Transparency Pays: Data Privacy on Digital Platform

Shumiao Ouyang Verina F. Que

University of Oxford *Nanyang Technological University*

March 23, 2026

Abstract

Platforms increasingly offer privacy controls, but does being *transparent* about these controls help or hurt the platform? We conduct a large-scale field experiment on one of the largest global platforms, with over 8,000 users, in which the treatment group receives information highlighting centralized data authorization tools, while the control group does not. Using field experiment and difference-in-differences specifications, we find that the treated group exhibits both higher direct engagement with the privacy tools and higher downstream platform engagement and trust. First, direct engagement with the privacy tools substantially increases. Second, downstream platform engagement and trust increase, with page views, active duration, and asset balances all rising, while contextual app-level data consent decisions remain unchanged at users' own privacy preference levels. Third, the direct and downstream effects are driven by distinct user segments: tool usage increases are concentrated among low-knowledge users (closing an awareness gap), while engagement and trust increases are concentrated among intrinsically privacy-sensitive users (consistent with enhanced perceived control). Consumers bear real privacy costs on platforms. Transparency decreases these costs by improving the decision environment through greater perceived control and clarity, while leaving consent-level data sharing with third-party apps unaffected. Transparency pays.

Keywords: digital privacy, transparency, field experiment, platform economics, perceived control

JEL Codes: D83, L86, M31, D91

1 Introduction

Digital platforms collect vast quantities of user data, yet the very regulations designed to curb this collection, such as GDPR, CCPA, and China's Personal Information Protection Law (PIPL), require platforms to offer users privacy controls. A central tension emerges: does being transparent about these controls help or hurt the platform? On one hand, heightened privacy salience could trigger a "chilling effect," priming users to restrict data sharing and disengage (Adjerid et al. (2019); Burtch et al. (2015); Acquisti et al. (2016)). On the other hand, transparency may increase perceived control and build trust (Tucker (2014)). Resolving this tension is of first-order importance: platforms must decide whether to actively communicate about privacy features or keep them quietly available, and regulators need evidence on whether disclosure mandates achieve their intended effects.

This paper provides causal evidence on whether transparency about privacy controls pays off for platforms. We leverage a large-scale field experiment conducted on Alipay, China's leading digital payment and super-app platform serving over one billion users, to estimate the effects of an information treatment that highlighted centralized data authorization tools. A key institutional feature motivates our design: while users have always been able to cancel data authorizations on an app-by-app basis, navigating to each app's individual settings page, Alipay introduced in 2024 a new centralized management tool that allows users to view and revoke all third-party app authorizations from a single interface. This centralized tool substantially lowers the cost of privacy management relative to the app-by-app alternative, but only for users who know it exists. Our experiment varies only whether users are informed about this tool, holding the tool itself constant across groups. We distributed a survey to 4 million randomly selected active users, assigning them to a treatment group, which received an additional survey question drawing attention to the location and function of these privacy management tools, or a control group receiving only benchmark questions. We received 4,136 responses in the treatment group and 4,138 in the control group, yielding a total sample of 8,274 respondents, and linked them to granular panel data on users' privacy actions, app-level consent decisions, and platform engagement over a 21-week window surrounding the intervention.

Our identification strategy combines this experimental randomization with a difference-in-differences (DiD) framework. The randomization ensures that treated and control groups are ex ante identical, as confirmed by balance tests across all survey dimensions, while the DiD design allows us to separate the treatment effect from common time trends by exploiting within-individual variation in outcomes before and after the information shock. Standard errors are clustered at the individual level throughout.

We document three sets of findings. *First*, the information treatment has strong direct effects on privacy tool usage. Treated users are 2.2 percentage points more likely to visit the data authorization list page (a 186% increase relative to the 1.18% pre-treatment baseline) and 1.8 percentage points more likely to actively cancel an app's data authorization (a 49% increase over baseline). Notably, the treatment does not increase visits to the general "settings" page, indicating that the effect is specific to the highlighted tools rather than reflecting a broad increase in platform exploration.

Second, the treated group exhibits higher downstream platform engagement and trust on average, while contextual data-sharing decisions remain unchanged. We find null effects on the number of third-party apps users authorize and on app-level consent rates, suggesting that data authorization decisions are made contextually at the point of app use and are not sensitive to platform-level privacy information. At the same time, treated users exhibit significantly higher page views (2.4%, $p < 0.05$), longer active duration (2.7%, $p < 0.05$), and larger asset balances held on the platform (6.4%, $p < 0.10$), the last being a revealed-preference measure of trust in a financial platform. These are average treatment effects. Heterogeneity analysis reveals that these engagement and trust gains are concentrated among users who self-identify as intrinsically privacy-sensitive: the treatment effect on page views, active duration, and asset balances is large and significant for high-intrinsic users, while the effect for non-intrinsic users is near zero. This pattern suggests that the aggregate engagement response is not diffuse but driven by users for whom perceived control over personal data is most salient.

Third, heterogeneity analysis suggests distinct patterns behind the direct and downstream effects. The engagement and trust increases are concentrated among users who self-identify as intrinsically privacy-sensitive (the "high-intrinsic" group); the triple-interaction terms for page views, duration, and asset balances are all significant and large. Notably, high-intrinsic users do not visit the privacy tools more than low-intrinsic users, a pattern consistent with the perceived-control channel identified by Tucker (2014): knowing the tools exist, rather than actively using them, may be what matters for trust. Conversely, the direct effects on tool visitation and authorization cancellation are concentrated among users with low baseline platform knowledge, consistent with the information treatment closing an awareness gap. This knowledge-based heterogeneity does not extend to downstream engagement, suggesting that the engagement increase operates through perceived control rather than through actual tool use.

This paper makes three contributions. First, we complement and replicate in a major new platform setting the foundational finding of Tucker (2014), that perceived control over personal data enhances user engagement, a decade later and in a high-stakes financial super-app

context. Tucker (2014) demonstrates that Facebook's introduction of enhanced privacy controls increased the click-through rate of personalized ads, consistent with perceived control reducing psychological resistance to data use. Our design holds the privacy controls constant and varies only transparency about their existence, providing causal evidence that the information channel alone is sufficient to generate engagement and trust gains; the treated group's engagement gains are not mediated by differential tool use, suggesting that knowing the option exists improves the decision environment. Complementing work on centralizing ex-ante consent (Farronato et al., 2025), we study a different channel: centralized *ex-post* management, the ability to review and revoke past authorizations from a single interface, which improves the perceived control of platform engagement even when rarely exercised.

Second, we establish that the information treatment does not produce a chilling effect on contextual data-sharing decisions. This finding speaks to a practical fear among platforms that highlighting privacy features may reduce the data flows sustaining their third-party app ecosystems. Our null result on app-level consent rates shows that these decisions remain at users' own privacy preference levels, driven by the immediate utility of the app, and are not revised in response to platform-level privacy information. This is consistent with the contextual nature of privacy decisions documented in the literature (Adjerid et al., 2019; Nissenbaum, 2010; Que, 2026). The implication is that consumers bear real privacy costs on platforms, and firms should not assume these costs disappear simply because users continue to share data. Our results suggest that transparency about data reversibility can reduce these costs, not by changing the consent-level tradeoff, but by improving the decision environment at the platform level, while leaving contextual data flows intact.

Third, we decompose the treatment effect along two dimensions of user heterogeneity, intrinsic privacy preferences and platform knowledge, to identify heterogeneous responses to the information treatment. The finding that engagement effects are concentrated among intrinsically privacy-sensitive users, while direct tool usage is concentrated among low-knowledge users, suggests that the cost of obscure privacy tools is not uniform: it scales with users' intrinsic privacy sensitivity θ_i. This connects to the heterogeneous privacy preferences documented by Lin (2022) and has practical implications for targeted communication strategies.

The remainder of the paper proceeds as follows. Section 2 lays out a simple conceptual framework that organizes our empirical analysis. Section 3 positions the paper within the relevant literatures. Section 4 describes the institutional setting, experimental design, data, and econometric specification. Section 5 presents the main results and heterogeneity analysis. Section 6 discusses implications and concludes.

