Consumer Value of Privacy: Evidence from an Online Retailer

Mimansa Bairathi
UCL School of Management, m.bairathi@ucl.ac.uk

Ankit Sisodia
Purdue University, asisodia@purdue.edu

Mayur Choudhary
Indiana University, mchoudh@iu.edu

We study whether and how much consumers value personal information using data from a large online fashion retailer's tiered loyalty program. Exploiting the quasi-experimental variation in the effective benefit the program's structure creates where a fixed loyalty points offered for data disclosure are more valuable when a consumer is closer to a tier-upgrade threshold, we find that consumers value privacy. As the effective payoff increases, consumers are more likely to disclosure personal information. Moreover, the average cost of disclosure differ sharply by information type and demographic information is values less compared to contact details and payment card information, indicating greater sensitivity to contact and financial data. Finally, consumers with alternative ways to earn rewards through higher past spending or prior participation in activities, have a higher cost of disclosing personal information and are less responsive to disclosure incentives. Together, the results provide causal, field-based evidence on the monetary value of personal data and show how incentive design and context shape consumers' willingness to share information.

Key words: privacy, value of data, digital platform, loyalty program

1. Introduction

Recent data breaches and the risk of loss or theft of personal information have intensified concerns about firms' collection, storage, and use of consumer data. In 2024, more than 5.5 billion user accounts were compromised worldwide, roughly 180 accounts per second.[1] The financial stakes are similarly large – the global average cost of a data breach reached $4.88 million in 2024, a 10% increase from the year before.[2] In response, regulators have enacted data protection laws such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA), which aim to expand consumers' control over the collection and sharing of their personal data.

[1] Source: https://surfshark.com/research/study/data-breach-recap-2024?srsltid=AfmBOoq4K6OWPiecf1dltrZ3y4nWFh8rgQz-_4fy5ovisqvdPqoYgVKo (Accessed October 7, 2025)

[2] Source: https://newsroom.ibm.com/2024-07-30-ibm-report-escalating-data-breach-disruption-pushes-costs-to-new-highs (Accessed October 7, 2025)

Consumers navigating digital platforms face a complex decision environment when asked to share personal information. On one hand, sharing data can yield benefits such as personalized recommendations, convenience, targeted advertising, and monetary rewards in exchange for data. On the other hand, data sharing can expose consumers to risks, including data loss or theft in the event of breaches, further sale or sharing of data with third parties, and targeted price discrimination. Consumer surveys show high levels of privacy concern among the public, with over 80% of Americans saying the risks of data sharing outweigh the corresponding benefits (Auxier et al. 2019). However, despite this apparent concern, consumers often disclose data easily (Athey et al. 2017), a phenomenon termed the "privacy paradox".

However, emerging perspectives suggest that consumer privacy decisions may not be paradoxical but instead reflect a combination of economic, contextual, and control considerations. A privacy calculus view proposes that individuals weigh potential benefits against perceived risks when deciding whether to disclose personal information (Culnan and Bies 2003). Nissenbaum (Nissenbaum 2004, 2011) emphasizes that the appropriateness of a data request depends on whether it aligns with norms of the surrounding context. Altman (Altman 1975) further argues that individuals continually adjust informational boundaries to maintain control over access to the self. Taken together, these perspectives imply that privacy decisions should vary with the strength of incentives, the sensitivity of information, and the availability of alternative ways to obtain similar benefits.

Hence, in this paper we examine whether and how much consumers value personal information. We then examine whether valuation differs across information types. Finally, we examine heterogeneity in valuation based on consumers' ability to obtain equivalent rewards through alternative ways.

To answer these questions, we use data from a leading Indian online fashion retailer that operates a tiered loyalty program. Higher tiers provide consumers more benefits, such as, exclusive discounts, coupons, faster service, creating incentives for consumers to accumulate loyalty points and upgrade to a higher tier. Consumers can earn points either through making purchases or by taking part in consumer engagement campaigns on the platform. We focus on a campaign in which the retailer offered consumers loyalty points in exchange for disclosing specific personal information such as date of birth, mobile phone

number, and credit or debit card details. In the data we observe whether each consumer accepted any of these offers.

