (4) operates with greater force across conditions: θ rises from 0.15 to 0.35, corresponding to a sign-based jump in perceived utility of approximately 7.5 (0.15×50) in the Small-Set Condition and 17.5 (0.35×50) in the Large-Set Condition—both large relative to the $|\Delta| \in \{10, 20, 30\}$ comparisons subjects face. This supports **H2** and is the structural component that distinguishes the present specification from a standard logit.

Third, sensitivity to the level of the salient component A also strengthens: κ rises from 0.01 in the Small-Set Condition to 0.20 in the Large-Set Condition, indicating that subjects over-weight A only when the environment becomes cognitively demanding—consistent with **H1**.

The relative weight placed on the sign-based cue versus the rational benchmark can be summarized by a single dimensionless measure, facilitating comparisons across treatments. We define

$$\alpha \equiv \frac{\theta\,\overline{|A|}}{\theta\,\overline{|A|} + |\Delta|} = \frac{\theta\,\frac{\overline{|A|}}{|\Delta|}}{\theta\,\frac{\overline{|A|}}{|\Delta|} + 1} \tag{7}$$

as the share of decision weight that the investor places on the sign of the salient component, with $1 - \alpha$ the complementary share placed on the rational benchmark. The ratio $\frac{\overline{|A|}}{|\Delta|}$ captures the salience of the default mode relative to the payoff difference: a larger value indicates that the salient component is larger relative to the difficulty of distinguishing the value of the focal asset from that of the outside option, making the sign-based cue more salient. The parameter θ captures the investor's reliance on the salient component. Thus, α increases with both the relative salience of the default mode and the investor's reliance on salient information.

Plugging in our estimates, α ranges from 0.2 (when $|\Delta| = 30$) to 0.43 (when $|\Delta| = 10$) in the Small Set Condition, and from 0.37 (when $|\Delta| = 30$) to 0.64 (when $|\Delta| = 10$) in the Large-Set Condition—an approximate doubling that summarizes the central message of this section: as the decision environment becomes more cognitively demanding, investors place substantially greater weight on the categorical, sign-based component of the decision.

8 Discussion and Conclusion

We close by drawing out two implications of our findings—for the structure of retail demand for financial assets (Section 8.1) and investor welfare cost of categorical trading (Section 8.2)—before concluding (Section 8.3).

8.1 Structure of Retail Demand

Beyond their direct investor welfare costs, our findings have implications for the structure of retail demand for financial assets. Standard frictionless asset-pricing models imply that demand responds smoothly to prices and fundamentals. A large literature, however, shows that demand is not perfectly elastic and that limits to arbitrage, transaction costs, and investor frictions can generate substantial price effects of non-fundamental demand (Shleifer, 1986; Koijen and Yogo, 2019; Gabaix and Koijen, 2022). While the first two mechanisms have received extensive empirical attention, cognitive constraints provide a less explored psychological foundation for such departures from frictionless demand in financial markets. Our findings identify a distinct mechanism: when decision complexity exceeds cognitive capacity, investors compress continuous price signals into coarse categories, generating locally discontinuous responses to otherwise continuous price changes. In contrast to cognitive noise, which smoothly attenuates the transmission from beliefs to actions (Charles et al., 2024), our mechanism changes the functional form of the response itself, producing discrete jumps at salient boundaries whose magnitude increases with decision complexity.

Two features of our evidence are particularly relevant for this interpretation. First, the departure from smooth demand is localized: trading probability jumps at the zero threshold even when the underlying price movement is arbitrarily small. Second, it is state-dependent: the magnitude of the jump systematically increases with decision complexity and is larger among investors with lower cognitive capacity. Thus, cognitive constraints do not simply make investors less responsive to information; they can change how information is represented and thereby generate non-smooth responses to continuous economic signals.

Whether these individual-level discontinuities aggregate into measurable discontinuities in fund flows, whether they extend to institutional investors, and whether similar patterns arise at other salient thresholds remain open questions. More broadly, our findings suggest a promising direction for connecting the psychology of information representation to the demand inelasticity documented in asset markets.

8.2 Investor Welfare Costs of Categorical Trading

We assess the welfare implications of categorical (sign-based) processing by first asking whether the zero-crossing jump carries any information about future fund performance. If it does not, the excess trades it triggers represent pure noise: they impose transaction costs on investors without improving portfolio allocation. Under two-way clustering by product and year-month, the coefficient on the positive price-change indicator is statistically indistinguishable from zero across all specifications—whether measuring raw or risk-adjusted 30-day forward returns, and whether or not we control for the magnitude of the price change (Appendix Table A.5). The sign of a small daily price change carries no information about subsequent fund performance, so the excess purchases triggered by the zero crossing are informationally empty.

We quantify investor welfare cost using the full sample of browsing observations on the platform. The buy-side RD estimate implies approximately 2,255 excess purchases per year—trades that

would not have occurred had the price change been marginally positive rather than marginally negative.[4] Each such trade generates a direct cost of approximately RMB 6 under the discounted subscription fee of 0.15% typical of online platforms, and approximately RMB 62 under the standard undiscounted rate of 1.5%. Aggregated across the platform's roughly 13,000 active investors, this amounts to a per-investor annual cost of approximately RMB 1 to RMB 10—a lower bound, as these figures exclude redemption fees on subsequent exits and management fees accrued during the holding period.[5]

Scaling these per-investor estimates to China's roughly 60 million active mutual fund investors yields an aggregate welfare cost on the order of RMB 60 million to RMB 600 million per year. An important feature of this estimate is that the underlying excess trading is more pronounced where complexity is higher: investors with larger consideration sets, less concentrated portfolios, and lower educational attainment, as well as trades involving more complex funds, all exhibit larger discontinuity. The welfare cost is therefore not uniform across the investor population but disproportionately borne by investors and products for which categorical processing has the strongest grip. This pattern provides an evidence-based rationale for simplifying decision environments, e.g., through interface design and product disclosure, as a way to reduce complexity-induced trading errors. We emphasize, however, that this calculation is an order-of-magnitude estimate, as it assumes that the categorical response documented in our sample is representative of the broader population of retail mutual fund investors, an assumption that we cannot directly verify.

8.3 Conclusion

Across field and experimental evidence, the paper documents a common mechanism operating across different decision settings. When the information structure of a decision exceeds investors' cognitive capacity, they compress continuous signals into coarse categories and act on those categories rather than on the underlying signal. This compression manifests as discrete jumps in trading behavior at salient boundaries—between "down" and "up" days for purchase decisions and between "loss" and "gain" positions for sell decisions. Three pieces of evidence point to cognitive constraints rather than stable preferences as the source of this compression: the discontinuity strengthens with decision complexity, it strengthens rather than weakens with investor effort, and it emerges under exogenous variation in complexity in a controlled experiment with a flat rational benchmark. Taken together, our findings show that complexity affects not merely the magnitude of investors' responses to information but the *functional form* of those responses, transforming a continuous mapping into a categorical one.