2 Conceptual Framework

We develop a simple model of consumer privacy management on a digital platform to organize the intuition behind our empirical analysis.[1] The model highlights a distinction between two layers of privacy management, a *central layer* operated by the platform and a *consent-level layer* at which users make app-by-app consent decisions, and shows how transparency at the central layer can improve the decision environment without altering consent-level choices.

Setting. Consider a platform hosting J third-party apps. User i faces privacy decisions at two layers. At the *consent-level layer*, the user encounters app j and makes a binary authorization decision $d_{ij} \in \{0, 1\}$: accept ($d_{ij} = 1$) to grant data access and receive the service, or reject ($d_{ij} = 0$) to forgo it. These decisions are contextual, made at the point of app use, driven by the immediate tradeoff between the app's service value v_{ij} and the user's privacy cost c_i, which reflects both intrinsic discomfort with data sharing and instrumental concerns about how the specific app will use the data. At the *central layer*, the platform provides tools that govern how users manage their data across all apps. In our setting, the platform operates a contextual consent model at the consent-level layer, where users authorize apps one by one, a natural interface for a super-app with heterogeneous services. The central layer complements this with centralized *ex-post management*: a single interface where users can review all past authorizations and revoke any of them. The key insight is that this tool reduces the stakes of each consent-level decision by making consent reversible. A user who knows she can revoke any past authorization does not need to deliberate as carefully at the point of app use, because she can try the service first and adjust later. This is analogous to a return policy: it does not change the information available at the point of sale, but it lowers the cost of buying because the buyer can reverse the transaction. Data reversibility transforms data consent from a perceived one-way door into a two-way door.

The model. We formalize this intuition with a utility framework at each layer.

At the *consent-level layer*, user i encountering app j obtains:

$$u_i(d_{ij}) = \begin{cases} v_{ij} - c_i + \varepsilon^1_{ij}, & \text{if } d_{ij} = 1 \\ \varepsilon^0_{ij}, & \text{if } d_{ij} = 0 \end{cases} \tag{1}$$

where v_{ij} is the service value, c_i is the privacy cost, which encompasses both intrinsic discomfort with data sharing and instrumental concerns about how the app will use the data, and $\varepsilon^1_{ij}, \varepsilon^0_{ij}$ are i.i.d. taste shocks. The acceptance probability is $\Pr(d_{ij} = 1) = F(v_{ij} - c_i)$, where $F(\cdot)$ is the CDF of $\varepsilon^0_{ij} - \varepsilon^1_{ij}$. These decisions are contextual: they depend on the

[1] We do not structurally estimate this model. The analysis is reduced-form; the model serves to clarify what our information treatment affects and how it maps onto user decisions at different layers of the platform.

app-specific tradeoff $v_{ij} - c_i$ and are made at the point of use.

At the *central layer*, user i chooses a level of overall platform engagement $e_i \geq 0$ to maximize:

$$U_i(e_i) = b(e_i) - \psi_i \cdot e_i \tag{2}$$

where $b(\cdot)$ is increasing and concave, and the marginal engagement cost is:

$$\psi_i = \bar{\psi} + \theta_i \cdot (1 - r_i) \tag{3}$$

Here $\bar{\psi}$ is a baseline friction, $\theta_i \geq 0$ is user i's intrinsic privacy sensitivity, and $r_i \in [0, 1]$ is the user's perceived ease of revoking past data authorizations. The term $\theta_i \cdot (1 - r_i)$ captures the privacy friction from perceived irreversibility: when the user believes past consent cannot easily be undone (r_i low), deeper engagement feels costly because each interaction generates data exposure perceived as permanent. When the user knows she can easily reverse any past authorization (r_i high), this friction diminishes, and the perceived downside risk of trying new services is lower, because a regretted decision can be corrected.

How transparency affects the model. Our experiment shifts r_i by being transparent about the centralized revocation tool, informing treated users of its existence and location. Define r_i^0 as baseline perceived reversibility and $r_i^1 > r_i^0$ as perceived reversibility after treatment. The first-order condition $b'(e_i^*) = \psi_i$ yields, by the implicit function theorem:

$$\frac{\partial e_i^*}{\partial r_i} = \frac{\theta_i}{-b''(e_i^*)} > 0 \tag{4}$$

The treatment effect on engagement is:

$$\Delta e_i^* = \frac{\theta_i \cdot (r_i^1 - r_i^0)}{-b''(e_i^*)} \tag{5}$$

Two features of this expression clarify what transparency does. First, the engagement response is proportional to intrinsic privacy sensitivity θ_i: transparency reduces intrinsic costs most for users who care most about privacy, because perceived control matters more to them. Second, the response is proportional to the awareness gap $(r_i^1 - r_i^0)$: transparency is more impactful for users who did not previously know that easy revocation was available.

Critically, the treatment shifts perceived reversibility r_i but does not change the app-specific value v_{ij} or privacy cost c_i. Learning that past authorizations can be revoked does not alter the immediate service value of any app, nor does it change the user's intrinsic or instrumental concerns about sharing data with that app at the moment of encounter.

Therefore:

$$\frac{\partial \Pr(d_{ij} = 1)}{\partial r_i} = 0 \tag{6}$$

This is the key distinction that separates the data-reversibility mechanism from an alternative in which the treatment shifts privacy costs. If the treatment lowered c_i, say by reassuring users that their data is well-protected and thereby reducing both intrinsic and instrumental concerns, we would expect both engagement to rise *and* app-level consent rates to increase. Instead, consent rates are unchanged: transparency operates through perceived reversibility of past decisions, not through the perceived cost of new ones.

Model insight on empirical evidence. Complementing work on centralizing ex-ante consent (Farronato et al., 2025), the model clarifies how transparency about ex-post management generates engagement and trust gains through a distinct channel: reducing the intrinsic privacy costs at the central layer by increasing perceived control, rather than simplifying the choice architecture at the consent-level layer.

The broader insight is that consumers bear real privacy costs on platforms, and firms should not assume these costs are negligible simply because users continue to share data. Counterintuitively, transparency about data reversibility *decreases* these costs. It does not change the consent-level tradeoff, since app-level consent is invariant to r_i and contextual data flows are unaffected, but it improves the decision environment at the central layer. Users who know they can reverse past consent face lower intrinsic friction, perceive greater control over their data, and consequently engage more deeply and trust the platform more. The punchline is that transparency pays: it is incentive-compatible with the platform because it builds trust without disrupting the data flows that sustain the third-party ecosystem. Section 5 provides the empirical evidence.

3 Related Literature

This paper contributes to several interconnected literatures spanning the economics of privacy, platform design, information provision, and behavioral responses to salience. We discuss each in turn, highlighting how our setting and findings extend existing knowledge.

I. Privacy Economics and Perceived Control. Our paper relates most directly to the literature on the economic consequences of privacy controls and disclosures on digital platforms. As Goldfarb and Que (2023) survey, the economics of digital privacy centers on the tension between the benefits of data flows (innovation, personalization, and matching) and the costs borne by consumers, both intrinsic (disutility from being observed) and instrumen-

tal (risks of discrimination, manipulation, or breach). A foundational finding in this area is that perceived control over personal data can enhance, rather than diminish, the effectiveness of data-dependent services. Tucker (2014) demonstrates that Facebook's introduction of enhanced privacy controls increased the click-through rate of personalized ads, consistent with the idea that perceived control reduces psychological resistance to data use. We complement and replicate this finding in a new major-platform setting a decade later: our design holds the privacy controls constant and varies only transparency about their existence, providing causal evidence that the information channel alone is sufficient to generate engagement and trust gains.

The empirical evidence on whether privacy salience helps or harms platforms is mixed. Burtch et al. (2015) find that privacy controls on a crowdfunding platform have both positive (comfort) and negative (priming) effects, with the net effect depending on the timing of disclosure. Adjerid et al. (2016) show that privacy policy reminders affect consent for health information exchanges only when combined with financial incentives. Most relevant to our downstream results, Adjerid et al. (2019) demonstrate that choice architecture-induced changes in upstream privacy choices do not cascade into downstream behaviors like browsing or transactions, a finding we extend by showing that platform-level information also fails to alter contextual app-level consent decisions.