Using this setting, we contribute to the literature on the economics of privacy by providing revealed preference evidence in a natural field setting. We complement prior work by conducted by Athey et al. (2017) in lab setting and online experiments conducted in Lin (2022) using Qualtrics, and examine consumers' privacy decisions in a large-scale field setting on a digital platform. In this environment, consumers face consequential, real trade-offs when deciding whether to disclose personal information. Our design thus complements incentivized experimental evidence by examining disclosure choices and the effect of context in a natural consumer–firm relationship and marketplace dynamics.

Our empirical strategy exploits quasi-experimental variation in the marginal value of loyalty points that arises based on proximity to the threshold for upgrading to the higher tier in the loyalty program. For a fixed number of loyalty points, consumers closer to the upgrade threshold derive a larger benefit from disclosure than consumers farther from the threshold. Using the within-consumer variation in the effective payoff in exchange for disclosure that arises due to variation in distance to upgrade due to transactions, we estimate a model with consumer fixed effects to absorb any time-invariant differences in privacy preferences across consumers. This mitigates the concern that consumers who shop more might thus be closer to the upgrade threshold and at the same time also have a higher preference for privacy.

We find evidence that consumers value privacy. The probability of disclosure increases as the distance to the upgrade threshold decreases, that is, as the effective payoff increases. This indicates that personal information has positive value.

We also find that valuation differs across information types. For a given payoff, consumers are less likely to disclose mobile numbers and credit or debit card details compared to date of birth.[3] This suggests that consumers consider contact information and financial information to be more sensitive compared to demographic information (Cranor et al. 2000). This shows that consumers value their personal information although they become more likely to disclose personal information as incentives increase, consistent with an incentives–benefits trade-off.

[3] Our results suggest consumers' valuation of mobile number and credit or debit card details does not differ significantly.

To quantify the monetary cost of disclosure, we leverage the structure of the tiered loyalty program. Since the effective benefit of the fixed number of loyalty points increases as consumers near the upgrade threshold, the same number of points can be more or less valuable depending on the consumers' distance to upgrade threshold. We observe instances in which a consumer does not disclose at one distance but chooses to disclose as the distance to upgrade threshold decreases, which allows us to infer a range within which the consumer's cost of disclosure lies. We aggregate these range across all consumers to recover an empirical distribution of the disclosure cost across different types of personal information. Using the monetary benefit of upgrading to the higher tier in the loyalty program, we recover the revealed cost of disclosing each piece of personal information in dollars.

We find that consumers place a positive, though economically modest, cost on disclosing personal information. In addition, this cost varies by the type of information. Among the consumers who disclose, for date of birth, the mean cost of disclosure is approximately \$0.10. The costs are higher for more sensitive data. The cost of disclosing mobile number is \$0.51, and credit or debit card details is \$0.52. These patterns indicate that while the absolute economic cost consumers place on disclosure of personal information is low, the valuations are statistically meaningful and heterogeneous across information types. Consumers value contact and financial information more highly than basic demographic data (Cranor et al. 2000).

Finally, we examine heterogeneity in consumers' cost of disclosure based on their ability to obtain equivalent loyalty points through alternative means. We focus on two channels through which consumers earn loyalty points. First, consumers can earn loyalty points by making purchases on the platform. We find that consumers with higher past spending on the platform have a significantly higher cost of disclosing personal information. This suggests that consumers who can accumulate loyalty points through purchasing have less incentive to compromise privacy for additional loyalty points. Second, consumers can earn loyalty points by engaging in activities or campaigns, such as, playing games on the platform. We find that consumers who previously participated in atleast one activity on the platform exhibit significantly higher disclosure costs. In contrast, consumers with no prior activity are more responsive to incentives. This pattern aligns with boundary-regulation theory proposed by Altman (1975) such that when viable methods to earn loyalty points

exist, consumers maintain tighter informational boundaries and demand more to disclose. On the other hand, when alternatives are limited, they are more willing to relax boundaries.