This mechanism also offers a new perspective on the disposition effect (Section 5). To the

[4] This figure applies the RD estimate of 0.3 percentage points as in Table 2 (after absorbing investor and fund fixed effects) to all negative-price-change browsing observations, unrestricted by bandwidth (686,116 investor × fund × day observations), scaled to a full year by a factor of 365/333.

[5] Redemption fees on Chinese mutual fund platforms typically follow a tiered schedule (e.g., 1.5% if held less than seven days, 0.5% if held less than six months, zero beyond one year). Including redemption costs would scale the per-trade estimate upward by an amount that depends on the holding-period distribution of these excess trades, which we do not pursue here.

extent that gain–loss asymmetry reflects categorical processing under complexity rather than stable preferences, simplifying the decision environment—for example, through interface design or product disclosure—could attenuate disposition-type behavior. This prediction is not implied by preference-based accounts and provides a potentially testable avenue for distinguishing the two explanations.

More broadly, our findings connect the psychology of information representation to the structure of financial demand and investor welfare. By showing that cognitive constraints can generate localized, complexity-dependent departures from smooth responses to prices, the paper provides a microfoundation for a previously less explored form of non-smooth investor demand. Several questions remain open: whether individual-level kinks aggregate into measurable discontinuities in fund-level demand, whether the mechanism extends to professional investors, and whether similar categorical responses arise at other psychologically salient thresholds in financial markets and beyond.

References

Banovetz, James and Ryan Oprea, "Complexity and Procedural Choice," *American Economic Journal: Microeconomics*, 2023, *15* (2), 384–413.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience Theory of Choice under Risk," *Quarterly Journal of Economics*, 2012, *127* (3), 1243–1285.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Cattaneo, Matias D, Michael Jansson, and Xinwei Ma, "Two-step estimation and inference with possibly many included covariates," *The Review of Economic Studies*, 2019, *86* (3), 1095–1122.

_ , _ , and _ , "Simple local polynomial density estimators," *Journal of the American Statistical Association*, 2020, *115* (531), 1449–1455.

Célérier, Claire and Boris Vallée, "Catering to Investors through Security Design: Headline Rate and Complexity," *Quarterly Journal of Economics*, 2017, *132* (3), 1469–1508.

Charles, Constantin, Cary Frydman, and Mete Kilic, "Insensitive Investors," *Journal of Finance*, 2024, *79* (4), 2473–2503.

Enke, Benjamin and Charles Shubatt, "Quantifying Lottery Choice Complexity," Technical Report w31677, National Bureau of Economic Research 2023.

_ and Thomas Graeber, "Cognitive Uncertainty," *Quarterly Journal of Economics*, 2023, *138* (4), 2021–2067.

_ , _ , Ryan Oprea, and Jing Yang, "Behavioral Attenuation," Technical Report w32973, National Bureau of Economic Research 2024.

Frydman, Cary and Lawrence J. Jin, "Efficient Coding and Risky Choice," *Quarterly Journal of Economics*, 2022, *137* (1), 161–213.

Gabaix, Xavier, "A Sparsity-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 2014, *129* (4), 1661–1710.

_ and Ralph S. J. Koijen, "In Search of the Origins of Financial Fluctuations: The Inelastic Markets Hypothesis," Working Paper 28967, National Bureau of Economic Research 2022.

_ and Thomas Graeber, "The Complexity of Economic Decisions," Technical Report w33109, National Bureau of Economic Research 2024.

Graeber, Thomas, Christopher Roth, and Marco Sammon, "Coarse Categories in a Complex World," Technical Report 364, University of Bonn and University of Cologne 2025.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw, "Identification and Estimation of Treatment Effects with a Regression Discontinuity Design," *Econometrica*, 2001, *69* (1), 201–209.

Imbens, Guido W. and Thomas Lemieux, "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 2008, *142* (2), 615–635.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 363–391.

Koijen, Ralph S. J. and Motohiro Yogo, "A Demand System Approach to Asset Pricing," *Journal of Political Economy*, 2019, *127* (4), 1475–1515.

Lee, David S. and Thomas Lemieux, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 2010, *48* (2), 281–355.

Odean, Terrance, "Are Investors Reluctant to Realize Their Losses?," *Journal of Finance*, 1998, *53* (5), 1775–1798.

Oprea, Ryan, "What Makes a Rule Complex?," *American Economic Review*, 2020, *110* (12), 3913–3951.

_ , "Complexity and Its Measurement," 2024. Forthcoming in Handbook of Experimental Methods in the Social Sciences.

_ , "Decisions under Risk Are Decisions under Complexity," *American Economic Review*, 2024, *114* (12), 3789–3811.

Puri, Indira, "Simplicity and risk," *The Journal of Finance*, 2025, *80* (2), 1029–1080.

Quispe-Torreblanca, Edwin, John Gathergood, George Loewenstein, and Neil Stewart, "Investor Logins and the Disposition Effect," *Management Science*, 2025, *71* (1), 219–239.

Shefrin, Hersh and Meir Statman, "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence," *Journal of Finance*, 1985, *40* (3), 777–790.

Shleifer, Andrei, "Do Demand Curves for Stocks Slope Down?," *The Journal of Finance*, 1986, *41* (3), 579–590.