Our contribution is to show that, in a high-stakes financial platform setting, the net effect of transparency about privacy features on engagement is unambiguously positive, operating through perceived control among privacy-sensitive users.

The broader economics of privacy literature has examined how privacy regulation affects market outcomes ((Acquisti et al., 2016; Goldfarb and Tucker, 2012; Miller and Tucker, 2009)). Acquisti et al. (2016) provide an influential survey of the behavioral economics of privacy, documenting systematic deviations between stated privacy preferences and revealed behaviors, the "privacy paradox." Goldfarb and Tucker (2011) show that privacy regulation (the EU's e-Privacy Directive) significantly reduced the effectiveness of online advertising. Our paper contributes to this literature by demonstrating that *voluntary* platform communication about privacy features, as opposed to mandatory regulation, can build trust without the adverse effects on data flows that regulatory mandates sometimes produce.

II. Heterogeneous Privacy Preferences. A growing literature recognizes that consumers differ systematically in their privacy preferences, and that this heterogeneity has important implications for platform strategy and welfare analysis. Lin (2022) develops a framework that disentangles privacy preferences into intrinsic and instrumental components, showing that intrinsic preferences, defined as the disutility from data sharing per se inde-

pendent of consequences, account for a substantial share of total privacy valuation. Our heterogeneity analysis maps directly onto this framework: we use a survey-based measure of intrinsic privacy sensitivity (self-assessed relative to peers) and show that downstream engagement effects are concentrated among high-intrinsic users.

Using survey and behavioral data from the same Alipay platform, Chen et al. (2025) document a "data privacy paradox" at the app level: users who express stronger privacy concerns paradoxically derive greater utility from data-intensive services, because the omitted variable driving both concern and usage is demand for digital services. Our findings complement theirs and offer a resolution from the platform design side: privacy-sensitive users respond most positively to information about privacy controls, suggesting that their elevated engagement reflects enhanced perceived control over data that they value sharing, not lower actual concern. When combined with the financial inclusion evidence in Ouyang (2022), who shows on the same platform that the data generated through cashless payment significantly expands credit access, a coherent picture emerges: privacy-sensitive users stand to gain most from data-intensive services, and providing them with credible revocation options unlocks this value by reducing the perceived risk of engagement. This is consistent with theoretical work on the value of choice and control ((Brandimarte et al., 2013)), which shows that even non-exercised options to restrict data sharing can reduce the psychological cost of participation.

More broadly, the literature on heterogeneous responses to privacy interventions underscores the importance of user-level variation. Athey et al. (2017) present a discrete-choice experiment revealing substantial heterogeneity in privacy valuations, with many users willing to share data for small incentives while a tail cares deeply about privacy. Our field experiment complements such survey-based approaches by documenting how heterogeneous preferences translate into differential behavioral responses to a real-world information intervention.

III. Information Provision and Salience. Our paper also contributes to the broader literature on how information provision affects consumer behavior, extending insights from product markets to digital privacy. In the food labeling context, Bollinger et al. (2011) show that mandatory calorie posting at Starbucks reduces average calories per transaction by 6%, attributing the effect to a combination of consumer learning and increased salience. Araya et al. (2022) find that warning labels on cereal packages decrease demand, with larger effects among lower-socioeconomic groups. In contrast, Adalja et al. (2023) demonstrate that mandatory GMO labeling has no incremental effect in markets where voluntary labels already exist, highlighting the importance of the informational baseline.

4 Empirical Strategy

This section describes the institutional setting, experimental design, data construction, and econometric framework. We begin with the platform context that motivates our study, then detail the field experiment, present descriptive evidence supporting our identification, and specify the estimating equations.

4.1 Institutional Background

Alipay, launched in 2004 by Alibaba Group, is one of the world's largest digital platforms, processing vast transaction volumes and serving over one billion users and millions of merchants. Originally a payment service, Alipay has evolved into a "super-app" integrating digital financial services (consumer credit, wealth management, and investment products) alongside its core payment functions (see Ouyang (2022), for evidence on how Alipay's payment infrastructure generates data that expands credit access). Crucially, it also aggregates daily life services through a library of over two million third-party mini-apps, covering transportation, local shopping, government services, and more.

This multi-layered architecture makes Alipay a compelling environment for studying privacy behavior. At the consent-level layer, users routinely grant data authorizations to third-party mini-apps to access specific services. Each mini-app may request access to different types of personal data, such as location, contacts, and payment history, for different purposes, and the user makes a binary consent decision at the point of app use. Que (2026) studies these consent-level authorization decisions and documents that they are shaped by contextual factors: a user's response to one app's data request can spill over to subsequent requests, with the effect depending on the nature of the service and whether the user reviewed the privacy terms. Chen et al. (2025) show that on this platform, users with stronger privacy concerns do not share less data, a "data privacy paradox" explained by the fact that privacy-concerned users also have higher demand for the digital services that require data sharing. These contextual authorization decisions coexist with centralized privacy management tools at the central layer that allow users to review and revoke authorizations across all apps from a single interface. The platform's financial dimension adds a meaningful trust margin: users can hold asset balances (e.g., in money market funds), providing a revealed-preference measure of trust in the platform's data practices that goes beyond stated preferences or engagement metrics.

In 2024, Alipay introduced a new centralized Data Authorization List (DAL) within its Central Privacy Setting page. The DAL allows users to view a comprehensive list of all third-party apps that have been granted data access, and to cancel authorizations directly

from this single interface rather than navigating to each app individually. The DAL was deployed platform-wide, meaning all users had access regardless of experimental assignment. Our experiment varies only whether users are *informed* about this tool.

4.2 The Field Experiment

We conduct a randomized field experiment on the Alipay platform in May 2024. The experiment takes the form of an A/B survey distributed through the message box at the center of Alipay's front page, a highly visible channel, to a random sample of 4 million active users. Participants are randomly assigned to one of two groups:

- **Control group:** Receives a benchmark survey measuring privacy preferences, concerns, data protection principles, intrinsic and instrumental privacy preferences, and platform knowledge (Questions Q1–Q5).
- **Treatment group:** Receives the identical benchmark survey plus an additional question (the "information treatment") that asks whether users know about and have used the data authorization lists in the privacy center, accompanied by a visual highlighting the tools' location.

The survey was distributed on May 14, 2024[2], with responses collected through May 22, 2024. We received 4,136 responses in the treatment group and 4,138 in the control group, yielding a total sample of 8,274 respondents.

Survey Design. The benchmark survey comprises five sections designed to measure users' understanding of and preferences for data sharing, all administered identically to both groups:

- **Q1 (Satisfaction):** "What is your level of satisfaction with Alipay's privacy protection?" (5-point scale)
- **Q2 (Concern):** "How concerned are you about sharing data with Alipay?" (Very/Mildly/Not concerned)
- **Q3 (Data protection principles):** Users tick all protections they value from the following options (multiple selections allowed): minimum collection, data leak prevention, limiting third-party sharing, no price discrimination, ability to retrieve data.

[2]The treatment was distributed on May 14, 2024. We define the pre-treatment period as weeks -7 through -1 and the post-treatment period as weeks $+1$ through $+10$ relative to the treatment week.

- **Q4 (Intrinsic and instrumental preferences):** Measures both intrinsic privacy preferences (Q4a: "Do you consider yourself more privacy-sensitive than people around you?") and instrumental motives for sharing data (Q4b: trust, resignation, no negative experience, service necessity).
- **Q5 (Platform knowledge):** Users indicate whether they "Know" or "Don't know" the location of: the central privacy setting, advertising personalization controls, and data authorization cancellation in each app.

The Information Treatment. The privacy management interface on Alipay has a two-layer structure. The first layer is the **Central Privacy Setting** (referred to as "setting" in the regression tables), the platform's top-level User Protection Center page. This page is general-purpose and broadly known: over 70% of users in both groups report awareness of it in Q5. The second layer, nested within the Central Privacy Setting, is the **Data Authorization List** (referred to as "DAL" in the regression tables), a specific tool that displays all third-party apps to which the user has granted data access. From this DAL page, users can review individual app authorizations and cancel them directly, revoking an app's data access without navigating to the app's own settings. The DAL is the key centralized management tool introduced in 2024; unlike the general setting page, most users are unaware of its existence.

