Campaign Progress Disclosure in Crowdfunding: Information Design and Counterfactual Evidence (Preliminary and Incomplete)

Keiichi KAWAI and Akira MATSUSHITA

November 26, 2025

1 Institutional Background

Campaign Structure and Investment Environment: Japan's regulatory framework requires that equity crowdfunding campaigns be conducted exclusively through licensed intermediaries. Our analysis is based on administrative access–log data from *Fundinno*, the dominant equity crowdfunding (ECF) platform in Japan in terms of both the number of campaigns hosted and the total amount of capital raised.

As many other ECF platforms, Fundinno operates under an *all–or–nothing rule*: a campaign succeeds only if its *funding target* is reached before the investment phase closes. Each campaign also specifies a *funding limit*, and once this limit is reached no further investments are accepted.

Investors receive equity shares, but these positions are highly illiquid and typically pay off only in the event of a future exit (e.g., IPO or acquisition). Because such exits are uncertain and often far in the future, forming precise valuations based solely on fundamentals at the time of investment is practically difficult. As a result, observable signals—such as a campaign's traction or perceived attention—can play an important role in shaping investors' expectations.

Preview and Investment Phases: Fundinno requires issuing firms to pass an internal screening process before they are permitted to launch a campaign on the platform. This review examines business feasibility, compliance with regulatory requirements, and the minimum documentation needed for investor protection. Only companies that pass this screening may post a campaign, and all campaigns must follow a standardized disclosure format when preparing their detail pages.

Once approved, each campaign follows a two–stage structure consisting of a *preview phase* and an *investment phase*. This structure determines when users can acquire information and when they may take investment actions, and it is central to how we interpret return timing and investment behavior in our model.

The preview phase typically lasts one to two weeks. During this period, users may freely access the campaign's detail page—which includes the full disclosure materials, the funding target and limit, and the scheduled opening time of the investment phase—but cannot invest. Users who view the detail page during this period are referred to as *previewers*; they are precisely the users who can form private assessments of the campaign before investment becomes possible.

The investment phase then opens for a short window, usually 48 hours.[1] Investments are accepted on a first–come basis until the funding limit is reached, or until the investment window closes, whichever comes first. Reaching the funding target marks the campaign as "successful" under the all–or–nothing rule, but investments may continue until the limit is exhausted.

Because the investable window is short, previewers must decide whether and,when to return to the campaign once the investment phase opens. The timing of these revisits is recorded in the platform's access logs and provides the main behavioral variation we exploit in our identification strategy.

Information Environment—Top Page and Detail Page: Users first encounter campaigns on Fundinno's top page, which presents a standardized campaign "snippet" in a grid layout. Each snippet displays the campaign title, industry category, a short description, the funding target and limit, and—during the investment phase—a real–time progress bar showing the fraction of funds raised.

To inspect the campaign more closely, users must click through to the campaign's detail page. The detail page contains the full disclosure package required by the platform's screening process, including a business summary, financial statements, founder information, the intended use of funds, and the firm's stated exit plan. This content follows a standardized format and remains fixed throughout the campaign. The detail page is accessible during both the preview and investment phases, and once the investment phase opens, it displays real–time funding progress at the bottom of the page.

Top–page snippets and detail pages therefore serve distinct roles. Both display real–time traction once investment becomes possible, but only the detail page provides substantive information about the firm, and only users who access the detail page during the preview phase can accumulate campaign-specific information before investment becomes possible. These users—whom we refer to as *previewers*—constitute the population of interest in our behavioral model because they are the only users for whom a return decision is well defined.

In our setting, a key practical consideration is how users can observe a campaign's funding progress. During the period covered by our data, issuing firms are typically small and do not systematically disseminate fundraising updates through external channels such as news outlets or social media, nor does the platform distribute progress information to third–party aggregators. Consequently, users can learn about a campaign's funding progress only by accessing Fundinno's top page or the campaign's detail page.

A crucial implication of this design is that users observe only the current aggregated level of progress at the moment they visit. The identity of investors, the number of visitors, and any information about concurrent monitoring are never disclosed, and users cannot observe the path or evolution of progress unless they visit repeatedly.

User Accounts and Logged-In Behavior: Participation on Fundinno requires users to complete a mandatory identity-verification process (KYC) before they become eligible to invest in any campaign. Throughout the paper, we use the term *users* to refer exclusively to these verified and registered investors, who constitute the population for which the platform records complete, user-level interaction histories.

Although portions of the top page and the detail page can be browsed without logging in, our analysis necessarily focuses on behaviors observed in verified sessions, because only logged-in activity is tracked at the user level.

[1] Historically, the investment phase on Fundinno lasted 48 hours, and the vast majority of campaigns in our dataset closed within this window. More recently, some campaigns have adopted longer investment periods, but these changes occur mostly after the time span covered by our data.

We do not expect this restriction to materially affect the measurement of the behaviors relevant for our model. First, any user who makes an investment must eventually log in, and all investment-related actions—including viewing the investment interface, initiating a commitment, or canceling one—are only available in logged-in sessions. Second, for users who ultimately make a commitment, the administrative logs display long, uninterrupted sequences of page views with no behavioral patterns suggestive of deliberate log-outs, indicating that nearly all meaningful interactions leading up to investment occur while logged in. Third, in the small number of cases in which users appear to switch devices or browsers, the platform records a consistent session footprint that allows us to reconstruct user identities; such cases represent only a minor fraction of our sample.

Table 1: User Groups and Terminology

A. Groups Used in Estimation	
`previewers`	Users who accessed the campaign's detail page at least once during the preview phase.
`early-returners`	`previewers` who, once the investment phase opens, return to the platform (top page or detail page) within the early–return window (e.g., 5 minutes). Used to construct the early–return rate p_j^e.
`non-returners`	`previewers` who return only after a substantial delay (e.g., 48 hours) or never return. Used to construct the non–return rate p_j^∞.
`immediate-investors`	A subset of `early-returners` who invest immediately upon their first return. These users form the numerator of the immediate–investment rate I_j.
B. Groups Used in Counterfactual / Simulation Analysis	
`preview-visitors`	Users who accessed any Fundinno page during the preview phase but *did not* reach the campaign's detail page. Used to measure the preview–conversion flow (snippet → detail).
`post-snippet-visitors`	Users who did not access the detail page during the preview phase but were in a state to observe real–time progress during the investment phase (via the top page). Relevant for simulating how disclosure rules affect entry into previewing.
`preview-investors`	Users who became `previewers` and later made an investment in campaign j.
`non-preview-investors`	Users who did *not* access the detail page during the preview phase but invested during the investment phase.