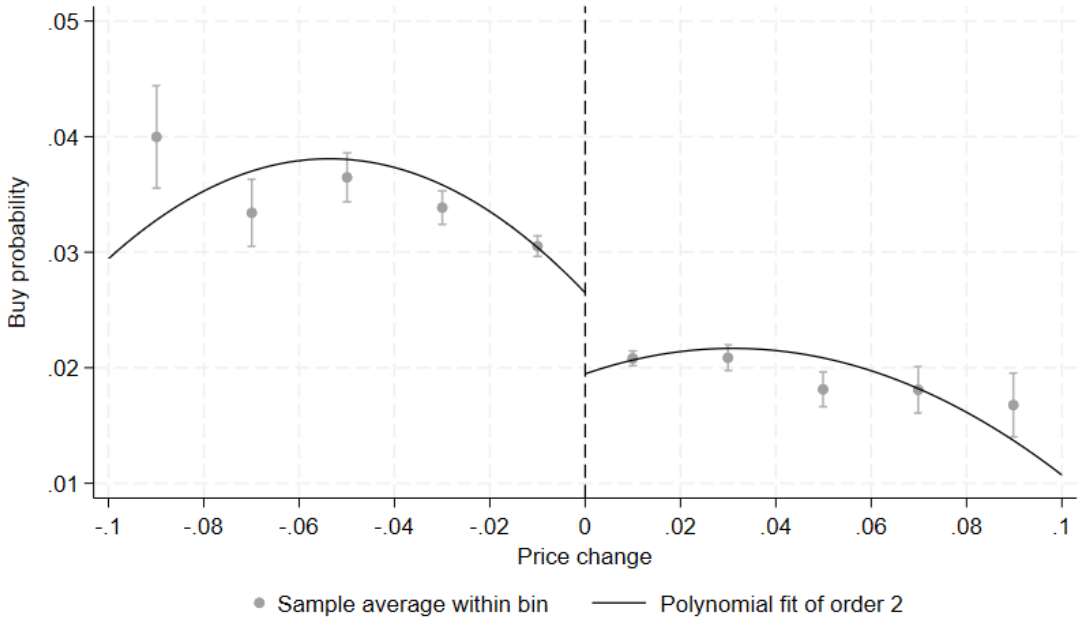
Figures

Figure 1: Manipulation Test at the Cutoff



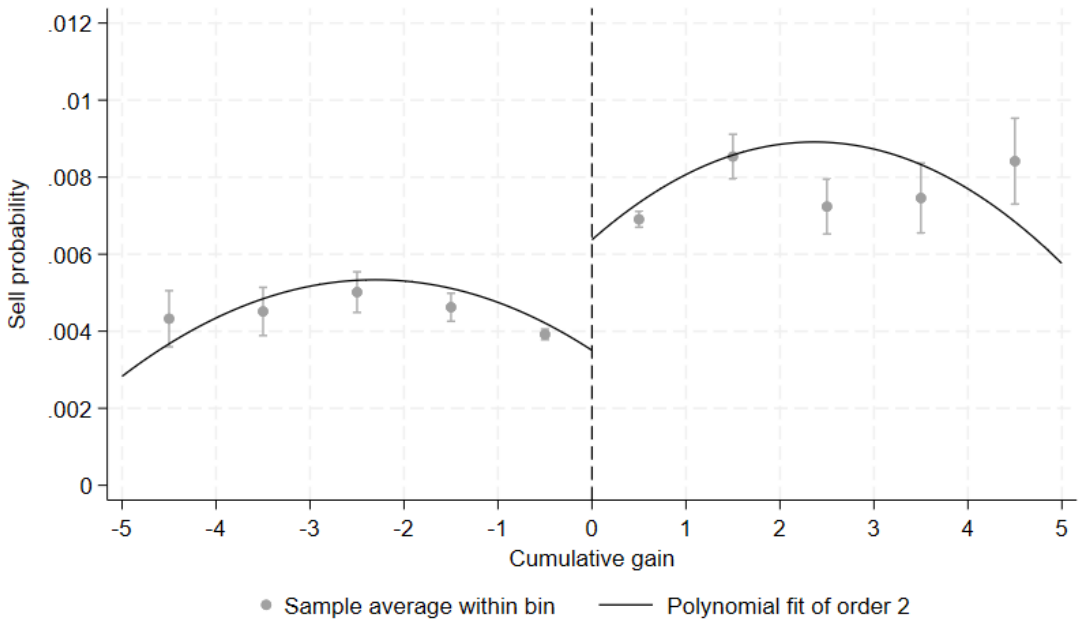
Note: In this figure, we provide a visual representation of the continuity of the running variable around the cutoff of zero. Specifically, we implement manipulation test of one-day price changes of traded assets at the cutoff of zero using the second-order local polynomials with a triangular kernel function Cattaneo et al. (2020). The solid curve represents the local polynomial density estimators and the shade area around the curve indicates 95% confidence intervals. Below the estimators is the histogram of one-day price changes.

Figure 2: Regression Discontinuity: Purchase Probability around Zero Price Change



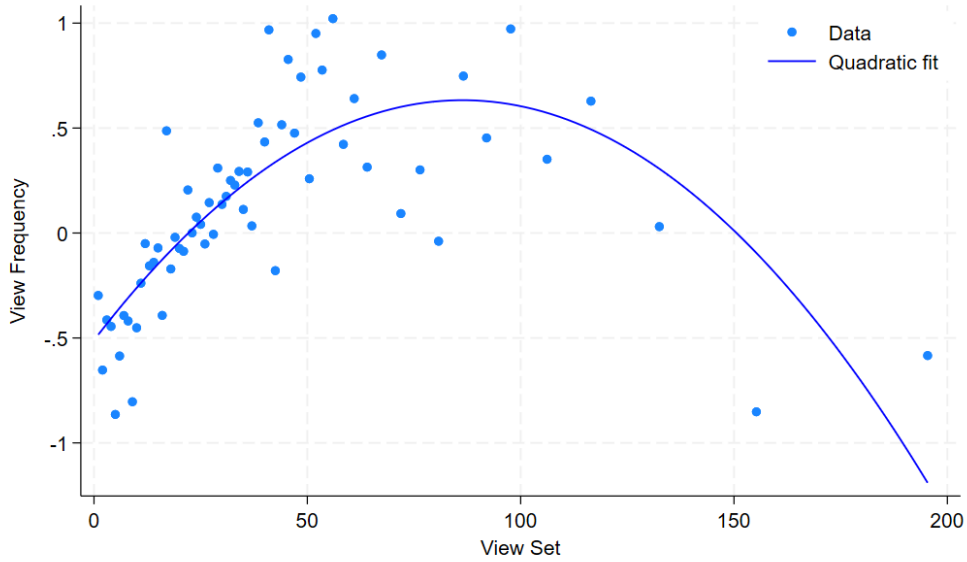
Note: This figure provides a graphical illustration of the treatment effect at the zero-price-change threshold. The running variable (one-day price change) is partitioned into bins of width 0.02. Each point represents the average purchase likelihood for observations within a bin, with vertical bars indicating 95% confidence intervals. The solid lines show separate second-order polynomial fits estimated on either side of the cutoff. Points to the left (right) of zero correspond to negative (positive) one-day price changes.

Figure 3: Regression Discontinuity: Sell Probability around Zero Cumulative Gain

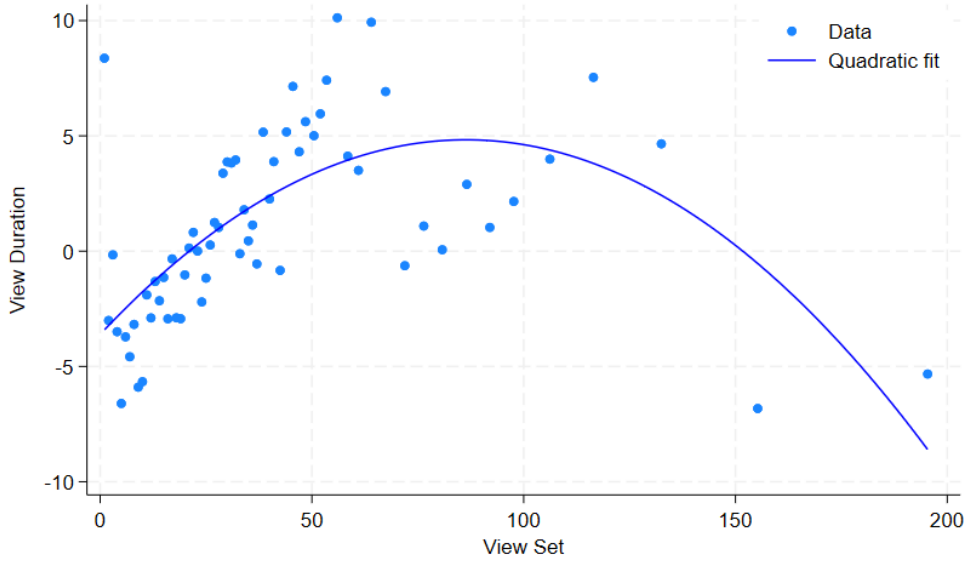


Note: This figure provides a graphical illustration of the treatment effect at the zero-cumulative-gain threshold. The running variable (cumulative gain) is partitioned into bins of width 1. Each point represents the average sell likelihood for observations within a bin, with vertical bars indicating 95% confidence intervals. The solid lines show separate second-order polynomial fits estimated on either side of the cutoff. Points to the left (right) of zero correspond to negative (positive) cumulative gain.