Our findings have implications for firms and policymakers. For firms that rely on first-party data collection, the results show that disclosure behavior is responsive to incentive design and to how data requests are incorporated into the consumer experience. Firms can therefore improve outcomes by timing requests near salient loyalty milestones and by tailoring strategies toward consumer segments with fewer alternative paths to benefits. At the same time, the nontrivial dollar valuations consumers place on personal information support the rationale behind growing regulatory efforts to enhance consumer protections. Policymakers should consider that even modest rewards can motivate consumers to share sensitive data, underscoring the need for transparent data-use policies and for ensuring that incentive mechanisms do not exploit behavioral frictions.

2. Literature

This paper relates to two streams of literature.

First, we contribute to the economics of privacy by analyzing how consumers trade off the benefits and risks of personal data disclosure. Prior research shows that disclosure decisions are influenced by perceived costs and benefits (Culnan and Bies 2003) and may appear inconsistent with consumers' stated privacy concerns, a phenomenon labeled the privacy paradox (Athey et al. 2017). We extend this stream of research by providing evidence from a real-world field setting in which consumers face actual economic trade-offs when deciding whether to disclose personal information. Our findings show that disclosure increases as incentives rise, indicating that privacy concerns do translate into behavior. At the same time, the cost of disclosure is economically low which might explain why consumers are willing to share data freely.

Recent perspectives argue that privacy decisions also depend on the appropriateness of the request. According to contextual integrity, consumers are sensitive to whether data requests align with the normative expectations of a situation (Nissenbaum 2004, 2011). Boundary regulation theory further emphasizes that individuals continually adjust informational boundaries to maintain a desired level of control (Altman 1975). We contribute to these discussions by demonstrating empirically consumers with alternative paths to loyalty points, such as earning points through purchases or participating in engagement activities,

maintain stronger boundaries, underscoring the joint role of context and control in privacy decisions.

Empirical studies also show that privacy preferences are multi-dimensional and mechanism-sensitive. For example, consumers differentiate between instrumental harms and intrinsic objections to data use (Lin 2022), and perceptions of invasiveness depend on how data is collected or used (Jerath et al. 2025). Our study contributes to this stream of research by documenting that the valuation of personal information differs substantially across information types. Contact and financial details are valued significantly more than demographic details, which highlights which types of consumer data is more sensitive and the conditions under which incentives can overcome privacy barriers.

Second, our study relates to research on loyalty program design and consumer behavior. Tiered loyalty programs segment customers by activity and provide progressively richer benefits as consumers approach higher tiers (Bergemann and Bonatti 2019, Nunes and Drèze 2006). A key mechanism is the goal-gradient effect—increased effort as individuals get closer to a goal (Kivetz et al. 2006, Nunes and Dreze 2006, Kopalle et al. 2012). Prior work has examined purchase responses to tier thresholds, but little is known about how non-purchase actions, such as disclosure of personal data, respond to these same incentives. We show that proximity to upgrade thresholds increases willingness to disclose and extend the goal-gradient logic to privacy-related effort. In addition, we show that consumers with strong alternative ways to progress in the program (e.g., high purchase activity) are less responsive to disclosure incentives. This insight connects loyalty program design with privacy management by demonstrating how incentive structures can be used to manage customer–firm data exchange while respecting privacy concern of consumers.

3. Empirical Setting & Data

3.1. Empirical Setting

Our data comes from a major online fashion retailer in India. The retailer operates a tiered loyalty program and customers can enroll in the program by registration and opt-in through the program page. There is no membership fee. Consumers can upgrade between tiers by accumulating loyalty points such that consumers between 0 and 499 points are in Tier 1 and consumers from 500 points or more are in Tier 2. Once achieved, tier status remains valid for 12 months. The Tier 2 status provides several benefits including coupons and exclusive access to discounts.

Consumers can accumulate loyalty points in two ways. First, they can earn loyalty points through purchases such that consumers earn 1 loyalty point for every 10 Indian rupees (INR) spent.[4] Second, consumers can also earn loyalty points by participating in activities or campaigns, such as playing games, on the platform.