User Classification and Behavioral Categories: For each campaign, we classify users according to their preview activity and their subsequent behavior once the investment phase opens, following the taxonomy in Table 1. Our estimation relies on four groups defined within the previewing population. Users who access the campaign's detail page during the preview phase are `previewers`; this is the population for which a return decision is well defined. Within this group, we distinguish `early-returners`—those who revisit the platform shortly after the investment phase opens—from `non-returners`, who return only after a substantial delay (e.g., more than 48 hours) or never return at all. A subset of early returners invest immediately upon their first return; we refer to these users as `immediate-investors`, who form the numerator of the immediate-investment rate used in estimation.

In addition to these estimation groups, several user types are relevant only for counterfactual and simulation analyses. `preview-visitors` are users who were exposed to the campaign snippet during the preview phase but did not reach the detail page, and thus determine the preview–conversion flow. `post-snippet-visitors` did not access the detail page during the preview phase but were in a position to observe real–time funding progress during the investment period (via the top page). We also distinguish `preview-investors` and `non-preview-investors`, which categorize investors according to whether they previewed the campaign before investing. These groups do not enter the estimation of structural parameters but are essential for counterfactuals that vary progress visibility or the preview environment.

Administrative Data Structure: Our analysis relies on two complementary data sources maintained by Fundinno. First, we observe user–level administrative access logs, which record every verified interaction with the platform at *one–second resolution*. For each user–campaign pair, the logs track page types (top page or detail page), timestamps of all visits, return times after the investment phase opens, investment attempts, and cancellations. These logs allow us to construct the behavioral measures used in our estimation—previewing behavior, early returns, non–returns, and immediate investments.[(2)]

Second, we observe a campaign–level dataset containing the full set of attributes disclosed by issuing firms and required by regulation. These include the campaign's business area (industry classification), headquarters location, funding target and funding limit, the scheduled opening and closing times of the preview phase, the scheduled opening time of the investment phase, and the minimum investable unit required by regulation. The minimum unit is typically around 100,000 JPY, and individual investors may invest up to 500,000 JPY per campaign, reflecting the regulatory cap on investment amounts for non-professional investors in equity crowdfunding. These campaign characteristics enter the structural model as observable primitives that shape investors' beliefs and participation incentives.

Together, the user-level logs and campaign-level attributes provide a comprehensive view of investor behavior and the economic environment faced by each campaign.

Link to Modeling Choices: In principle, one could model investment behavior in terms of a latent "fundamental quality" of the firm and assume that investors form noisy signals about it. In the setting of ECF, however, we find such a representation less compelling. Campaigns on Fundinno are launched by very young firms whose future cash flows are highly uncertain, and the equity purchased through the platform is extremely illiquid, paying off only in the event of a distant and uncertain exit. Disclosure materials therefore provide only a limited basis for valuing intrinsic firm quality.

What previewers *can* evaluate is something more immediate: whether the business appears capable of attracting and sustaining the attention, visibility, and engagement that are essential for eventual commercial viability. For this reason, we find that it is more natural—and more consistent with the institutional setting—to think of each campaign in terms of two elements:

1. *attention potential* — the campaign's underlying ability to attract and sustain attention, both on the platform in the short run and in the market more broadly in the long run, as inferred by previewers from disclosure materials; and

2. *an attention requirement* — a campaign–specific minimum level of such traction that investors consider necessary for the business to be viable enough to justify an equity commitment.

[(2)] Fundinno's logs contain device identifiers and session metadata that allow us to link multiple access events to the same user account when a user log in in the middle of the session. These cases represent a very small share of the data, and session continuity is observed for the overwhelming majority of users.

Previewers are likely to form subjective impressions of a campaign's attention potential during the preview phase. These impressions are not based on any observed traction—none is observable before investment opens—but instead reflect users' interpretations of the firm's materials, narrative, and presentation. Once the investment window opens, these preview-phase impressions are therefore likely to shape behavior along two conceptually distinct margins: *when* a user chooses to return, and *whether* she ultimately chooses to invest.

First, the timing of a previewer's return is most naturally understood as depending only on her preview-phase impression of attention potential. Before revisiting the platform, users cannot observe progress, the behavior of others, or any signals about how the campaign is performing. Thus, the only basis for deciding whether to check early is the strength of the impression formed during the preview phase. We interpret that any campaign attributes that could plausibly influence both return timing and the investment decision can be interpreted as already incorporated into this private impression. Under this interpretation, attention potential captures the part of the user's evaluation that is relevant for the urgency of returning, whereas the attention requirement does not contribute independently to this timing margin.

Second, once users do return, they may update their preview-phase impressions based on whatever public information is available at that moment—most notably, the funding progress displayed upon their arrival. Such updating need not occur for all users: those who return very early have little new information to incorporate, whereas later returners may substantially revise their initial assessments depending on the campaign progress. After any such updating, users naturally ask whether the campaign appears capable of generating the level of attention they view as necessary for long-run viability. This is where the attention requirement enters: it governs the extensive margin of participation by determining whether the (possibly updated) impression is sufficient to justify an equity commitment. Users whose impressions fall short may continue to watch the campaign or decide not to invest, whereas those who regard the requirement as met have little reason to delay.

Finally, many previewers return very early in the investment window, often within the first few minutes. Although some progress may have begun to accumulate by that time—especially for highly attractive campaigns—early returners observe only a static snapshot and typically do not stay on the page long enough, nor return frequently enough, for their behavior to be plausibly interpreted as systematic monitoring of others' actions. Consistent with this, the data show that even for campaigns with rapid early progress, a substantial share of very-early returners do not invest.

Given this institutional context and the observed patterns in the data, it is natural to interpret early investment behavior by previewers as being based almost entirely on previewers' private impressions formed during the preview phase. At such early moments, the amount of public information available is extremely limited, making it unlikely that users condition their decisions on others' behavior. Thus, in our interpretation, early returners who invest are doing so primarily on the basis of their private assessment of the campaign's attention potential, while early returners who do not invest are those whose private impressions fall short of their own attention requirement. Under this view, early returns reflect the strength of the initial impression, whereas investment requires that the (possibly updated) assessment exceed the attention requirement. Early returns are therefore better understood as a natural response to strong preview-phase impressions rather than an attempt to learn from others.

Taken together, these observations suggest that return timing and investment decisions are governed by distinct components of the previewer's evaluation: the strength of her preview-phase impression and the campaign's attention requirement, possibly combined with the limited public information available at the moment of return. In the next section, we formalize a model that captures exactly these behavioral elements and shows how the two margins—returning and investing—jointly reveal a campaign's underlying attention potential and its attention requirement.

2 Model of Investor Behavior

Based on the interpretation developed in the previous section, we construct a model that captures how previewers form private impressions during the preview phase, how these impressions determine the timing of their return once investment becomes possible, and how they subsequently decide whether to invest after observing the limited public information available at the moment of return.