Figure 4: Product-level Attention and Decision Complexity



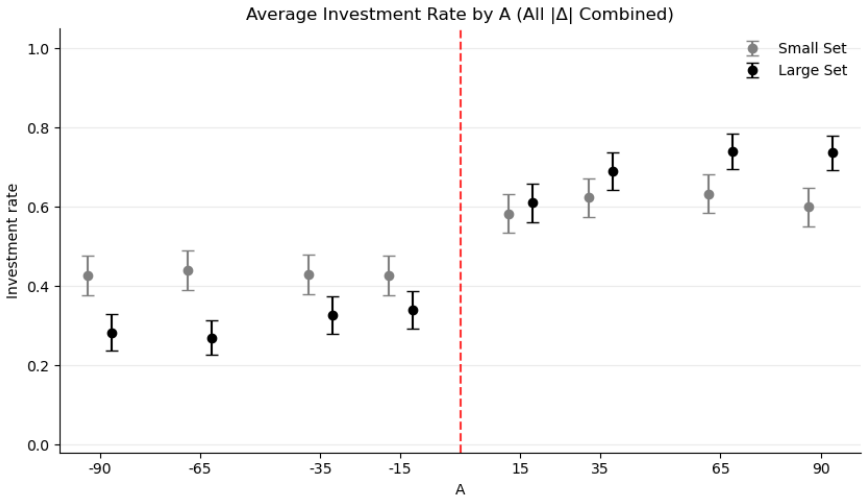
Panel A. View Frequency



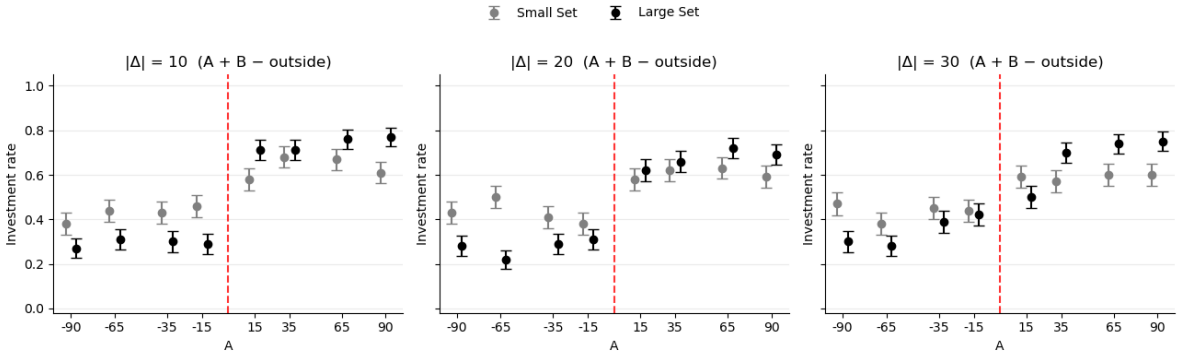
Panel B. View Duration

Note: This figure illustrates the relationship between product-level attention effort and decision complexity. Attention effort is proxied by product-level view frequency (Panel A) and view duration (Panel B), while decision complexity is measured by the size of the consideration set. We first residualize the attention measures by regressing them on investor, product, year, and month fixed effects, consistent with the main analysis. We then sort observations into 100 equally sized bins based on consideration set size. The figure plots the mean residual attention measure within each bin. The solid curve is obtained from a quadratic regression of the residualized attention measure on consideration set size.

Figure 5: Experiment Results



Panel A. Full Sample



Panel B. Results by Δ

Note: This figure presents the experimental results. Panel A reports the average investment rate across different values of A for the full sample under Large and Small Set conditions, while Panel B breaks down the investment rate by levels of Δ. Vertical bars represent standard errors.

Tables

Table 1: Summary Statistics

Variable	N	mean	sd	p5	p50	p95
Panel A. Account Level						
Age	11713	46.16	12.59	30	40	70
Gender (F)	11713	0.44	0.50	0	0	1
Holding (RMB)	11713	198208.62	562774.80	1314.25	51769.88	818988.38
Number of products	11713	8.39	23.28	1	5	24
Monthly transaction	11713	7.28	18.15	0	4	23
Panel B. Buy Sample (Investor × Asset × Potential Buy Day)						
Price change	605727	-0.0004	0.0426	-0.0770	0.0001	0.0750
Buy Dummy	605727	0.0267	0.1613	0	0	0
Transaction amount (RMB)	16191	4818.14	39645.68	3.00	400.00	14250.00
Consideration set (one week prior to the potential trading day)						
Browse freq of the asset	605727	7.27	14.08	1	3	24
Browse length of the asset (Second)	605727	91.46	134.73	3	36	357
Total browse freq	605727	142.64	282.35	3	49	590
Total browse length (Second)	605727	1608.45	3459.10	40	644	5546
Complexity						
Set size	605727	48.89	90.93	3	22	172
Portfolio HHI	605727	0.2229	0.2320	0.0289	0.1409	0.8105
Fee structure	605231	4.69	1.58	3.08	5.38	6.92
Document complexity	605231	2.90	2.62	0.23	3.57	6.45
Panel C. Sell Sample (Investor × Asset × Potential Sell Day)						
Cumulative gain change (000 RMB)	2245474	0.2083	5.3755	-7.7379	-0.0054	9.4984
Sell Dummy	2245474	0.0056	0.0745	0	0	0
Transaction amount (RMB)	12531	11535.09	83903.58	11.95	1202.07	43828.58
Consideration set (one week prior to the potential trading day)						
Browse freq of the asset	2245474	3.54	9.71	0	1	14
Browse length of the asset (Second)	2245474	40.03	91.52	0	6	196
Total browse freq	2245474	58.75	146.01	0	8	293
Total browse length (Second)	2245474	571.52	1447.84	0	92	2617
Complexity						
Set size	2245474	30.69	52.72	1	15	110
Portfolio HHI	2245474	0.1532	0.1585	0.0194	0.1051	0.4648
Fee structure	2245461	4.70	1.58	3.08	5.38	6.92
Document complexity	2245461	2.90	2.64	0.23	3.53	6.58

Note: The table presents the summary statistics of the variables used in our analysis.