The platform conducted three separate activities or campaigns where consumers were offered a fixed number of loyalty points in exchange for disclosure of different pieces of personal information. From June 12 to 18, 2019, the platform offered consumers ten points in exchange for their date of birth. Similarly, the platform offered consumers loyalty points in exchange for their mobile number. They conducted this activity twice, from December 15 to 19, 2018, where they offered consumers 50 points in exchange for their mobile number and again from June 12 to 18, 2019, where they offered consumers 30 points in exchange for their mobile number. Finally, the platform also offered consumers loyalty points in exchange for their credit or debit card details during these two period. In December they offered 50 points and in June they offered 20 points in exchange for the credit or debit card details. Each of these three activities are independent of each other and consumers were notified about each of these three activities separately. Finally, as a part of these activities, consumers can only earn points once in exchange for disclosing a specific piece of information.

3.2. Data

We use three sets of data from the retailer. The first data set records consumer purchase and engagement behavior on the platform. It records every date on which a consumer earns loyalty points, the number of loyalty points earned, and the manner in which the consumer earns these points, either by making a purchase or completing an activity. The second data set records the total number of loyalty points a consumers has on any given date. The third data set records the details of a consumer including the total purchase on the platform in the past and the activities they completed in the past.

Our data includes 88,966 consumers enrolled in the retailer's loyalty program between October 1, 2018, and June 18, 2018. Table 1 reports summary statistics. Panel A presents consumer characteristics measured prior to the disclosure period. Lifetime measures including total spend, items purchased, and items returned, are constructed from transaction

[4] One US dollar translates to 84.74 INR (as of December 2, 2024).

data spanning July 1, 2016 through through the beginning of the disclosure period. We also calculate platform visits in the 60 days preceding the disclosure period. The summary statistics indicate substantial heterogeneity in transactional behavior including past spending and visit frequency. In addition, the average customer in our data is highly engaged with the platform. Average lifetime spend is INR 22,989.5, with 25.7 items purchased and 6.4 items returned. In the 60 days preceding the campaign, customers visit the platform 18.3 times on average. Given the number of visits, it is unlikely that the typical consumer failed to notice the campaign offering loyalty points in exchange for personal information.

Panel B demonstrates the summary statistics of loyalty points during the disclosure periods. Again, we find substantial variation in loyalty points across consumers and consequently variation in distance, in loyalty points, to the upgrade threshold (mean = 270.27 points, SD = 142.85). We exploit this variation for identification in our empirical analysis. Finally, on average a consumer earns 118.86 loyalty points during the disclosure period. This suggests that the points offered in exchange for personal information - ranging from 10 to 50, represent meaningful incentives.

Table 1 Summary Statistics

Variable	Statistic
A: Consumer Level	
N	88,966
Mean Lifetime Total Spend in INR (SD)	22,989.5 (37,968.1)
Mean Lifetime Items Bought (SD)	25.7 (34.0)
Mean Lifetime Total Returns (SD)	6.4 (12.7)
Mean Visits in last 60 days (SD)	18.3 (15.3)
B: Observation Level	
Mean Points (SD)	281.58 (130.57)
Mean Distance to Next Tier in Points (SD)	270.27 (142.85)
Mean Change in Points (SD)	118.86 (183.50)

Note: Lifetime characteristics are calculated using data from July 1, 2016 to the disclosure period. Visits are measured in the 60-day window preceding the privacy disclosure period.

4. Estimation and Results

Table 2 reports outcomes from the disclosure campaign. Overall disclosure rates are low. 1.23% of consumers share their date of birth, and the rates are lower still for mobile number and credit or debit card details - 0.31% and 0.12%, respectively. In this field setting

where consumers interact with an online retailer, very few choose to exchange personal information for rewards. These figures are consistent with consumers taking part in other campaigns on the platform.

Table 2 Disclosure of Personal Information

Variable	Statistic
Date of birth	
Total Eligible Customers	59,480
Disclosure %	1.23%
Mobile number	
Total Eligible Customers	88,966
Disclosure %	0.31%
Credit or debit card details	
Total Eligible Customers	88,966
Disclosure %	0.12%

Note: The firm did not ask for birthday information during December 2018.