2.1 Setup

Campaign Attention Structure: Each campaign j is characterized by two latent parameters: its *attention potential* θ_j and its *attention requirement* $\overline{\theta}_j$. As argued previously, we adopt this attention-based representation—rather than an unobserved technological "quality"—because previewers can form meaningful impressions of how much attention a campaign is likely to attract, whereas they have limited ability to infer its long-run fundamentals.

The attention potential θ_j represents the campaign's inherent tendency to attract and sustain attention. Previewers do not observe θ_j directly; instead, they form subjective and noisy impressions of it when they read the disclosure materials during the preview phase. In contrast, the attention requirement $\overline{\theta}_j$ represents the minimum level of traction that investors view as necessary for the business to be viable enough to justify an equity commitment. This viability benchmark is campaign-specific and reflects features such as the funding target, business area, and project scale. These two latent parameters govern distinct behavioral margins: the attention potential shapes return timing, whereas the attention requirement governs the investment decision once the user has revisited the page.

Signal and Belief Formation: When a previewer i views campaign j's detail page during the preview phase, she receives a noisy private signal of the campaign's attention potential:

$$\hat{\theta}_{ij} = \theta_j + \varepsilon_{ij}, \qquad \varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2_{\varepsilon,j}). \tag{1}$$

The idiosyncratic noise term ε_{ij} captures investor-specific variation in how the campaign's presentation is interpreted. A small $\sigma^2_{\varepsilon,j}$ corresponds to a clearer or more uniformly interpretable campaign, while larger dispersion reflects greater divergence in previewers' impressions.

This private impression forms the basis for subsequent behavior during the investment phase: it determines both the previewer's inclination to return and the posterior belief she will update once additional public information becomes available.

Return Behavior: We model this return-timing decision formally. The intended return time t_{ij} is

$$\log t_{ij} = -a\hat{\theta}_{ij} + b + \nu_{ij}, \tag{2}$$

where $a > 0$ and b are common parameters across campaigns, and ν_{ij} captures user-specific responsiveness or promptness. Each user i has a latent responsiveness factor $u_i \sim \mathcal{N}(0, \sigma^2_u)$, and the subset of users who preview campaign j may differ from the overall population due to targeting or selection. Hence, within campaign j we approximate $\nu_{ij} \mid j \sim \mathcal{N}(0, \sigma^2_{\nu,j})$, where $\sigma^2_{\nu,j}$ reflects heterogeneity in responsiveness among users who previewed that campaign. Empirically, we restrict attention to each user's most recent three campaigns, assuming that u_i remains stable over this short horizon. We assume that ν_{ij} is independent of both ε_{ij} and θ_j.

Crucially, the attention requirement $\overline{\theta}_j$ plays no independent role in determining return timing. Any campaign attribute that could plausibly affect the urgency with which a previewer chooses to return—such as perceived difficulty of the funding target or concerns about eventual viability—is already embedded in her preview-phase impression of θ_j. Return behavior therefore depends solely on beliefs about θ_j, whereas $\overline{\theta}_j$ affects only the investment decision after the user has returned.

Furthermore, any investor-specific interpretation of the campaign's qualitative presentation is already absorbed into the private signal $\hat{\theta}_{ij}$. For this reason, the attention requirement $\overline{\theta}_j$ is naturally treated as a campaign-specific but common benchmark shared by all investors.

Investing Behavior: When a previewer revisits the campaign page during the investment phase, she may observe additional public information—most notably the campaign's current funding progress. Let s_{ij} denote the information available to user i at the moment of her return, consisting of her private signal $\hat{\theta}_{ij}$ and whatever public information is visible upon arrival. Based on s_{ij}, she forms a posterior expectation about the campaign's attention potential and invests if

$$\mathbb{E}\left[\, \theta_j \mid s_{ij} \,\right] \geq \overline{\theta}_j. \tag{3}$$

The amount of public information contained in s_{ij} varies with return timing. Early returners observe essentially no progress and therefore rely almost entirely on their preview-phase private impressions, whereas later returners may revise their beliefs more substantially if progress has accumulated by the time they arrive. Thus, the investment margin is shaped by how the user's (possibly updated) posterior compares to the attention requirement, while the return-timing margin is governed entirely by the strength of her initial impression.

Normalization of Latent Attention Potential θ_j. To make campaigns comparable, we normalize the latent attention potential θ_j to follow a standard normal distribution:

$$\theta_j \sim \mathcal{N}(0, 1). \tag{4}$$

This normalization does not imply that "campaign quality" is normally distributed. Rather, θ_j is a latent index of attention potential, aggregating many small factors—such as the business idea's appeal, thematic relevance, or alignment with current investor interests—that jointly determine how strongly a campaign tends to attract attention.

If these components contribute additively and independently, their sum is approximately normal by the central limit theorem. The normalization simply fixes the scale of this latent index, allowing us to compare campaigns on a common metric and to interpret θ_j as a standardized measure of attention- drawing ability.

Summary of Model Parameters: To summarize, our model is fully specified by the following set of parameters:

$$\{\theta_j, \overline{\theta}_j, \sigma_{\varepsilon,j}, \sigma_{\nu,j}\}_j \quad \text{and} \quad \{a, b\}.$$

For each campaign j, θ_j denotes the latent *attention potential*—how strongly the campaign tends to attract investor interest once its details are viewed—while $\overline{\theta}_j$ represents the common benchmark level of attention required for the campaign to be perceived as economically viable. The variance $\sigma^2_{\varepsilon,j}$ captures dispersion in investors' private interpretation of the campaign's appeal, and $\sigma^2_{\nu,j}$ measures heterogeneity in responsiveness among previewers who have viewed the campaign. The common parameters a and b govern how differences in θ_j translate into return timing. In the next section, we show how variation in observable user behavior within and across campaigns can be used to recover these parameters. Together, these parameters determine how previewers' initial private assessments, responsiveness, and observable campaign features jointly generate variation in return and investment behavior across campaigns. In the next section, we discuss how variation in observable user behavior within and across campaigns can be exploited to recover these parameters from the data.

2.2 Assumptions for Interpreting the Data

We now clarify the assumptions that govern how the access-log data are interpreted within the framework of the model. Before doing so, we summarize the terminology used throughout the analysis, as these observable categories map directly onto the behavioral margins of the model.

We use the terminology summarized in Table 1. *Previewers* are users who viewed the campaign's detail page during the preview phase. Among them, *early-returners* are those who return to the platform (either the top page or the campaign detail page) within the early–return window (five minutes after the investment phase opens). A subset of these early-returners who invest immediately upon that first return constitute the *immediate-investors*. Previewers who do not return within forty–eight hours, or never return, are classified as *non-returners*.