Table 2: Regression Discontinuity: Buy Decisions

Dep. var.: Buy	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pos Price Chg	-0.003**	-0.002*	-0.003***	-0.002*	-0.004***	-0.002*	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Bandwidth	0.011 (opt.)	0.015	0.5	0.008 (opt.)	0.5	0.002 (opt.)	0.5
Polynomial	Order 2	Order 2	Order 2	Order 1	Order 1	N	N
Individual FE	Y	Y	Y	Y	Y	Y	Y
Product FE	Y	Y	Y	Y	Y	Y	Y
Year Month FE	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.111	0.113	0.114	0.109	0.114	0.115	0.114
N	235,195	284,120	601,454	188,172	601,454	81,072	601,454

Note: This table presents the relationship between the likelihood of an investor purchasing an asset and the price change of that asset within a single day. *Buy* is a dummy variable that equals one if the asset is bought and zero if it is not. *Pos Price Chg* is a dummy variable that equals one if the change in price since the previous day is greater than or equal to zero. We control polynomials, different for each side, of order two in columns (1)–(3), order one in columns (4)–(5), and no polynomials in columns (6)–(7). We show the RD regressions at the optimal and other bandwidths: column (1) uses the optimal bandwidth of 0.011 selected under a uniform kernel, while column (2) uses a wider bandwidth of 0.015 selected under a triangular kernel for robustness. Columns (4) and (6) adopt bandwidths of 0.008 and 0.002, optimal under the respective polynomial functions. Columns (3), (5), and (7) use a global support of 0.5 for an examination of external validity. Investor, product, year, and month fixed effects are controlled in all specifications. Standard errors are clustered at the investor level. * p<0.1, ** p<0.05, *** p<0.01.

Table 3: Buy Decisions: Complexity Heterogeneity

Dep. var.: Buy	Set Size	Portfolio HHI	Fee Structure	Document Complexity	Education	Profession
	(1)	(2)	(3)	(4)	(5)	(6)
Pos. Price Chg × Complex	−0.0036*	−0.0073***	−0.0025*	−0.0038***	−0.0041**	−0.0053***
	(0.0019)	(0.0018)	(0.0014)	(0.0013)	(0.0019)	(0.0016)
Pos. Price Chg	−0.0011	0.0003	−0.0025*	−0.0020	−0.0013	0.0014
	(0.0016)	(0.0013)	(0.0014)	(0.0015)	(0.0014)	(0.0015)
Complex	0.0029	0.0037*	—	—	—	—
	(0.0018)	(0.0022)				
Polynomial	Y	Y	Y	Y	Y	Y
Individual FE	Y	Y	Y	Y	Y	Y
Product FE	Y	Y	Y	Y	Y	Y
Year-Month FE	Y	Y	Y	Y	Y	Y
Adj. R^2	0.111	0.111	0.114	0.114	0.100	0.111
N	235,195	235,195	228,326	228,326	205,355	235,195

Note: This table reports a pooled two-way interaction regression. *Buy* is a dummy variable that equals one if the asset is bought and zero otherwise. *Pos Price Chg* is a dummy variable that equals one if the change in price since the previous day is greater than or equal to zero. *Complex* is a dummy variable that indicates more complex decisions. Columns (1)–(2) report decision-environment complexity related to asset choices, Columns (3)–(4) product complexity, and Columns (5)–(6) investor cognitive capacity. The consideration set consists of assets the investor viewed on the trade day and during the preceding week. In Column (1), *Complex* equals one if consideration set size is above the sample median; in Column (2), it equals one if portfolio holding HHI is below the sample median. In Columns (3)–(4), *Complex* equals one if the fund's fee structure or document complexity score strictly exceeds the median within its fund type, respectively. For Columns (5)–(6), *Complex* equals one if the investor does not hold a college degree (*Education*), or if the investor is not employed in the financial industry (*Profession*). "—" indicates the group main effect is collinear with an absorbed fixed effect (product FE for Fee Structure and Document Complexity; investor FE for Education and Profession) and is therefore not separately identified. All specifications use the optimal bandwidth (0.011), and include a second-order polynomial, different on both sides, and investor, product, and year-month fixed effects. Standard errors are clustered at the investor level. * p<0.1, ** p<0.05, *** p<0.01.

Table 4: Regression Discontinuity: Sell Decisions

Dep. var.: Sell	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pos. Cum. Perf.	0.0008**	0.0008***	0.0011***	0.0010***	0.0011***	0.0013***	0.0013***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0002)
Bandwidth	0.813 (opt.)	0.954 (opt.)	10	0.538 (opt.)	10	0.142 (opt.)	10
Polynomial	Order 2	Order 2	Order 2	Order 1	Order 1	N	N
Individual FE	Y	Y	Y	Y	Y	Y	Y
Product FE	Y	Y	Y	Y	Y	Y	Y
Year Month FE	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.025	0.026	0.027	0.024	0.027	0.021	0.027
N	1,341,687	1,396,999	2,049,548	1,200,119	2,049,548	780,732	2,049,548

Note: This table presents the relationship between the likelihood of selling an asset and cumulative gains. *Sell* is a dummy variable equal to one if the asset is sold and zero otherwise. *Pos. Cum. Perf.* equals one if cumulative gain measured in thousand RMB is greater than or equal to zero. We control polynomials, different for each side, of order two in columns (1)–(3), order one in columns (4)–(5), and no polynomials in columns (6)–(7). We show the RD regressions at the optimal and other bandwidths: column (1) uses the optimal bandwidth of 0.813 selected under a uniform kernel, while column (2) uses an alternative optimal bandwidth of 0.954 selected under a triangular kernel. Columns (4) and (6) adopt bandwidths of 0.538 and 0.142, optimal under the respective polynomial functions. Columns (3), (5), and (7) use a global support of 10 for an examination of external validity. Investor, product, year, and month fixed effects are controlled in all specifications. Standard errors are clustered at the investor level. * p<0.1, ** p<0.05, *** p<0.01.

Table 5: Sell Decisions: Complexity Heterogeneity

Dep. var.: Sell	Set Size	Portfolio HHI	Fee Structure	Document Complexity	Education	Profession
	(1)	(2)	(3)	(4)	(5)	(6)
Pos. Cum. Perf. × Complex	0.0010**	0.0012*	0.0003	0.0000	−0.0012	0.0007
	(0.0005)	(0.0007)	(0.0004)	(0.0004)	(0.0008)	(0.0006)
Pos. Cum. Perf.	0.0002	0.0002	0.0007*	0.0008**	0.0015**	0.0002
	(0.0004)	(0.0005)	(0.0004)	(0.0004)	(0.0007)	(0.0005)
Complex	0.0002	0.0011**	—	—	—	—
	(0.0003)	(0.0005)				
Polynomial	Y	Y	Y	Y	Y	Y
Individual FE	Y	Y	Y	Y	Y	Y
Product FE	Y	Y	Y	Y	Y	Y
Year Month FE	Y	Y	Y	Y	Y	Y
Adj. R^2	0.025	0.025	0.025	0.025	0.033	0.028
N	1,341,687	1,341,687	1,341,566	1,341,566	907,659	751,744