The treatment group receives an additional question after all benchmark questions, asking whether users know about and have used the DAL located within the Central Privacy Setting. Figure 1 shows the treatment screen. The key design feature is that this question simultaneously *informs* users about the DAL's existence and location while appearing as a natural survey question, minimizing demand effects. By directing attention to the second-layer tool (the DAL) nested within the first-layer page (the Central Privacy Setting), the treatment provides information about a specific, low-awareness tool rather than general privacy awareness.

4.3 Data and Descriptive Evidence

We construct a panel dataset at the individual-day level covering March 15 to August 31, 2024, spanning approximately 8 weeks before and 12 weeks after the treatment date.

Outcome Variables. We organize our outcome variables into three categories:

- **Direct privacy actions:** Binary indicators for whether a user ever visits the Data Authorization List (DAL), visits the Central Privacy Setting page, or actively cancels

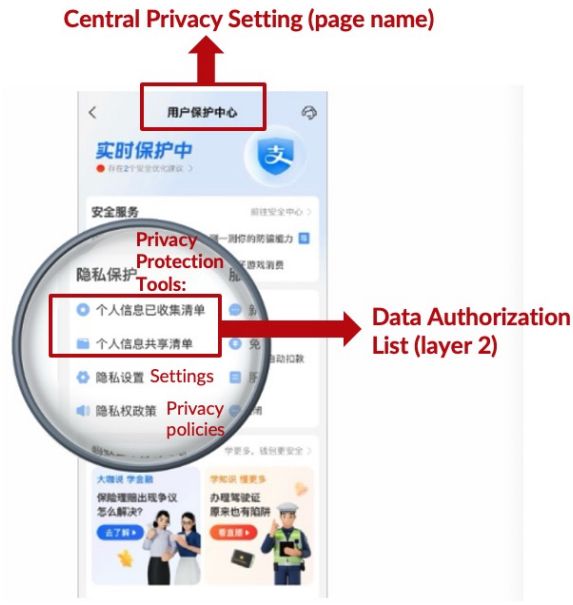


Figure 1: The Information Treatment: Data Authorization List in the Central Privacy Setting

Note: Screenshot of the treatment question shown to users in the treatment group. The top of the page shows the Central Privacy Setting ("setting" in the regression tables). The magnified section highlights the Data Authorization List ("DAL" in the regression tables), the second-layer tool where users can view all third-party app authorizations and cancel them directly. The original interface is in Chinese.

an app's data authorization via the DAL, measured separately for the pre- and post-treatment periods.

- **App-level data consent (weekly):** The number of third-party apps a user authorizes (downloads with data consent) and the ratio of authorized apps to total apps downloaded, aggregated weekly.
- **Platform engagement and trust (weekly):** Log page views, log active duration, and log asset balance held on the platform, aggregated weekly.

Heterogeneity Measures. We define two key moderating variables from the benchmark survey responses, which are measured identically for both treatment and control groups:

- **Intrinsic privacy sensitivity:** A binary indicator equal to 1 if the user responded "Agree" to Q4a ("Do you consider yourself more privacy-sensitive than people around you?").
- **Platform knowledge:** A binary indicator equal to 1 if the user responded "Know" to Q5(2), indicating awareness of advertising personalization controls. We use this

measure rather than the other two Q5 items because awareness of advertising settings is less likely to reflect privacy attitudes per se, and more likely to proxy for how actively a user navigates platform settings—making it a cleaner measure of general platform familiarity.

4.3.1 Sample Characteristics and Balance

Figure 2 presents the distribution of survey responses across all benchmark questions, comparing treatment and control groups. The distributions are visually balanced across all dimensions, confirming the validity of the randomization.

Several patterns merit discussion. In Q1, over half of users in both groups report the highest satisfaction level (level 5) with Alipay's privacy protection. In Q2, approximately two-thirds are "Very" or "Mildly" concerned about sharing data, a concern unsupported by Q4b, where the dominant motives for sharing are "Need to use the service" ($\approx 78\%$) and "Trust the platform" ($\approx 75\%$), consistent with the pragmatic trade-offs documented by Chen et al. (2025). A striking $\approx 73\%$ of users in both groups consider themselves more privacy-sensitive than their peers (Q4a), providing a large high-intrinsic subsample for heterogeneity analysis. Finally, Q5 reveals a gradient in platform knowledge: while over 70% know the general privacy center, knowledge drops to ≈ 66–67% for more specific tools like advertising personalization controls and data authorization cancellation in each app, motivating our investigation of knowledge-based heterogeneity.

Figure 3 complements the survey-based balance check with observed demographic characteristics. The sample is predominantly young working-age adults: the 26–35 age group accounts for the largest share in both treatment (36.7%) and control (36.1%), followed by the 36–45 cohort ($\approx 29\%$). Users aged 18–25 and 46–55 each constitute roughly 13–15%, while those over 56 represent approximately 6–7%. In terms of educational attainment, the largest identifiable group demonstrates strong balance across treatment and control conditions. Gender composition is also well-balanced, with approximately 55% male and 45% female users in both groups.

4.4 Empirical Specification

We use two estimating equations corresponding to the two types of outcome variables.

Direct Effects. For the direct privacy action outcomes, each user contributes exactly one binary observation in the pre-treatment period and one in the post-treatment period. Because there are only two observations per individual, individual fixed effects are not identified;