These groups correspond to the two behavioral margins in the model: (i) the return-timing margin, captured by the distinction between early-returners and non-returners, and (ii) the investment margin, captured by immediate-investors versus those who return but do not invest.

With this terminology in place, we now describe the assumptions that govern how investors' latent decisions translate into the observable access-log data. These assumptions specify how timing behavior and early investments are reflected in the data, and how measurement noise is handled. We distinguish three layers that jointly define the data-generating process consistent with the model: (i) observation and measurement within a campaign, (ii) behavioral interpretation of early investments, and (iii) stability and noise cancellation across campaigns.

(i) Within-Campaign Observation and Logging Noise: The access logs record platform visits with fine but not perfectly precise timestamps. Minor server- or device-level delays may introduce small random variation, but our empirical analysis relies only on coarse distinctions in return timing (*early-returners vs. non-returners*, as defined in Table 1), which are robust to such noise.

Visits occurring long after the investment phase has opened (beyond the 48-hour window defined earlier) are far more likely to reflect browsing of newly opened campaigns rather than continued monitoring of the same campaign. These late visits are therefore excluded from the analysis.

Assumption 1 *Timestamp noise may be present, but it does not alter the classification of previewers into early-returners (those who revisit within five minutes of launch) and non-returners (those who do not return within the subsequent forty-eight hours).*

(ii) Interpretation of Early Investments: Under our interpretation, the decision to return to a campaign shortly after the investment phase opens is governed solely by the strength of a previewer's initial private impression of its attention potential θ_j. This return decision is therefore based on a different and substantially lower behavioral threshold than the decision to commit capital.

Given this separation of motives, it is natural to treat the investment choices of early-returners as being driven entirely by their private signals. Previewers who return almost immediately have had no meaningful opportunity to update their beliefs using public progress, and the fact that they return so quickly already indicates that their continued monitoring is not conditioned on observing others' actions. Thus, among early-returners, those who invest immediately upon their first return are interpreted as doing so because their private signal $\hat{\theta}_{ij}$ exceeds the common viability benchmark $\overline{\theta}_j$, while those whose private signal falls short simply do not invest.

Assumption 2 *Among early-returners, a previewer becomes an immediate-investor if and only if her private signal about the campaign's attention potential exceeds the attention requirement $\overline{\theta}_j$. Because essentially no public progress has accumulated within the early-return window, this investment decision reflects only the previewer's private signal, without influence from others' actions.*

(iii) Across–Campaign Stability and Noise Cancellation: Many users preview multiple campaigns. We assume that latency or interface delays in the access logs are primarily user–specific rather than campaign–specific, so that the measurement noise affecting recorded return times remains approximately constant for each user across campaigns within a short time window. Under this interpretation, differences in return timing across campaigns can be attributed to behavioral variation rather than logging artifacts.

Empirically, each user previews only a small number of campaigns within any short horizon. We therefore restrict attention to a user's most recent three campaigns, treating this window as short enough for the user–specific responsiveness component u_i to remain effectively constant. Within this window, variation in observed return times across campaigns can be interpreted as driven by campaign–level heterogeneity—rather than by changes in individual responsiveness—which is the key source of identifying variation for θ_j and $\sigma_{\nu,j}$.

Assumption 3 *Measurement noise in recorded return times is user–specific and cancels out across campaigns. Users' responsiveness is stable across their most recent three campaigns, allowing within–user variation in return timing to identify campaign–level differences.*

Together, these assumptions define how the access-log data map onto the parameters of the model. We now define the empirical measures that summarize user behavior within and across campaigns. These observables serve as the basis for estimating the model parameters:

$$\{\theta_j, \overline{\theta}_j, \sigma_{\varepsilon,j}, \sigma_{\nu,j}\}_j \quad \text{and} \quad \{a, b\}.$$

2.3 Observed Variables

The assumptions above clarify how the access-log data can be interpreted in terms of users' underlying behavior. We now describe the empirical quantities that can be constructed from these logs and that will serve as the basis for identifying the model's structural parameters.

Observed Return Behavior: For each campaign j, we observe the fractions of its previewers who are classified as (i) *early returners* (those who revisit within t_j^e) and (ii) *non–returners* (those who do not revisit before t_j^l). We denote these empirical shares by $\widehat{p}_j^e$ and $\widehat{p}_j^\infty$.

Under Assumption 1, timestamp noise does not change classification, so these empirical shares coincide with the model-implied probabilities:

$$\widehat{p}_j^e = p_j^e \equiv \Pr(\text{early-returner of } j \mid \text{previewer of } j),$$

$$\widehat{p}_j^\infty = p_j^\infty \equiv \Pr(\text{non-returner of } j \mid \text{previewer of } j).$$

We set the cutoffs to $t_j^e = 5$ minutes and $t_j^l = 48$ hours after opening.

Observed Early Investing Behavior: Among early-returners of campaign j, we observe the fraction who invest immediately upon returning. Let the empirical share be $\widehat{I}_j$. Formally,

$$\widehat{I}_j = \frac{\#\{\text{immediate-investors of } j\}}{\#\{\text{early-returners of } j\}}.$$

Under Assumption 2, these immediate investments occur before previewers can observe others' actions. Therefore, $\widehat{I}_j$ coincides with the model-implied probability

$$\begin{aligned} I_j &\equiv \Pr(\text{immediate-investor of } j \mid \text{early-returner of } j) \\ &= \frac{\Pr(\text{immediate-investor of } j \mid \text{previewer of } j)}{\Pr(\text{early-returner of } j \mid \text{previewer of } j)}. \end{aligned}$$

Cross-Campaign Variation in Return Timing. For users who preview multiple campaigns, we construct the empirical variance of within-user differences in log return times: $\widehat{\delta}_{j,k} \equiv \mathrm{Var}_i\left[\log t_{ij} - \log t_{ik}\right]$ for $k = j+1, j+2$. Under Assumption 3 the user-specific intended return time noise ν_{ij} cancels out across campaigns:

$$\log t_{ij} - \log t_{ik} = -a(\theta_j - \theta_k) - a(\varepsilon_{ij} - \varepsilon_{ik}).$$

Hence, the model-implied moment is

$$\delta_{j,k} \equiv \mathrm{Var}_i[\log t_{ij} - \log t_{ik}] = a^2(\sigma_{\varepsilon,j}^2 + \sigma_{\varepsilon,k}^2).$$

We interpret the empirical quantity $\widehat{\delta}_{j,k}$ as the sample analogue of the model-implied moment $\delta_{j,k}$. Because each $\delta_{j,k}$ loads additively on the pair $(\sigma_{\varepsilon,j}^2, \sigma_{\varepsilon,k}^2)$, the full collection $\{\widehat{\delta}_{j,j+1}, \widehat{\delta}_{j,j+2}\}_j$ supplies a system of linear moment conditions that uniquely identifies the set of noise variances $\{\sigma_{\varepsilon,j}^2\}_j$ (and equivalently the correlation parameters $\{\rho_j\}_j$).