Note: This table reports a pooled two-way interaction regression. *Sell* is a dummy variable that equals one if the asset is sold and zero otherwise. *Pos. Cum. Perf.* is a dummy variable that equals one if cumulative gains are greater than or equal to zero. *Complex* is a dummy variable that indicates more complex decisions. Columns (1)–(2) report decision-environment complexity related to asset choices, Columns (3)–(4) product complexity, and Columns (5)–(6) investor cognitive capacity. The consideration set consists of assets the investor viewed on the trade day and during the preceding week. In Column (1), *Complex* equals one if consideration set size is above the sample median; in Column (2), it equals one if portfolio holding HHI is below the sample median. In Columns (3)–(4), *Complex* equals one if the fund's fee structure or document complexity score strictly exceeds the median within its fund type, respectively. For Columns (5)–(6), *Complex* equals one if the investor does not hold a college degree (*Education*), or if the investor is not employed in the financial industry (*Profession*). "—" indicates the group main effect is collinear with an absorbed fixed effect (product FE for Fee Structure and Document Complexity; investor FE for Education and Profession) and is therefore not separately identified. All specifications use the optimal bandwidth (0.813), a second-order polynomial, different on both sides, and investor, product, and year-month fixed effects. Standard errors are clustered at the investor level. * p<0.1, ** p<0.05, *** p<0.01.

Table 6: Regression Discontinuity: Product-level Attention Effort Heterogeneity

	Buy		Sell	
Dep. var.	Frequency	Duration	Frequency	Duration
	(1)	(2)	(3)	(4)
Pos. Signal × Effort	-0.0105***	-0.0032**	0.0035***	0.0029***
	(0.0023)	(0.0013)	(0.0007)	(0.0006)
Pos. Signal	0.0012	-0.0016	-0.0007	-0.0004
	(0.0012)	(0.0015)	(0.0004)	(0.0003)
Effort	0.0324***	0.0199***	0.0043***	0.0037***
	(0.0035)	(0.0024)	(0.0006)	(0.0005)
Bandwidth	0.011	0.011	0.813	0.813
Polynomial	Y	Y	Y	Y
Individual FE	Y	Y	Y	Y
Product FE	Y	Y	Y	Y
Year Month FE	Y	Y	Y	Y
Adj. R^2	0.117	0.114	0.027	0.026
N	235,195	235,195	1,341,687	1,341,687

Note: This table reports a pooled two-way interaction regression testing whether the RD discontinuity is larger among investors who exert more cognitive effort, proxied by product-level browsing frequency (columns (1), (3)) and browsing duration (columns (2), (4)) in the week before a potential trading day. *Buy* (resp. *Sell*) is a dummy variable that equals one if the asset is bought (resp. sold) and zero otherwise. *Pos. Signal* is a dummy variable that equals one if the change in price since the previous day (Buy, columns (1)–(2)) or the cumulative gain since purchase (Sell, columns (3)–(4)) is greater than or equal to zero. *Effort* is a dummy variable that equals one if browsing frequency or duration exceeds the sample median. We control for a second-order polynomial in the forcing variable, separately on each side of the cutoff. Investor, product, and year-month fixed effects are included in all columns. Standard errors are clustered at the investor level. * p<0.1, ** p<0.05, *** p<0.01.

Table 7: Product-level Attention Effort and Decision Complexity

	View Frequency				View Duration			
Dep. var.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
View set	2.109***	5.988***	1.964***	5.405***	18.414***	45.151***	16.586***	39.620***
	(0.386)	(0.973)	(0.337)	(0.899)	(3.505)	(8.822)	(2.856)	(7.885)
View set^2		-1.839***		-1.625***		-12.676***		-10.875***
		(0.506)		(0.431)		(4.529)		(3.787)
Bandwidth	0.011	0.011	0.5	0.5	0.011	0.011	0.5	0.5
Polynomial	Y	Y	Y	Y	Y	Y	Y	Y
Individual FE	Y	Y	Y	Y	Y	Y	Y	Y
Product FE	Y	Y	Y	Y	Y	Y	Y	Y
Year Month FE	Y	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.448	0.449	0.468	0.469	0.324	0.324	0.322	0.322
Observations	601,454	601,454	235,195	235,195	601,454	601,454	235,195	235,195

Note: This table examines the relationship between attention effort and decision complexity for purchase decisions. Attention effort is measured by product-level view frequency in columns (1)–(4) and view duration in columns (5)–(8), while decision complexity is proxied by the size of the consideration set (in hundreds of funds). We regress the attention measures on the consideration set size and its quadratic term. Columns (1), (2), (5), and (6) use the optimal bandwidth of 0.011 around the cutoff, whereas columns (3), (4), (7), and (8) use a broader bandwidth of 0.5. All specifications control for separate second-order polynomials in the forcing variable on either side of the cutoff, as well as investor, product, and year-month fixed effects. Standard errors are clustered at the investor level. * p<0.1, ** p<0.05, *** p<0.01.

Table 8: Experiment Result

Dep. Var: Invest	(1) All	(2) $\|\Delta\| = 10$	(3) $\|\Delta\| = 20$	(4) $\|\Delta\| = 30$	(5) All	(6) $\|\Delta\| = 10$	(7) $\|\Delta\| = 20$	(8) $\|\Delta\| = 30$
A_value	0.001***	0.002***	0.001**	0.001***	0.000	0.001**	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Large set	-0.021	-0.019	-0.040	-0.005	-0.052	-0.121**	-0.074	0.038
	(0.019)	(0.023)	(0.025)	(0.021)	(0.034)	(0.051)	(0.049)	(0.045)
Large set × *A_value*	0.002***	0.002***	0.002***	0.002***	0.001***	0.000	0.001*	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
A_pos					0.172***	0.135**	0.217***	0.160***
					(0.047)	(0.066)	(0.061)	(0.050)
Large set × *A_pos*					0.060	0.205**	0.068	-0.086
					(0.065)	(0.093)	(0.084)	(0.085)
N	4,762	1,591	1,582	1,589	4,762	1,591	1,582	1,589
Adj. R^2	0.076	0.093	0.073	0.061	0.085	0.107	0.088	0.063

Note: This table presents the regression of participants' investment decision on the A values and the treatment High complexity. *Invest* is a dummy variable that equals one if the participant decides to invest in the focal asset and zero if in outside options. *A_value* is the value of A. *A_pos* is a dummy variable that equals one if *A_value* ≥ 0. High is a dummy variable that equals one if the participant is from the Large Set condition. Standard errors are clustered at the participant level. * p<0.1, ** p<0.05, *** p<0.01.