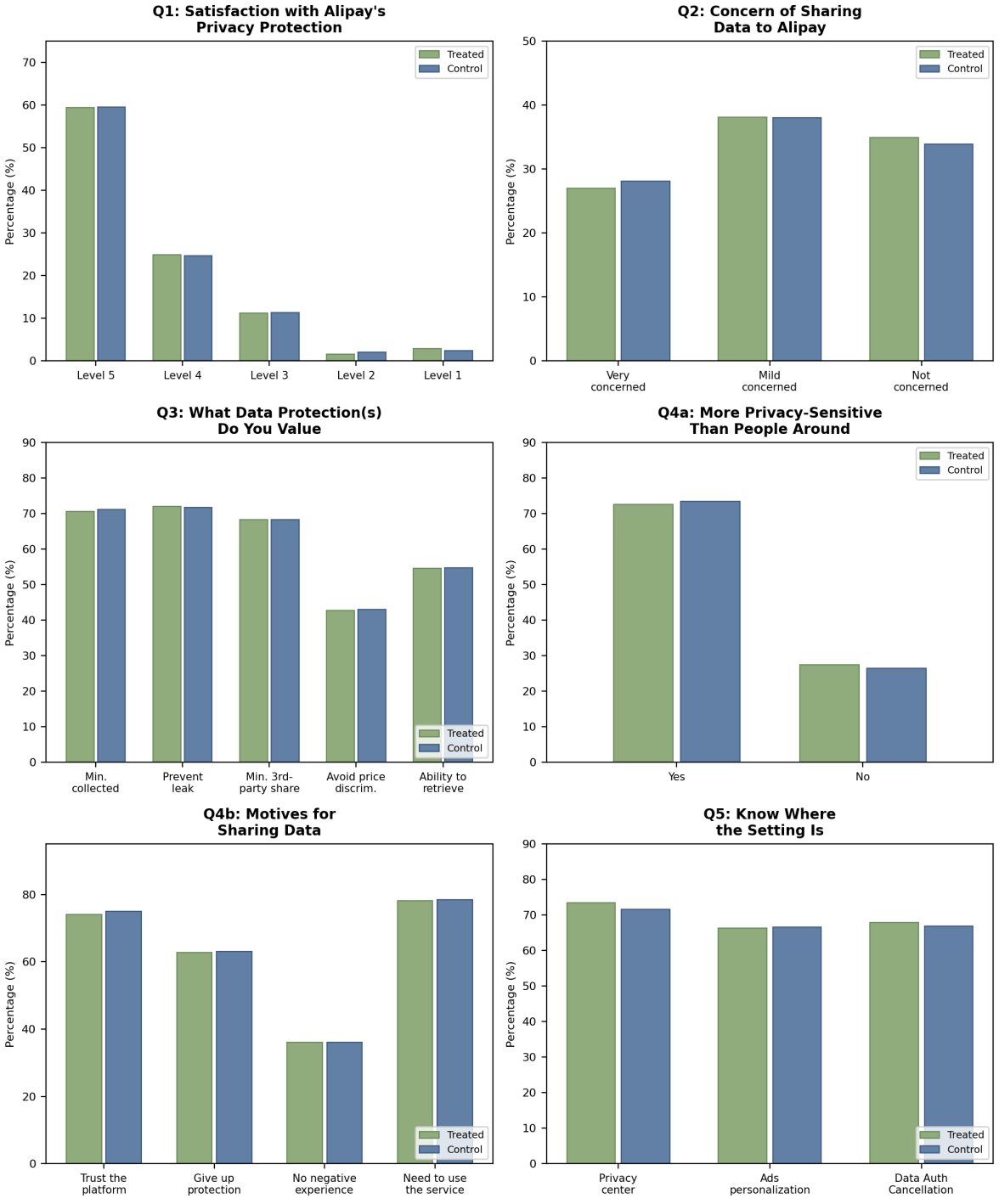


Figure 2: Balance Between Treated and Control Users: Distribution of Survey Responses

Note: Distribution of responses to each benchmark survey question, separately for treated and control users. Questions Q1–Q5 are administered identically to both groups. The close alignment across all dimensions confirms balance from the randomization.

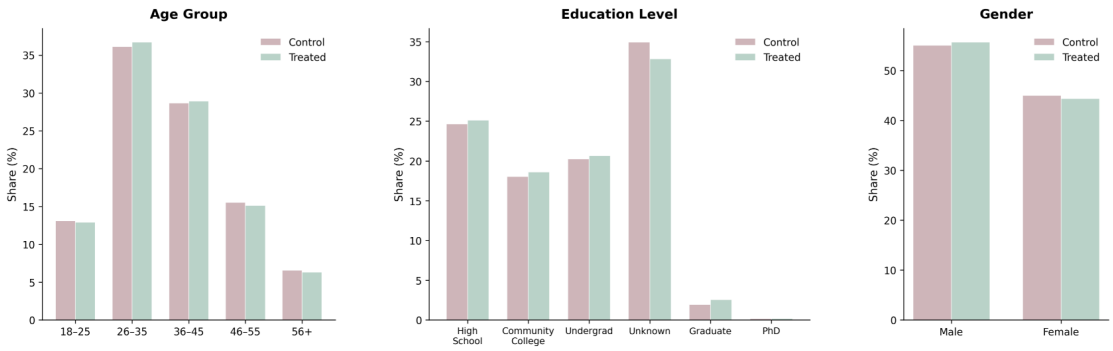


Figure 3: Balance on Demographic Characteristics: Age, Education, and Gender

Note: Demographic composition of the survey sample by treatment group. The close alignment on age, educational, and gender distributions across treatment and control confirms the validity of the randomization.

we instead estimate a cross-sectional difference-in-differences:

$$Y_{id} = \alpha + \beta_1 \cdot \text{POST}_d + \beta_2 \cdot \text{Treat}_i + \delta \cdot \text{POST}_d \cdot \text{Treat}_i + \varepsilon_{id} \tag{7}$$

where Y_{id} is the binary outcome for individual i in period $d \in \{\text{pre}, \text{post}\}$, POST_d is an indicator for the post-treatment period, Treat_i is an indicator for assignment to the treatment group, and ε_{id} is the error term. The coefficient δ is the difference-in-differences estimator, capturing the average treatment effect on the treated.

Downstream Effects: Main Specification. For the downstream outcomes (app-level consent and platform engagement), we observe a weekly panel for each user. To estimate the overall average treatment effect, we use a two-way fixed effects difference-in-differences model:

$$Y_{it} = \alpha_i + \lambda_t + \beta_1 \cdot \text{POST}_t + \beta_2 \cdot \text{Treat}_i + \beta \cdot \text{POST}_t \cdot \text{Treat}_i + \varepsilon_{it} \tag{8}$$

where Y_{it} is the outcome for individual i in week t, α_i denotes individual fixed effects, λ_t denotes week fixed effects, POST_t is an indicator for the post-treatment period (weeks +1 through +12 relative to treatment), Treat_i is an indicator for assignment to the treatment group, and ε_{it} is the error term. Note that POST_t is absorbed by λ_t and Treat_i is absorbed by α_i in estimation. The coefficient β captures the average treatment effect on the treated.

Downstream Effects: Event Study. To examine the dynamics of the treatment effect and validate the parallel trends assumption, we replace the single post-treatment indicator

with a full set of bi-weekly period interactions:

$$Y_{it} = \alpha_i + \lambda_t + \sum_{k \neq -1} \mu_k \cdot \mathbf{1}[b(t) = k] + \eta \cdot \text{Treat}_i + \sum_{k \neq -1} \beta_k \cdot \mathbf{1}[b(t) = k] \cdot \text{Treat}_i + \varepsilon_{it} \quad (9)$$

where $b(t)$ maps week t to its bi-weekly bin k, with $k < 0$ denoting pre-treatment bins and $k \geq 1$ denoting post-treatment bins. The bin $k = -1$ (the bi-weekly period immediately before treatment) is omitted as the reference period. Note that $\sum_{k \neq -1} \mu_k \cdot \mathbf{1}[b(t) = k]$ is absorbed by the week fixed effects λ_t, and $\eta \cdot \text{Treat}_i$ is absorbed by the individual fixed effects α_i in estimation. Each coefficient β_k captures the treatment-control difference in week bin k relative to the reference period. Parallel trends requires $\beta_k \approx 0$ for all $k < 0$, while $\beta_k > 0$ for $k \geq 1$ would indicate a positive treatment effect that persists or evolves over time.

Heterogeneity. To examine heterogeneous treatment effects, we augment each main specification with a triple interaction term. For direct effects (Tables 3 and 4), we augment equation (7):

$$Y_{id} = \alpha + \beta_1 \cdot \text{POST}_d + \beta_2 \cdot \text{Treat}_i + \delta \cdot \text{POST}_d \cdot \text{Treat}_i + \gamma \cdot H_i \cdot \text{POST}_d \cdot \text{Treat}_i + \varepsilon_{id} \quad (10)$$

where δ captures the treatment effect for low-H_i users and γ captures the differential effect for high-H_i users, so that $\delta + \gamma$ is the total treatment effect for the high group.

For downstream effects (Table 5), we augment equation (8):

$$Y_{it} = \alpha_i + \lambda_t + \beta \cdot \text{POST}_t \cdot \text{Treat}_i + \gamma_1 \cdot H_i \cdot \text{POST}_t + \gamma_2 \cdot H_i \cdot \text{Treat}_i + \delta \cdot H_i \cdot \text{POST}_t \cdot \text{Treat}_i + \varepsilon_{it} \quad (11)$$

where β is the treatment effect for the low-H_i group and $\beta + \delta$ is the total treatment effect for the high-H_i group. Note that H_i, Treat_i, and $H_i \cdot \text{Treat}_i$ are time-invariant and thus absorbed by the individual fixed effects α_i; the identified coefficients are β, γ_1, and δ. For the event study, we replace $\text{POST}_t \cdot \text{Treat}_i$ and $H_i \cdot \text{POST}_t \cdot \text{Treat}_i$ with their bi-weekly counterparts $\mathbf{1}[b(t) = k] \cdot \text{Treat}_i$ and $H_i \cdot \mathbf{1}[b(t) = k] \cdot \text{Treat}_i$, tracing out the dynamic heterogeneous effects across event-time bins.

5 Results

This section presents the empirical results in three parts. We first examine the direct effects of the information treatment on privacy tool usage (Section 5.1), then analyze downstream effects on data-consent decisions and platform engagement (Section 5.1.2), and finally decompose the effects along two dimensions of user heterogeneity (Section 5.2). Throughout, we link the findings back to the framework in Section 2.