Summary of Observables. Together, these quantities constitute the empirical observables used in the identification analysis:

$$\mathcal{O} = \{\, p_j^e,\ p_j^\infty,\ I_j,\ \widehat{\delta}_{j,j+1},\ \widehat{\delta}_{j,j+2}\}_j.$$

3 Mapping Observables to Model Parameters: Identification

The observable return and investment patterns described in the previous section encode information about all components of the structural model. Before detailing the mapping, we state the main identification result.

Proposition 1 *All structural parameters $\{\theta_j,\ \bar{\theta}_j,\ \sigma_{\varepsilon,j},\ \sigma_{v,j}\}_{j=1}^{J}$ and the global parameters $\{a, b\}$ are point identified from the empirical observables $\mathcal{O} = \{p_j^e,\ p_j^\infty,\ I_j,\ \delta_{j,j+1},\ \delta_{j,j+2}\}_j$ implied by the access-log data.*

To make the connection between the observables and the model primitives transparent, it is useful to reexpress the campaign-specific structural parameters in terms of four composite objects that map directly into the data:

$$\tau_j \equiv \sqrt{a^2\sigma_{\varepsilon,j}^2 + \sigma_{v,j}^2}, \qquad \eta_j \equiv -a\theta_j + b, \qquad \bar{h}_j \equiv \frac{\bar{\theta}_j - \theta_j}{\sigma_{\varepsilon,j}}, \qquad \rho_j \equiv \frac{-a\sigma_{\varepsilon,j}}{\tau_j}. \tag{5}$$

These parameters are jointly invertible with respect to the original set of parameters $\{\theta_j,\ \bar{\theta}_j,\ \sigma_{\varepsilon,j},\ \sigma_{v,j}\}$. The reparameterization is convenient because (η_j, τ_j) summarize the location and scale of the return–time distribution, while $(\bar{h}_j, \rho_j)$ characterize the normalized investment threshold and the correlation between early-return information and investment decisions.

With this reparameterization, the mapping from observables to parameters can be described in four steps. First, within-campaign return observables (p_j^e, p_j^∞) identify (η_j, τ_j). Second, because θ_j is standardized, the cross-campaign distribution of η_j identifies (a, b). Third, cross-campaign variation $\delta_{j,k}$ isolates the noise component and thus identifies ρ_j (or $\sigma_{\varepsilon,j}$). Finally, early-investment behavior I_j recovers the normalized investment threshold $\bar{h}_j$ conditional on (p_j^e, τ_j, ρ_j).

(i) Identifying (η_j, τ_j) from Return Behavior: We begin by showing how within-campaign return behavior identifies the two parameters (η_j, τ_j) that govern the location and scale of the latent log return-time distribution. These parameters can be recovered directly from the early-return probability p_j^e and the no-return probability p_j^∞, because under Assumption 1 these probabilities correspond exactly to the lower and upper tails of that distribution, implying a one-to-one mapping to (η_j, τ_j).

More specifically, we can write the intended return time as $\log t_{ij} = \eta_j + \tau_j \widetilde{w}_{ij}$, where $\widetilde{w}_{ij} \equiv \frac{-a\varepsilon_{ij} + v_{ij}}{\tau_j} \sim \mathcal{N}(0,1),\ \widetilde{w}_{ij} \perp \theta_j$. Therefore,

$$p_j^e = \Pr\left(\log t_{ij} \le \log t_j^e\right) = \Phi\left(\frac{\log t_j^e - \eta_j}{\tau_j}\right) \tag{6}$$

and

$$p_j^\infty = \Pr\left(\log t_{ij} > \log t_j^l\right) = 1 - \Phi\left(\frac{\log t_j^l - \eta_j}{\tau_j}\right).$$

Lemma 1 *The within-campaign return parameters (η_j, τ_j) are uniquely identified from the early-return and no-return probabilities (p_j^e, p_j^∞):*

$$\tau_j = \frac{\log t_j^l - \log t_j^e}{\Phi^{-1}(1 - p_j^\infty) - \Phi^{-1}(p_j^e)}, \text{ and } \eta_j = -\frac{\left(\log t_j^l - \log t_j^e\right)\Phi^{-1}\left(p_j^e\right)}{\Phi^{-1}\left(1 - p_j^\infty\right) - \Phi^{-1}\left(p_j^e\right)} + \log t_j^e. \tag{7}$$

(ii) Identifying (a, b) from Variation in Expected Return Timing η_j: Having pinned down the campaign-specific return parameters (η_j, τ_j), we can now turn to the global parameters (a, b). Bacause $\theta_j \sim \mathcal{N}(0,1)$ and $\eta_j \equiv -a\,\theta_j + b$ with $a > 0$, the cross-campaign variation in η_j provide us with:

$$a = \sqrt{\mathrm{Var}_j[\eta_j]} \qquad \text{and} \qquad b = \mathbb{E}_j[\eta_j]. \tag{8}$$

(iii) Identifying $\sigma_{\varepsilon,j}$ from Cross-campaign Variation in Return Timing $\{\delta_{j,k}\}_{k=j+1,j+2}$: Having identified the global scaling parameter a from the cross-campaign variation in η_j, we now treat a as known and proceed to identify the noise component $\sigma_{\varepsilon,j}$ (and therefore the correlation parameter $\rho_j = -a\sigma_{\varepsilon,j}/\tau_j$). Our approach exploits the cross-campaign variation in return timing.

For any user i who previews both campaigns j and $k \in \{j+1, j+2\}$, Assumption 3 implies $\nu_{ij} = \nu_{ik}$, and therefore

$$\log t_{ij} - \log t_{ik} = -a(\theta_j - \theta_k) - a(\varepsilon_{ij} - \varepsilon_{ik}).$$

Taking variances across users yields the observable moment

$$\delta_{j,k} \equiv \text{Var}_i[\log t_{ij} - \log t_{ik}] = a^2(\sigma^2_{\varepsilon,j} + \sigma^2_{\varepsilon,k}), \qquad k = j+1, j+2. \tag{9}$$

Thus, for each campaign j, we observe two linear moment conditions involving $\sigma^2_{\varepsilon,j}$ and its neighbors.