Table 9: Structure Estimation

	Small set		Large set	
	Model 1	Model 2	Model 1	Model 2
β	14.79	13.84	37.02	30.06
κ		0.01		0.20
θ		0.15		0.35
Loglikelihood	-1197.49	-1137.74	-1552.56	-1338.46
AIC	2396.99	2281.49	3107.12	2682.92
BIC	2402.77	2298.82	3112.89	2700.24

Appendix

A Variable Definition

Variable	Definition
Account, Trading, and Browsing Variables	
Buy	An indicator variable that equals 1 if the asset is purchased on the potential trading day, and 0 otherwise.
Sell	An indicator variable that equals 1 if the asset is sold on the potential trading day, and 0 otherwise.
Pos Price Chg	An indicator variable that takes the value of 1 if the price change from the day before the potential trading day to the potential trading day is greater than or equal to 0, and 0 otherwise.
Pos. Cum. Perf.	An indicator variable that equals 1 if the cumulative gain since purchase is greater than or equal to 0, and 0 otherwise.
Price Change	The one-day change in the asset's price relative to the price the investor most recently browsed (the running variable in the main buy-side RD specifications).
Cumulative Gain Change	The investor's cumulative gain (in thousand RMB) on a held position since purchase (the running variable in the main sell-side RD specifications).
Transaction Amount	The RMB value of the buy (or sell) transaction, conditional on a transaction occurring.
Set Size	An investor's consideration set is the number of assets viewed in the week before a potential trading day. Set Size is the number of assets in this consideration set.
Browse Freq	The number of times the asset was viewed in the week up to the potential trading day.
Browse Length	The total duration (in seconds) spent viewing the asset in the week up to the potential trading day.
Annual std	The standard deviation of the asset's price over the one-year period up to the potential trading day.
Holding	The average value of portfolio held by the investor during the observed sample period.
Moderating Variables of Complexity	
Set Size (Large/Small)	If Set Size is below the sample median consideration set size, it is classified as Small; if above or equal, Large.
Portfolio HHI (Large/Small)	Small HHI refers to the Herfindahl–Hirschman Index of the investor's portfolio being below the median of observations in the same month, and Large HHI otherwise. If the investor does not hold any asset in that month, the HHI is set to 1.
Fee Structure (High/Low)	High if the fund's fee-structure complexity score strictly exceeds the median within its fund type; Low otherwise.
Document Complexity (High/Low)	High if the fund's document-complexity score strictly exceeds the median within its fund type; Low otherwise.
Education (College/Not)	College if the investor has a college degree or higher; Not College otherwise.

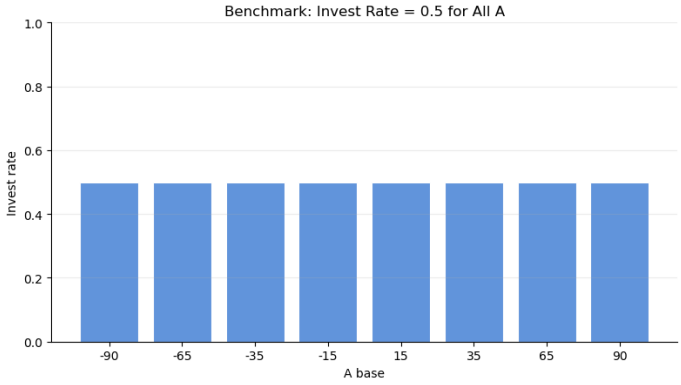
Continued on next page

Variable	**Definition**
Profession (Finance/Not)	Finance if the investor is employed in the financial industry; Non-finance otherwise.
Effort — Frequency-based (High/Low)	If the number of times an investor views an asset in the week before a potential trading day exceeds the sample median, it is classified as High; otherwise, Low.
Effort — Duration-based (High/Low)	If the total time an investor spends viewing an asset in the week before a potential trading day exceeds the sample median, it is classified as High; otherwise, Low.

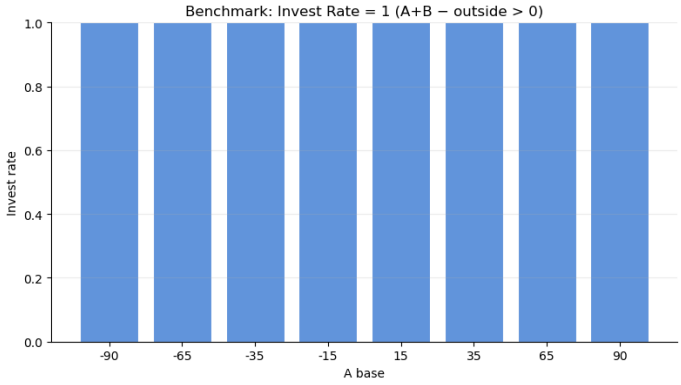
B Additional Figures

Figure A.1: Rational Benchmarks for Experiments

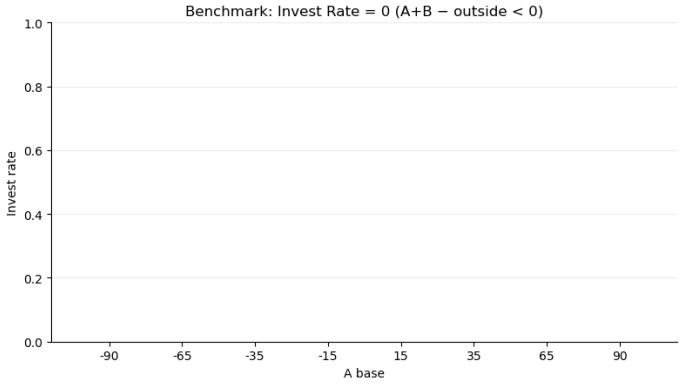
Panel A: Rational Benchmark for all Decision Problems



Panel B: Rational Benchmark when Focal Asset Dominants (Positive Δ)



Panel C: Rational Benchmark when Focal Asset Dominants (Negative Δ)



Note: These figures provide the rational benchmark investment rates for the experimental decision problems. Panel A assumes the rational benchmark for all A values. Panel B shows the benchmark when the focal asset dominates with a positive Δ, and Panel C shows the benchmark when the focal asset dominates with a negative Δ. The vertical axis represents the invest rate.

C Additional Tables

Table A.1: Validation Tests

Dep. var.:	(1) Annual std	(2) Browse Freq	(3) Browse Length	(4) Holding
Pos Price Chg	-0.000	0.031	-0.801	-2441.111
	(0.000)	(0.083)	(1.043)	(3871.391)
Bandwidth	0.011 (opt)	0.011 (opt)	0.011 (opt)	0.011 (opt)
Polynomial	Y	Y	Y	Y
Individual FE	Y	Y	Y	Y
Product FE	Y	Y	Y	Y
Year Month FE	Y	Y	Y	Y
Adj. R^2	0.734	0.449	0.296	0.381
N	234509	235195	235195	235195

Note: This table verifies that there is no jump in the property of the asset or the investors' browsing behavior prior to the potential trading date around the cutoff of the running variable, namely the price change of the asset within a single day. *Annual std* is the annual standard deviation of the asset price. *Browse Freq* and *Browse Length* are the browsing frequency and length (in seconds) of the asset one week prior to the potential trading date. *Holding* is the mean holding of the investor in that month. *Pos Price Chg* is a dummy variable that equals one if the change in price since the previous day is greater than or equal to 0. Investor, product, year, and month fixed effects are controlled in all specifications. Standard errors are clustered at the investor level. * p<0.1, ** p<0.05, *** p<0.01.