5.1 Main Effect of Information Treatment

5.1.1 Direct Effects on Privacy Tool Usage

Table 1 presents estimates of Equation (9) for the three direct outcome variables. Column (1) reports the effect on visiting the Data Authorization List (DAL), the second-layer tool highlighted by the treatment. The coefficient on $Post \times Treat$ is 0.022 ($p < 0.01$), indicating that treated users are 2.2 percentage points more likely to visit the DAL after the treatment. Given the pre-treatment average of just 1.18%, this represents a 186% increase in the probability of visiting the tool, a large effect driven by the very low baseline.

Column (3) shows the effect on actively canceling an app's data authorization via the DAL. The coefficient of 0.018 ($p < 0.01$) indicates that treated users are 1.8 percentage points more likely to revoke an authorization, a 49% increase relative to the pre-treatment average of 3.69%. This result confirms that the information treatment not only increased DAL visitation but also triggered substantive privacy management actions.

Column (2) provides an important placebo-like check. The outcome is whether a user visits the Central Privacy Setting, the first-layer page that is broadly known to the population (over 70% of users report awareness of it in Q5). The null coefficient (0.012, $p > 0.10$) indicates that the treatment did not increase visits to this general page. This specificity is informative: the treatment directed users to the second-layer DAL tool rather than prompting a general increase in privacy-related exploration of the first-layer setting.

5.1.2 Downstream Effects on Engagement and Data Consent

Table 2 presents estimates of Equation (10) for downstream outcomes. The results reveal a striking asymmetry: the information treatment has no effect on contextual data-sharing decisions but significantly increases platform engagement and trust.

Columns (1) and (2) examine app-level data authorization decisions. Both the number of apps authorized (column 1) and the authorization rejection rate (column 2) show precisely estimated null effects, with point estimates near zero. This is consistent with the contextual

Table 1: Direct Effects of the Information Treatment

	dummy_dal (shock) (1)	dummy_setting (2)	dummy_cancel (3)
Post	0.005**	0.047***	0.027***
	(0.002)	(0.006)	(0.004)
Post × Treat	0.022***	0.012	0.018***
	(0.004)	(0.009)	(0.006)
Observations	16,548	16,548	16,548
Individual FE	Yes	Yes	Yes
Clustered error	user	user	user
Pre-shock avg.	1.18%	10.56%	3.69%

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

insensitivity result in Equation (8): contextual decisions about granting data access to individual apps are driven by the immediate utility of the app at the point of use and are not sensitive to platform-level information about privacy tools.

Columns (3)–(5) examine platform engagement and trust. Page views increase by 2.4% ($p < 0.05$), active duration increases by 2.7% ($p < 0.05$), and asset balances increase by 6.4% ($p < 0.10$). The asset balance result is particularly notable: it captures a financial commitment to the platform that reflects genuine trust, not merely casual browsing. Together, these results are consistent with Equation (6): the information treatment raises perceived ease of revocation r_i, which reduces the friction cost $\psi_i = \bar{\psi} + \theta_i(1 - r_i)$ and increases optimal engagement e_i^*.

Table 2: Downstream Effects of the Information Treatment

	app auth		engagement/activity		
	download	rejection rate	page view	duration	asset bal.
	(1)	(2)	(3)	(4)	(5)
Post × Treat	0.002	−0.002	0.024**	0.027**	0.064*
	(0.006)	(0.004)	(0.012)	(0.012)	(0.038)
Observations	124,110	124,110	165,480	165,480	165,480
Individual FE	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes
Clustered error	user	user	user	user	user

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$. Engagement outcomes are log-transformed.

The event study evidence in the next subsection provides a dynamic view of these effects and validates the parallel trends assumption.

5.1.3 Event Study Evidence

To validate the parallel trends assumption and examine the dynamics of the treatment effect, Figure 4 presents event study estimates for the three engagement outcomes. We estimate a fully dynamic version of Equation (10) that replaces the single $POST_t \cdot Treat_i$ interaction with a series of bi-weekly period interactions.

Two features of Figure 4 are noteworthy. First, the pre-treatment coefficients are close to zero and statistically insignificant across all three outcomes, confirming that the parallel trends assumption holds. Second, the treatment effects emerge gradually in the post-period rather than appearing as a sharp jump, consistent with transparency building trust over time rather than producing an immediate behavioral response. For page views (panel a), the coefficients grow steadily from approximately 0.01 in the first post-treatment period to 0.07 by the end of the observation window. This growing pattern suggests the effects may persist or even strengthen beyond our sample period.

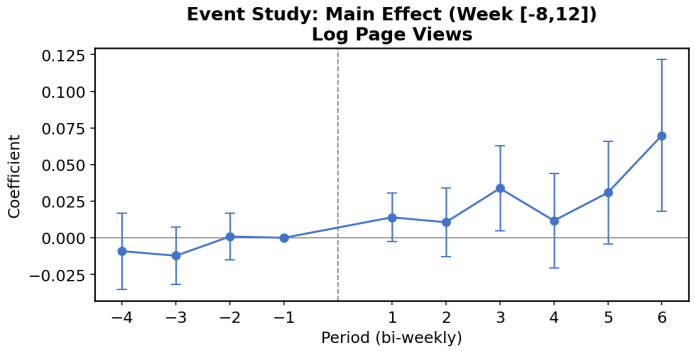
5.2 Heterogeneity

We now decompose the treatment effects along two dimensions, intrinsic privacy sensitivity and platform knowledge, to distinguish between the mechanisms identified in our conceptual framework.

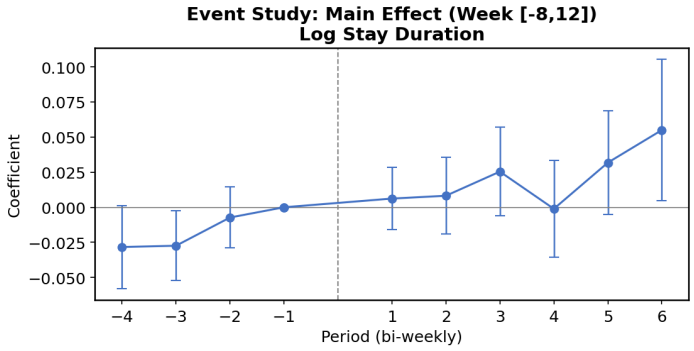
5.2.1 Intrinsic Privacy Preferences

Downstream Effects. The central finding of this paper is that the engagement and trust gains from the information treatment are entirely concentrated among users with high intrinsic privacy sensitivity. Estimating equation (11), we find that the triple interaction $Post \times Treat \times Intrinsic$ is large and significant for all three engagement outcomes: page views (-0.045, $p < 0.05$), active duration (-0.041, $p < 0.05$), and asset balances (-0.172, $p < 0.01$). The $Post \times Treat$ coefficient for the non-intrinsic baseline is near zero and statistically insignificant across all outcomes. The aggregate positive effects reported in Table 2 are therefore entirely attributable to the high-intrinsic subgroup.

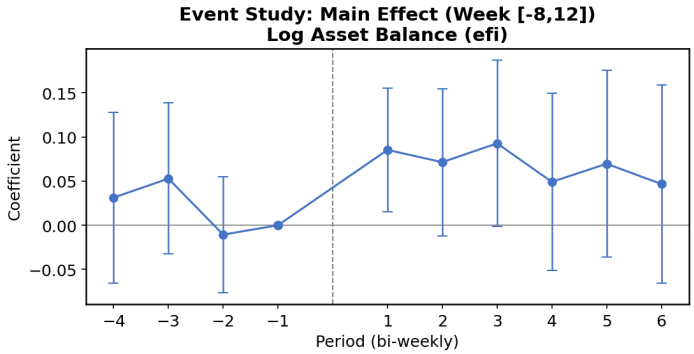
This result is the empirical punchline of the perceived-control mechanism formalized in Section 4. Recall from equation (5) that the treatment effect on engagement is $\Delta e_i^* = \theta_i \cdot (r_i^1 - r_i^0)/(-b''(e_i^*))$: it is proportional to the user's intrinsic privacy sensitivity θ_i. When users learn that past data authorizations can be centrally revoked, the engagement friction $\theta_i(1 - r_i)$ falls most sharply for those with the highest θ_i. For non-intrinsic users, privacy friction is modest to begin with, so information about revocability produces little change in



(a) Log Page Views



(b) Log Active Duration



(c) Log Asset Balance

Figure 4: Event Study: Main Effects on Platform Engagement

behavior. For intrinsic users, by contrast, the reduction in perceived irreversibility meaningfully lowers the cost of engagement, unlocking latent demand that was previously suppressed by privacy concerns. The asset balance result—a 17.2 percentage point increase for the intrinsic group—is particularly striking. Holding financial assets on a platform is a deliberate and reversible commitment that directly reflects trust; the magnitude of this effect suggests that the information treatment triggered a genuine shift in how high-intrinsic users perceive their relationship with the platform.