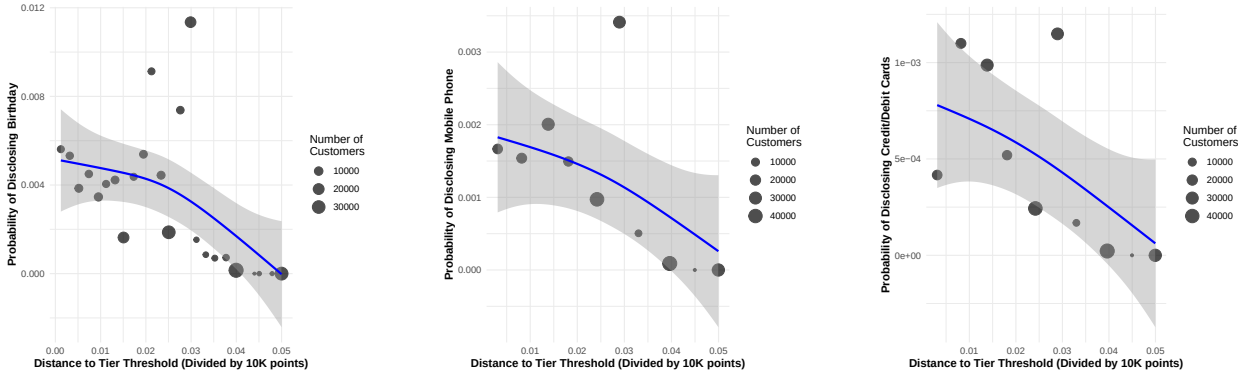
To causally identify whether consumers value their personal information and examine the effect of changes in payoff on consumers' willingness to disclose personal information, we exploit the structure of the retailer's loyalty program. The program assigns consumers to tiers, with higher tiers offering greater benefits. Hence, consumers have incentive to accumulate loyalty points in order to cross the threshold into a higher tier. As a result, the effective marginal benefit from earning a fixed number of loyalty points in exchange for disclosing personal information depends on how close a consumer is to the upgrade threshold.

To illustrate the intuition, consider two consumers A and B. Suppose the retailer offers 20 loyalty points in exchange for disclosing date of birth. Consumer A is 50 points below the threshold required to upgrade to the next tier, while consumer B is 200 points below that same threshold. If both disclose their date of birth, both would receive 20 points. However, the effective payoff is higher for A than for B. For A, the additional 20 points move them substantially closer to the upgrade threshold, and may even be pivotal in crossing it, thereby unlocking higher-tier benefits. For B, the same 20 points are less valuable because B remains far from the upgrade threshold. In other words, although the nominal reward is identical, its effective value is increasing in proximity to the tier threshold.

Hence, our empirical strategy leverages this plausibly exogenous heterogeneity in effective payoffs generated by a consumer's distance to the upgrade threshold. Since we focus on the distance to the upgrade threshold, we focus our analysis on customers in Tier 1 and examine their progression toward Tier 2. We test whether consumers closer to the threshold, who face a higher effective payoff for disclosure, are more likely to share personal information than consumers farther from the threshold. Two institutional features support this identification approach. First, disclosure requests were introduced platform-wide on fixed calendar dates and not targeted to specific users based on user behavior. Second, because consumers join the program at different times and accumulate points at different rates, there is substantial cross-sectional variation in how close each consumer is to the upgrade threshold at the moment a disclosure opportunity appears. These differences generates a quasi-random variation in effective payoff when presented with the opportunity to disclosure personal information.

Figure 1a plots the average probability that a consumer discloses their date of birth as a function of their distance (in loyalty points) from the upgrade threshold. The probability of disclosure declines with distance such that consumers closer to upgrading are more likely to disclose their date of birth. Figure 1b shows a similar pattern for disclosure of mobile phone numbers and Figure 1c shows a similar relationship for the disclosure of credit or debit card details. These plots provide face validity for our identification strategy.