To express the system compactly, let J denote the number of campaigns (indexed $1, \ldots, J$). Define a $(2J-3) \times J$ matrix $\mathbf{A}$ whose $i_{(j,k)}$-th row (corresponding to the pair $k = j+1$ or $j+2$) has ones in columns j and k, and zeros elsewhere. Let $\mathbf{s}$ be the J-vector of noise variances with i-th entry $\sigma^2_{\varepsilon,i}$ and define the $(2J-3)$-dimensional vector $\mathbf{d}$ whose $i_{(j,k)}$-th entry is $\delta_{j,k}$. Then the moment conditions in (9) can be written compactly as $\mathbf{As} = \mathbf{d}/a^2$.[(3)]

While it appears that we have more moment conditions than the number of parameters $\{\sigma^2_{\varepsilon,j}\}_j$, the system nevertheless has a unique solution. For any $\mathbf{x} \neq \mathbf{0}$,

$$\mathbf{x}^\top \mathbf{A}^\top \mathbf{A}\mathbf{x} = \|\mathbf{A}\mathbf{x}\|^2 = \sum_{(j,k):\, k=j+1,j+2} (x_j + x_k)^2 > 0,$$

so $\mathbf{A}^\top\mathbf{A}$ is positive definite and hence nonsingular. It follows that $\text{rank}(\mathbf{A}) = J$ and the unique solution is

$$\mathbf{s} = (\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{d}/a^2.$$

The corresponding closed-form solution is summarized below:

Lemma 2 *For each campaign j, $\sigma_{\varepsilon,j}$ is uniquely identified because the system of equations* (9) *admits a unique solution $\{\sigma^2_{\varepsilon,j}\}$, given by*

$$\sigma^2_{\varepsilon,j} = \frac{\delta_{j,k} + \delta_{j,l} - \delta_{k,l}}{2a^2}, \qquad k = j+1,\ l = j+2. \tag{10}$$

(iv) Identifying the Normalized Investment Threshold $\bar{h}_j$ from Early-Investment Behavior: Having recovered the noise component $\sigma_{\varepsilon,j}$—and therefore the correlation parameter $\rho_j \equiv -a\sigma_{\varepsilon,j}/\tau_j$—we now turn to the normalized investment threshold $\bar{h}_j \equiv (\bar{\theta}_j - \theta_j)/\sigma_{\varepsilon,j}$, which determines how investors translate early-return information into early-investment decisions. This parameter is identified from the observed early-investment frequency within each campaign.

[(3)] For example, when $J = 4$, $\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$, $\mathbf{s} = (\sigma^2_{\varepsilon,1}, \ldots, \sigma^2_{\varepsilon,4})^\top$, and $\mathbf{d} = (\delta_{1,2}, \delta_{1,3}, \delta_{2,3}, \delta_{2,4}, \delta_{3,4})^\top$.

A previewer invests in campaign j upon returning if and only if $\widetilde{\varepsilon}_{ij} \equiv \frac{\varepsilon_{ij}}{\sigma_{\varepsilon,j}} \geq \bar{h}_j$, and returns early if and only if $\widetilde{w}_{ij} \equiv \frac{-a\varepsilon_{ij}+\nu_{ij}}{\tau_j} \leq \Phi^{-1}(p_j^e)$ as identified in (6). Therefore, the observed early-investment frequency among early returners is

$$I_j = \frac{\Pr(\widetilde{\varepsilon}_{ij} \geq \bar{h}_j,\ \widetilde{w}_{ij} \leq \Phi^{-1}(p_j^e))}{p_j^e} = 1 - \frac{\Pr\left(\widetilde{\varepsilon}_{ij} \leq \bar{h}_j, \widetilde{w}_{ij} \leq \Phi^{-1}\left(p_j^e\right)\right)}{p_j^e}$$

Furthermore, $(\tilde{\varepsilon}_{ij}, \tilde{w}_{ij})$ is jointly standard normal with correlation ρ_j because both $\tilde{\varepsilon}_{ij}$ and $\tilde{w}_{ij}$ are standard normal, their correlation equals their covariance:

$$\operatorname{corr}(\tilde{\varepsilon}_{ij}, \tilde{w}_{ij}) = \operatorname{Cov}\left(\frac{\varepsilon_{ij}}{\sigma_{\varepsilon,j}}, \frac{-a\varepsilon_{ij} + \nu_{ij}}{\tau_j}\right) = \frac{-a\sigma_{\varepsilon,j}}{\tau_j} = \rho_j.$$

That is,

$$1 - \frac{\Pr\left(\widetilde{\varepsilon}_{ij} \leq \overline{h}_j, \widetilde{w}_{ij} \leq \Phi^{-1}\left(p_j^e\right)\right)}{p_j^e} = I\left(\overline{h}_j; \rho_j\right) \equiv 1 - \frac{\Phi_2\left(\overline{h}_j, \Phi^{-1}\left(p_j^e\right); \rho_j\right)}{p_j^e},$$

where $\Phi_2(x, y; \rho) \equiv \int_{-\infty}^{x} \Phi\left(\frac{y-\rho t}{\sqrt{1-\rho^2}}\right) \phi(t)\, dt$ is the bivariate standard normal distribution function with correlation ρ. Since $I(\cdot)$ is strictly increasing in $\overline{h}_j$, we obtain the follwoing lemma.

Lemma 3 *For each campaign j, once the correlation parameter ρ_j is uniquely identified, the normalized investment threshold $\overline{h}_j$ is uniquely identified as the unique solution h to the equation $I_j = I\left(h; \rho_j\right)$.*

4 Data and Methodology

Table 2: Summary Statistics for Previewer Behavior

	All	Successful	Unsuccessful	Difference
# previewers	634.380	683.205	519.226	163.979***
	(207.461)	(208.605)	(152.259)	
Share early-returners p_j^e	0.088	0.100	0.060	0.040***
	(0.069)	(0.075)	(0.039)	
Share non-returners p_j^∞	0.382	0.372	0.405	-0.033***
	(0.076)	(0.074)	(0.078)	
Share immediate-investors I_j	0.338	0.394	0.208	0.186***
	(0.203)	(0.200)	(0.143)	

Notes: Each cell reports the campaign-level mean, with standard deviation in parentheses. The last column reports the difference in means between successful and unsuccessful campaigns (Successful − Unsuccessful). Stars indicate significance levels from a Welch two-sample t-test: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

4.1 Data

We utilize the following datasets provided by Fundinno:

1. **`campaign_df` (527 campaigns)**: For each campaign, this dataset includes the company name, `company_id`, `campaign_id`, funding target, funding limit, start time of the preview period, start time of the investing period, end time of the investing period, and the prefecture of the company's headquarters.(4) *(Complete dataset, no missing rows.)*

2. **`user_df` (44,512 registered users)**(5): For each registered user (i.e., those eligible to invest), this dataset includes `user_id`, prefecture of residence, and registration and approval dates.(6) *(Complete dataset, no missing rows.)*

3. **`orders_df`**: For each pair of `campaign_id` and `user_id` where an investment-related action (e.g., investment, cancellation, or application to be waitlisted) is recorded, this dataset includes the timestamp of the action, the type of activity, the pledged amount, and the canceled amount. *(Complete dataset, no missing rows.)*

4. **`access_log_df` (40,123,306 rows)**: For each campaign, during the preview or investment phase, this dataset records the timestamp of each activity (e.g., visit to the top page, campaign detail page

(4) The dataset also includes additional details such as the area of business, self-reported stock price, and whether the company holds patents, has achieved profitability, received venture capital support, or its tenure. These attributes are recorded but not explicitly modeled due to the inclusion of campaign fixed effects.