Event Study Evidence. Figure 5 presents the dynamic version of this result, plotting bi-weekly estimates of the triple interaction $Post \times Treat \times Intrinsic$ for each engagement outcome. Three features of these plots are worth emphasizing.

First, the pre-treatment coefficients are tightly clustered around zero in all three panels. This confirms that the differential response of high-intrinsic users is not an artifact of pre-existing divergent trends: the two groups were on parallel trajectories before the information shock, and the separation arises only after the treatment date.

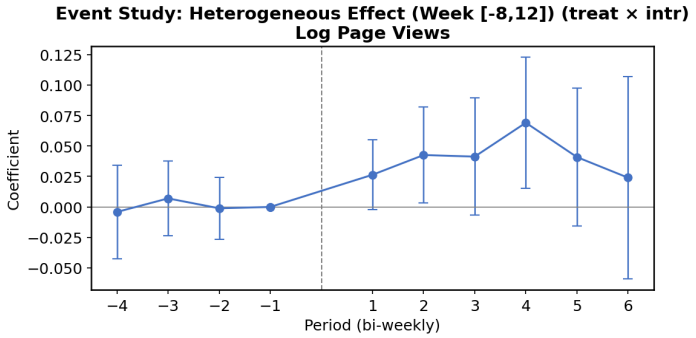
Second, the treatment effect emerges and *persists* across all post-treatment bi-weekly bins. The point estimates remain negative and economically meaningful throughout the 12-week post-treatment window, and if anything increase in magnitude over the first several bins before stabilizing. This persistence rules out a novelty interpretation in which users briefly explore the platform after learning about the tools and then revert to baseline. Instead, the sustained divergence is consistent with a durable shift in perceived control: once high-intrinsic users learn that their data exposure is reversible, this belief shapes ongoing engagement decisions over an extended horizon.

Direct Effects: A Mechanism Check. Table 3 examines whether intrinsic privacy sensitivity moderates the direct effects on privacy tool usage. All triple interaction terms ($Post \times Treat \times Intrinsic$) are small and statistically indistinguishable from zero across all three outcomes. High-intrinsic users are no more likely than low-intrinsic users to visit the data authorization list, visit the settings page, or cancel an app's data authorization after receiving the information treatment.

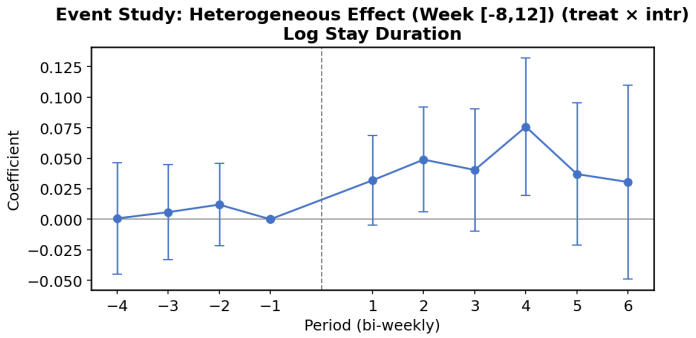
This null result is not a failure of the analysis but a key piece of the mechanism. If the engagement gains among high-intrinsic users were driven by increased tool usage—that is, if intrinsic users actually exercised greater revocation activity following the treatment—we would expect the triple interaction to be positive and significant here. It is not. The evidence therefore rules out a pure information channel operating through behavioral tool use, and instead supports the perceived-control interpretation: knowing the option to revoke exists reduces engagement friction for intrinsic users *even when the option is not exercised*. This is consistent with Tucker (2014) and with the option-value literature showing that the availability of a reversibility mechanism improves decisions even when it is rarely invoked (Brandimarte et al., 2013).

5.2.2 Platform Knowledge

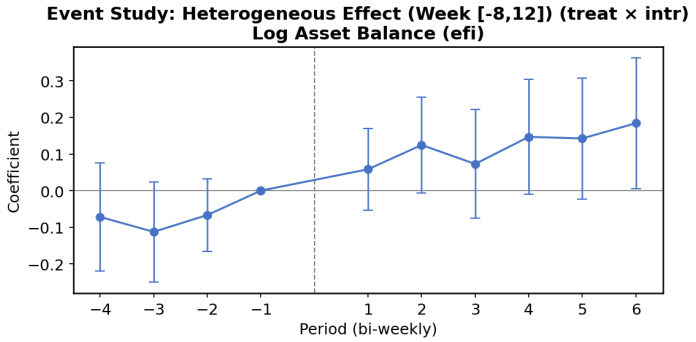
Direct Effects. Table 4 reports the direct effects moderated by platform knowledge. The triple interaction $Post \times Treat \times Knowledge$ is negative and significant for DAL visits



(a) Log Page Views



(b) Log Active Duration



(c) Log Asset Balance

Figure 5: Event Study: Heterogeneous Effects by Intrinsic Privacy Sensitivity

($-0.016, p < 0.10$) and authorization cancellations ($-0.023, p < 0.10$), indicating that the direct treatment effects are concentrated among low-knowledge users. This is consistent with Equation (7): the information treatment closes an awareness gap that is most binding for users who were previously unaware of the privacy tools.

Downstream Effects. Table 5 shows that the knowledge-based heterogeneity does not extend to downstream engagement outcomes. All triple interactions are small and statistically insignificant, confirming that the engagement and trust increases are not driven by the

Table 3: Direct Effects by Intrinsic Privacy Sensitivity

	dummy_dal (shock) (1)	dummy_setting (2)	dummy_cancel (3)
Post	0.003	0.037***	0.019**
	(0.005)	(0.011)	(0.008)
Post × Treat	0.023***	0.021	0.019
	(0.007)	(0.016)	(0.012)
Post × Intrinsic	0.003	0.014	0.010
	(0.005)	(0.013)	(0.010)
Post × Treat × Intrinsic	−0.001	−0.011	−0.001
	(0.009)	(0.019)	(0.014)
Observations	16,548	16,548	16,548
Individual FE	Yes	Yes	Yes
Clustered error	user	user	user

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 4: Direct Effects by Platform Knowledge

	dummy_dal (shock) (1)	dummy_setting (2)	dummy_cancel (3)
Post	0.005	0.041***	0.024***
	(0.003)	(0.006)	(0.005)
Post × Treat	0.033***	0.023**	0.033***
	(0.008)	(0.011)	(0.010)
Post × Knowledge	−0.004	−0.008	−0.010
	(0.004)	(0.007)	(0.007)
Post × Treat × Knowledge	−0.016*	−0.015	−0.023*
	(0.009)	(0.013)	(0.012)
Observations	16,548	16,548	16,548
Individual FE	Yes	Yes	Yes
Clustered error	user	user	user

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

low-knowledge users who exhibited the largest direct effects. This dissociation between the direct and downstream heterogeneity patterns is the key empirical finding that allows us to separate the information channel (Channel 1) from the perceived-control channel (Channel 2). If engagement increases were simply a byproduct of increased tool use, we would expect them to be concentrated among the same low-knowledge users who exhibited the largest direct effects. This is not what we observe. Instead, the engagement increases are driven by the high-intrinsic users who exhibited the largest downstream effects.