Table A.2: Buy Trade: Robustness Tests

Panel A: Restricting to Dates with a Buy Trade

Dep. var.: Buy	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pos Price Chg	-0.012**	-0.010**	-0.010***	-0.007	-0.023***	-0.009*	-0.016***
	(0.006)	(0.005)	(0.003)	(0.005)	(0.003)	(0.005)	(0.003)
Bandwidth	0.011 (opt.)	0.015	0.5	0.002	0.5	0.008	0.5
Polynomial	Y	Y	Y	N	N	linear	linear
Individual FE	Y	Y	Y	Y	Y	Y	Y
Product FE	Y	Y	Y	Y	Y	Y	Y
Year Month FE	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.156	0.155	0.158	0.170	0.158	0.153	0.158
N	63386	77781	169777	19070	169777	49610	169777

Panel B: Asset Returns as the Running Variable

Dep. var.: Buy	(1)	(2)	(3)	(4)	(5)	(6)
Pos Price Chg	-0.004**	-0.003**	-0.004***	-0.006***	-0.003**	-0.002**
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
Bandwidth	0.006 (opt.)	0.007	0.001	0.05	0.004	0.05
Polynomial	Y	Y	N	N	Linear	Linear
Individual FE	Y	Y	Y	Y	Y	Y
Product FE	Y	Y	Y	Y	Y	Y
Year Month FE	Y	Y	Y	Y	Y	Y
Adj. R^2	0.115	0.117	0.125	0.114	0.114	0.114
N	225589	249583	75240	601454	171607	601454

Note: This table presents the robustness tests of the relationship between the likelihood of an investor purchasing an asset and the price change of that asset within a single day. In Panel A, the sample is restricted to dates on which the investor has made a purchase. In panel B, asset price return relative to the previous-day browsed price is used as the running variable. *Buy* is a dummy variable that equals one if the asset is bought and zero if it is not. *Pos Price Chg* is a dummy variable that equals one if the change in price since the previous day is greater than or equal to zero. Investor, product, year, and month fixed effects are controlled in all specifications. Standard errors are clustered at the investor level. * p<0.1, ** p<0.05, *** p<0.01.

Table A.3: Assets Investors Less Interested in

Panel A. Effect of one-day positive price change on investors who viewed two days but not one day before the trade date

Dep. var.: Buy	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pos Price Chg	-0.001	-0.001	0.001	-0.001	-0.002**	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Bandwidth	0.011 (opt.)	0.015	0.5	0.002	0.5	0.008	0.5
Polynomial	Y	Y	Y	N	N	linear	linear
Individual FE	Y	Y	Y	Y	Y	Y	Y
Product FE	Y	Y	Y	Y	Y	Y	Y
Year Month FE	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.078	0.077	0.075	0.100	0.075	0.076	0.075
N	163584	194657	379305	60751	379305	132977	379305

Panel B. Effect of one-day price change on investors who viewed on previous day vs. not on previous day

Dep. var.: Buy	(1) d1 vs. d2	(2) d1 vs. d2d3	(3) d1 vs. d2-d7	(4) d2Only	(5) d2d3Only	(6) d2-d7Only
Pos Price Chg	-0.001	0.000	0.001**	-0.001	0.000	0.001**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Pos Price Chg × View D-1	-0.003*	-0.003***	-0.004***			
	(0.001)	(0.001)	(0.001)			
Bandwidth	0.011 (opt.)	0.011 (opt.)	0.011 (opt.)	0.011 (opt.)	0.011 (opt.)	0.011 (opt.)
Polynomial	Y	Y	Y	Y	Y	Y
Individual FE	Y	Y	Y	Y	Y	Y
Product FE	Y	Y	Y	Y	Y	Y
Year Month FE	Y	Y	Y	Y	Y	Y
Adj. R^2	0.100	0.098	0.091	0.078	0.083	0.085
N	404103	515369	860257	163584	275550	620502

Note: This table presents the relationship between the likelihood of an investor purchasing an asset and the price change of that asset within a single day. The sample includes investors who viewed the asset page at least once during the previous three days. *Buy* is a dummy variable that equals one if the asset is bought and zero otherwise. *Pos Price Chg* is a dummy variable that equals one if the change in price since the previous day is greater than or equal to zero. *View D-1* is a dummy variable that equals one if the investor viewed the asset page on the previous day and zero otherwise. Investor, product, year, and month fixed effects are controlled in all specifications. Standard errors are clustered at the investor level. * p<0.1, ** p<0.05, *** p<0.01.

Table A.4: Restricting to Dates with a Sell Trade

Dep. var.: Sell	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pos. Cum. Perf.	0.0021***	0.0022***	0.0028***	0.0025***	0.0025***	0.0031***	0.0032***
	(0.0008)	(0.0008)	(0.0007)	(0.0007)	(0.0007)	(0.0008)	(0.0005)
Bandwidth	0.813 (opt.)	0.954 (opt.)	10	0.538 (opt.)	10	0.142 (opt.)	10
Polynomial	Order 2	Order 2	Order 2	Order 1	Order 1	N	N
Individual FE	Y	Y	Y	Y	Y	Y	Y
Product FE	Y	Y	Y	Y	Y	Y	Y
Year Month FE	Y	Y	Y	Y	Y	Y	Y
Adj. R^2	0.056	0.056	0.058	0.054	0.058	0.046	0.058
N	494,916	515,000	743,512	443,136	743,512	285,460	743,512

Note: This table presents the relationship between the likelihood of an investor selling an asset and the cumulative gain of that asset. The sample is restricted to dates on which the investor has sold at least one other position in her portfolio. *Sell* is a dummy variable that equals one if the asset is sold and zero otherwise. *Pos. Cum. Perf.* is a dummy variable that equals one if cumulative gain is greater than or equal to zero. Investor, product, year, and month fixed effects are controlled in all specifications. Standard errors are clustered at the investor level. * p<0.1, ** p<0.05, *** p<0.01.

Table A.5: Future Returns Around the Zero Crossing

	Raw 30d Return		Risk-Adj 30d Return	
	(1)	(2)	(3)	(4)
Pos Price Chg	−0.0000	−0.0016	−0.0018	0.0007
	(0.0010)	(0.0020)	(0.0014)	(0.0013)
Price Change		0.1445		−0.2273
		(0.2156)		(0.1355)
Product FE	Y	Y	Y	Y
Year-Month FE	Y	Y	Y	Y
Adj. R^2	0.359	0.359	0.168	0.168
N	115,336	115,336	112,167	112,167

Note: This table tests whether the zero-crossing event predicts future fund performance. The unit of observation is product × date, restricted to observations within the optimal bandwidth of 0.014 around zero one-day price change. The dependent variable is the 22-trading-day forward return, computed either as the raw cumulative return (columns 1–2) or as the risk-adjusted return based on a two-factor model (columns 3–4). *Pos Price Chg* is a dummy variable that equals one if the one-day price change is non-negative. *Price Change* is the magnitude of the one-day price change (the running variable in the main RD specifications). Columns 1 and 3 include only the sign indicator; columns 2 and 4 additionally control for the magnitude. Product and year-month fixed effects are included. Standard errors are double-clustered by product and year-month. * p<0.1, ** p<0.05, *** p<0.01.