(5) We have excluded all users who are not eligible to invest.

(6) The dataset also includes user-level variables such as year of birth, self-reported income, and investment experience, which are not directly used in this analysis due to the inclusion of user fixed effects.

view, or investment in the campaign), the session_id, the type of activity, and the user_id if the activity was performed by a signed-in user. *(Some rows are missing.)*

While these datasets provide comprehensive coverage of user and campaign activities, certain limitations remain. A primary issue is that users can view campaign detail pages without logging in, resulting in some access log entries lacking a user_id. These entries may correspond to browsing behavior by registered users who did not log in during their sessions, making it challenging to associate these actions with specific users.

However, the session_id enables us to track a user's activity within a single session. If a user logs in during the session—for instance, logging in is required to make an investment—we can associate all actions performed within that session, including those completed prior to logging in (such as browsing campaign detail pages or visiting the site), with the corresponding user_id. Moreover, as making a pledge necessitates logging into a registered account, we can reliably trace the user_id for all investors.

Additionally, as is common with access logs, some data appear to be missing in the following ways:

- **Unrecorded Activities**: 1.27% of the 76,248 unique user-campaign_id pairs with recorded investments in orders_df lack any corresponding activities (e.g., views or investment actions) in access_log_df.
- **Partial Records**: An additional 0.23% of the same user-campaign_id pairs have investment actions logged in access_log_df but no prior viewing behavior.

These 1,104 pairs have been excluded from our analysis, leaving 1,489,064 unique pairs of user_id and campaign_id with complete records of all variables for the estimation.

4.2 Descriptive Statistics

5 Estimation Results

Figure 1 and Table 4 summarize our estimation results. In Figure 1, the horizontal axis shows the predicted number of investors among previewers under the counterfactual scenario in which users rely only on their private signals and cannot observe others' behavior. The vertical axis shows the actual number of investors among previewers in the observed disclosure environment. Points above the 45-degree line therefore indicate campaigns where disclosure led to more investors relative to the counterfactual, while points below the line indicate campaigns where disclosure discouraged investment.

Our estimation results show a systematic pattern: many campaigns that ultimately failed to reach their funding target lie below the 45-degree line, suggesting that these campaigns would have attracted more investors if progress information had not been disclosed. By contrast, campaigns that ultimately reached their target (but not the funding limit) tend to lie above the line, indicating that disclosure amplified participation.

Table 4 provides a summary of these results by reporting the distribution of the ratio of actual investors (vertical axis in Figure 1) to counterfactual investors (horizontal axis). This ratio captures the relative effect of disclosure on participation. For about 75% of failed campaigns, the ratio is below one, confirming that disclosure reduced investment. For successful-but-unfilled campaigns, the median ratio is below but much closer to one, consistent with disclosure encouraging investment.

Table 3: Descriptive Statistics of Campaigns and Users

Campaign Metrics ($N = 527$)	**Mean**	**Std**	**25%**	**50%**	**75%**	**Max**
Funding Target (JPY)	13,887,260.38	7,799,753.71	9,990,000.00	12,500,000.00	16,000,000.00	80,000,000.00
Funding Limit (JPY)	48,394,737.95	21,781,349.10	30,091,000.00	45,000,000.00	60,000,000.00	99,990,000.00
Pledged Amount (JPY)	24,077,864.24	19,451,787.81	9,522,000.00	19,500,000.00	32,300,000.00	99,990,000.00
Success Ratio (%)	194.31	155.07	79.79	150.00	275.63	1,010.11
User Engagement by Campaign ($N = 44,512$)	**Mean**	**Std**	**25%**	**50%**	**75%**	**Max**
Visitors to Campaigns: $\sum_i \texttt{visit}_{i\to j}$	2241.92	1010.20	1488.50	2261.00	2870.50	4859.00
Views of Campaigns: $\sum_i \texttt{view}_{i\to j}$	1090.25	344.16	834.50	1037.00	1302.00	2534.00
Viewers among Visitors: $\sum_i \texttt{visit-view}_{i\to j}$	506.62	180.82	377.50	483.00	600.00	1294.00
Investors per Campaign: $\sum_i \texttt{inv}_{i\to j}$	142.38	117.58	50.00	111.00	198.50	591.00
User Engagement Across Campaigns ($N = 44,512$)	**Mean**	**Std**	**25%**	**50%**	**75%**	**Max**
Visited Campaigns: $\sum_j \texttt{visit}_{i\to j}$	26.54	51.83	4.00	8.00	24.00	520.00
Viewed Campaigns: $\sum_j \texttt{view}_{i\to j}$	12.91	34.80	1.00	3.00	9.00	511.00
Viewed after Visited: $\sum_j \texttt{visit-view}_{i\to j}$	6.00	18.15	0.00	1.00	5.00	461.00
Campaigns Invested In: $\sum_j \texttt{inv}_{i\to j}$	1.69	5.72	0.00	0.00	1.00	277.00
User with At Least One View ($N = 35,561$)	**Mean**	**Std**	**25%**	**50%**	**75%**	**Max**
Visited Campaigns: $\sum_j \texttt{visit}_{i\to j}$	32.09	56.58	5.00	11.00	32.00	520.00
Viewed Campaigns: $\sum_j \texttt{view}_{i\to j}$	16.16	38.26	2.00	4.00	13.00	511.00
Viewed after Visited: $\sum_j \texttt{visit-view}_{i\to j}$	7.51	20.03	1.00	2.00	6.00	461.00
Campaigns Invested In: $\sum_j \texttt{inv}_{i\to j}$	2.11	6.33	0.00	0.00	2.00	277.00
User with At Least One Pledge ($N = 17,758$)	**Mean**	**Std**	**25%**	**50%**	**75%**	**Max**
Visited Campaigns: $\sum_j \texttt{visit}_{i\to j}$ 54.64	71.72	11.00	28.00	66.00	520.00	
Viewed Campaigns: $\sum_j \texttt{view}_{i\to j}$	28.20	50.75	4.00	11.00	28.00	511.00
Viewed after Visited: $\sum_j \texttt{visit-view}_{i\to j}$	12.71	26.80	1.00	5.00	12.00	461.00
Campaigns Invested In: $\sum_j \texttt{inv}_{i\to j}$	4.23	8.45	1.00	2.00	4.00	277.00