Table 5: Downstream Effects by Platform Knowledge

	app auth		engagement/activity	
	confirm	confirm/total	page view	asset bal.
	(1)	(2)	(3)	(4)
Post × Treat	−0.002	−0.007	0.017	0.043
	(0.009)	(0.006)	(0.018)	(0.072)
Post × Knowledge	−0.011	−0.011**	−0.009	−0.014
	(0.008)	(0.005)	(0.017)	(0.063)
Post × Treat × Knowledge	0.006	0.007	0.007	0.037
	(0.012)	(0.008)	(0.023)	(0.087)
Observations	124,110	124,110	124,110	124,110
Individual FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Clustered error	user	user	user	user

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

6 Conclusion

This paper provides causal evidence that transparency about privacy management tools, holding the tools themselves constant, significantly increases platform engagement and trust without reducing data sharing in the platform's third-party app ecosystem. The mechanism is that transparency increases perceived control, which reduces the intrinsic privacy costs users bear at the platform level and improves their decision environment. Using a large-scale field experiment on Alipay involving 8,274 respondents drawn from 4 million randomly selected active users, we show that an information treatment highlighting centralized data authorization controls produces three distinct effects.

First, the treatment substantially increases direct engagement with the highlighted privacy tools: treated users are 186% more likely to visit the data authorization list and 49% more likely to cancel an app's data authorization, relative to pre-treatment baselines. Second, the treatment increases downstream platform engagement, with page views rising by 2.4%, active duration by 2.7%, and asset balances by 6.4%, while having precisely estimated null effects on contextual app-level data consent decisions. Third, the direct and downstream effects are driven by distinct user segments: direct tool usage is concentrated among low-knowledge users (consistent with closing an awareness gap), while engagement and trust increases are concentrated among intrinsically privacy-sensitive users (consistent with enhanced perceived control).

Implications for Platform Strategy. These findings offer actionable insights for platform operators navigating the tension between user privacy empowerment and ecosystem

data needs. The key takeaway is that transparency pays. Consumers bear real privacy costs on platforms, and firms should not assume these costs are negligible simply because users continue to share data. Counterintuitively, transparency about data reversibility decreases these costs, not by changing consent-level choices, but by improving the decision environment at the platform level. Users who perceive greater control over their data face lower intrinsic friction, engage more deeply, and trust the platform more, without reducing the contextual data flows that sustain third-party services.

The heterogeneity results suggest that a differentiated communication approach may be optimal. For users with high intrinsic privacy sensitivity, the primary benefit of privacy communication is the trust and engagement it generates through enhanced perceived control, even without active tool use. For users with low platform knowledge, the primary benefit is awareness of tools they would otherwise not discover. Platforms could leverage these insights by tailoring privacy communications to user segments, though we note this recommendation is speculative and awaits direct experimental testing.

Implications for Privacy Regulation. Our findings also speak to ongoing regulatory debates about privacy disclosure requirements. Regulations like GDPR and PIPL mandate that platforms provide privacy controls, but there is limited evidence on whether informing users about these controls achieves the regulators' intended goals or produces unintended consequences. Our results suggest that information provision successfully directs users toward privacy tools (especially less sophisticated users) and builds platform trust, without producing the chilling effects on data sharing that platforms may fear. This supports the case for disclosure mandates as a complement to substantive privacy protections, with the caveat that our evidence comes from a setting where the platform itself, rather than a regulator, delivers the information.

Limitations. We acknowledge several limitations that point toward future research. First, our observation window captures effects over approximately ten weeks post-treatment; whether the trust and engagement effects persist, grow, or decay over longer horizons remains an open question. The event study evidence (Figure 4) suggests a growing pattern that may continue beyond our sample, but this is speculative. Second, while randomization addresses selection into treatment, the specific survey-based delivery mechanism may interact with user attention in ways that differ from other communication channels (e.g., push notifications, in-app banners). Testing alternative delivery mechanisms would strengthen external validity. Third, our measure of intrinsic privacy sensitivity relies on self-report (Q4a), which, while practically convenient and well-motivated by Lin (2022), may suffer from social desirability bias.

Fourth, we study a single platform in a specific regulatory environment; whether our findings generalize to platforms with different data practices, user demographics, or regulatory regimes is an empirical question.

Future Research. Several extensions would be valuable. Direct experimental comparison of different communication strategies, varying the framing, timing, and channel of privacy information, could inform optimal platform communication design. Structural estimation of the perceived-control mechanism, building on the framework in Lin (2022), could quantify the welfare implications of transparency. Finally, studying the general equilibrium effects of platform-wide privacy disclosure, including potential responses by third-party app developers, would complete the picture of how transparency reshapes digital ecosystems.

Overall, our findings demonstrate that the perceived tension between user privacy empowerment and platform data needs is less acute than commonly assumed. Transparency about privacy features operates through perceived control rather than through chilling effects: it improves the decision environment by giving users clarity and the freedom to reverse past choices, which reduces intrinsic costs and builds trust. Transparency pays, not despite the existence of consumer privacy costs, but because of them.

References

Alessandro Acquisti, Curtis Taylor, and Liad Wagman. The economics of privacy. *Journal of Economic Literature*, 54(2):442–492, 2016.

Aaron Adalja, Jura Liaukonyté, Emily Wang, and Xinrong Zhu. GMO and non-GMO labeling effects: Evidence from a quasi-natural experiment. *Marketing Science*, 42(2):233–250, 2023.

Idris Adjerid, Alessandro Acquisti, Rahul Telang, Rema Padman, and Julia Adler-Milstein. The impact of privacy regulation and technology incentives: The case of health information exchanges. *Management Science*, 62(4):1042–1063, 2016.

Idris Adjerid, Alessandro Acquisti, and George Loewenstein. Choice architecture, framing, and cascaded privacy choices. *Management Science*, 65(5):2267–2290, 2019.

Sebastián Araya, Andrés Elberg, Carlos Noton, and Daniel Schwartz. Identifying food labeling effects on consumer behavior. *Marketing Science*, 41(5):982–1003, 2022.

Susan Athey, Christian Catalini, and Catherine E. Tucker. The digital privacy paradox: Small money, small costs, small talk. NBER Working Paper 23488, National Bureau of Economic Research, 2017.

Bryan Bollinger, Phillip Leslie, and Alan Sorensen. Calorie posting in chain restaurants. *American Economic Journal: Economic Policy*, 3(1):91–128, 2011.

Laura Brandimarte, Alessandro Acquisti, and George Loewenstein. Misplaced confidences: Privacy and the control paradox. *Social Psychological and Personality Science*, 4(3):340–347, 2013.

Gordon Burtch, Anindya Ghose, and Sunil Wattal. The hidden cost of accommodating crowdfunder privacy preferences: A randomized field experiment. *Management Science*, 61(5):949–962, 2015.

Long Chen, Yadong Huang, Shumiao Ouyang, and Wei Xiong. Data privacy, data sharing and credit access. Technical report, Working Paper, 2025.

Chiara Farronato, Andrey Fradkin, and Tesary Lin. Designing consent: Choice architecture and consumer welfare in data sharing. Working Paper 34025, National Bureau of Economic Research, 2025.

Avi Goldfarb and Verina F. Que. The economics of digital privacy. *Annual Review of Economics*, 15:267–286, 2023.

Avi Goldfarb and Catherine E. Tucker. Privacy regulation and online advertising. *Management Science*, 57(1):57–71, 2011.

Avi Goldfarb and Catherine E. Tucker. Shifts in privacy concerns. *American Economic Review*, 102(3):349–353, 2012.

Tesary Lin. Valuing intrinsic and instrumental preferences for privacy. *Marketing Science*, 41(4):663–681, 2022.

Amalia R. Miller and Catherine E. Tucker. Privacy protection and technology diffusion: The case of electronic medical records. *Management Science*, 55(7):1077–1093, 2009.

Helen Nissenbaum. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press, 2010.

Shumiao Ouyang. Cashless payment and financial inclusion. Working paper, Princeton University, 2022.

Verina F. Que. How do past privacy choices shape the future? Working Paper 6432042, SSRN, 2026. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6432042.

Catherine E. Tucker. Social networks, personalized advertising, and privacy controls. *Journal of Marketing Research*, 51(5):546–562, 2014.