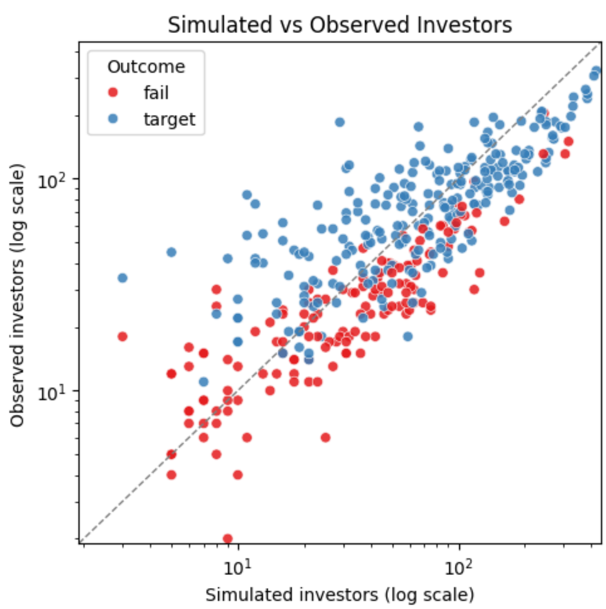

Figure 1: Model Implied Investors and Actual Investors

Table 4: Investors under Full Dislcosure / Investors under No Disclosure

Outcome	Count	Mean	Std. Dev.	25%	50%	75%
Fail	146	0.890	0.678	0.546	0.713	1.000
Target	235	1.335	1.302	0.659	0.907	1.495

References

6 Appendix: Figures

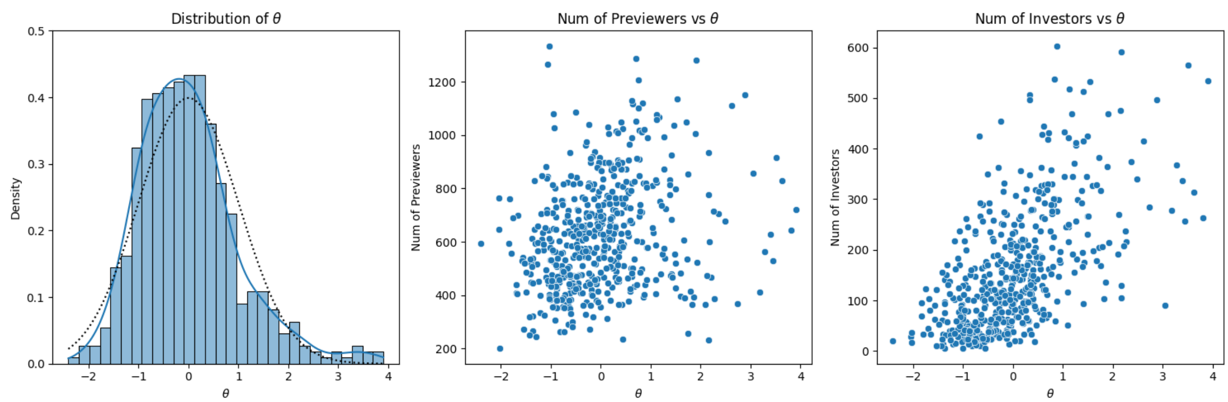

Figure 2: Distribution of Campaign Quality

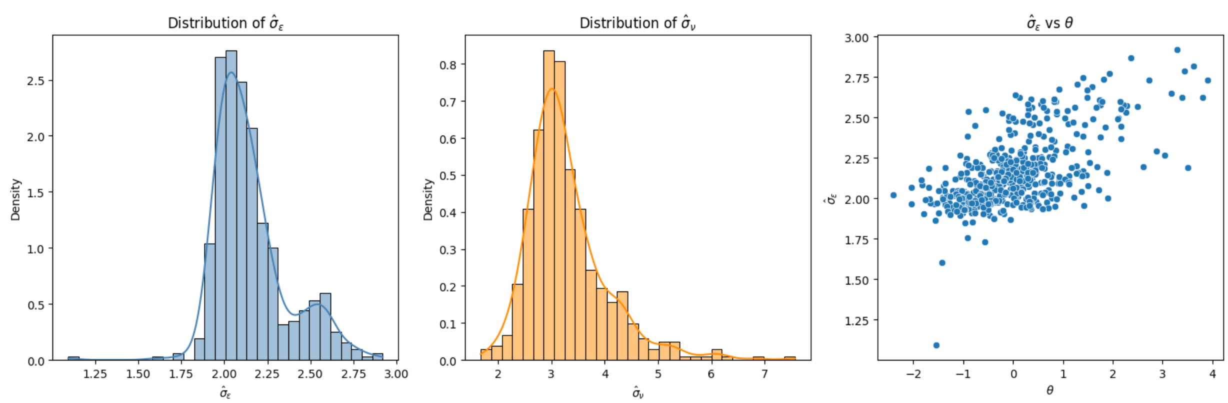

Figure 3: Distribution of Private Signal/Timing Noise $(\sigma_\varepsilon, \sigma_\nu)$

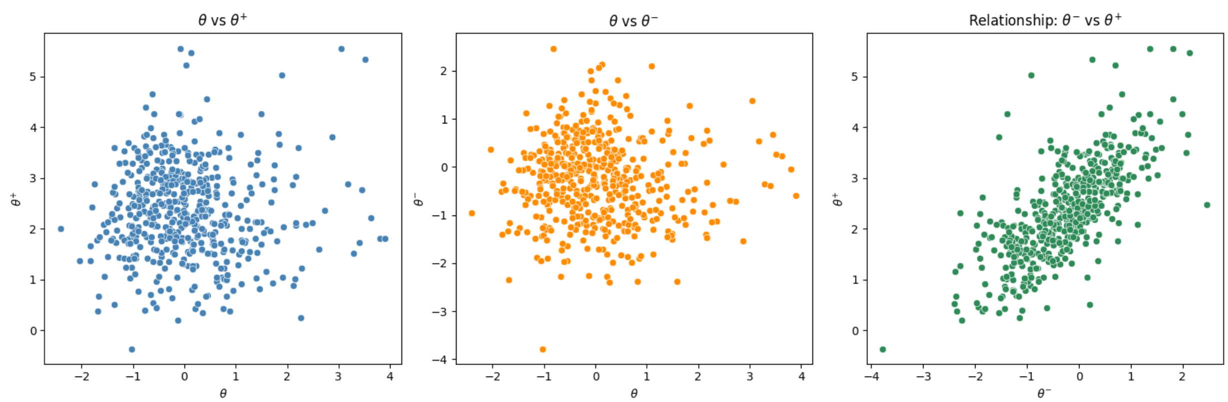

Figure 4: Distribution of Investment/Dropout Thresholds (θ^+, θ^-)