Figure 1 Likelihood of information disclose based on proximity to upgrade threshold



(a) Birthday **(b) Mobile number** **(c) Credit or debit card details**

A potential concern could be that unobserved heterogeneity in privacy preferences could confound the relationship between distance to the upgrade threshold and disclosure behavior. For example, consumers who are digitally more savvy may be, both, more concerned

about data privacy and purchase more online, thereby accumulating more loyalty points be closer to the upgrade threshold. If so, distance to the threshold could be correlated with underlying privacy preference, biasing our estimates. We address this concern by including consumer fixed effects, which absorb all time-invariant individual heterogeneity in privacy preferences.

To measure the effect of payoff on disclosure, we estimate the following specification:

$$\mathbb{1}\{\text{If disclose}\}_{it} = \beta_0 + \beta_1 \text{Distance (in points)}_{it} + \alpha_i + \epsilon_{it} \tag{1}$$

where the dependent variable If disclose_{it} is an indicator equal to 1 if consumer i discloses a given piece of personal information at time t. $\text{Distance (in points)}_{it}$ is consumer i's distance, in loyalty points, from the upgrade threshold at time t[5], and α_i denotes consumer fixed effects. We estimate equation (1) using ordinary least squares (OLS) and cluster standard errors at the consumer level to account for serial correlation in the dependent variable. Although the dependent variable is binary, we prefer the linear probability model because the coefficient β_1 has a straightforward marginal interpretation and the inclusion of high-dimensional fixed effects makes nonlinear alternatives (e.g., logit with fixed effects) less transparent.

Table 3 presents the results across all information types. Column (1) shows that the probability of disclosing date of birth decreases with the distance to the upgrade threshold ($p < 0.01$). This indicates that consumers respond to incentives - they are more willing to disclose date of birth when the effective payoff from disclosure is higher. Column (2) presents results for disclosure of mobile phone number. Again, the probability of disclosure declines with distance to the threshold ($p < 0.01$), implying that consumers are more likely to share their mobile number when the implied payoff is larger. Column (3) shows a similar pattern for disclosure of credit or debit card details. The likelihood of providing payment card information decreases with distance to the threshold ($p < 0.01$), consistent with consumers requiring higher payoffs to disclose more sensitive financial information.

Taken together, these results suggest that consumers attach value to their personal data and are willing to trade personal information for benefits. Moreover, the magnitude of responsiveness varies by information type, consistent with consumers viewing contact and financial information as more sensitive than demographic information.

[5] We divide the tier points by a factor of 10,000 for ease of interpretation.

Table 3 Effect of distance on probability of disclosing

	(1) Date of birth	(2) Mobile number	(3) Credit or debit card details
Distance (in points)	-1.212307***	-0.1164873***	-0.0565535***
	(0.049185)	(0.0078318)	(0.0060873)
Consumer F.E.	Yes	Yes	Yes
Observations	110,708	233,398	173,443

Notes: Clustered standard errors in parentheses.

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

4.1. Robustness Checks

To examine whether consumers decision to disclosure personal information is sensitive to incentives, we conduct several robustness checks.

First, rather than use a continuous measure of distance to the upgrade threshold, we use a binary measure which indicates if disclosing the personal information would move the user across the upgrade threshold. Table 4 shows that across all information types, users who would cross to the higher tier upon disclosure are more likely to disclose the personal information compared to users who would not cross the tier.

Table 4 Robustness: Effect of crossing the tier on probability of disclosure

	(1) Birthday	(2) Mobile no.	(3) Credit/debit card
If tier cross	0.0081082***	0.00416953***	0.00114918**
	(0.0020415)	(0.00086197)	(0.00047705)
User F.E.	Yes	Yes	Yes
Observations	110,708	233,398	173,443

Notes: Clustered standard errors in parentheses. Distance in points divided by 10,000.

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

Second, one concern might be that our findings may be driven by user inattention rather than response to incentives. If many users are not actively engaged with the platform and are thus unaware of their proximity to the upgrade threshold, the observed relationship between disclosure and distance may simply reflect noise rather than deliberate decision making. To address this, we first restrict the sample to users who visited the platform

within the 15 days leading up to the disclosure campaign, a group that is more engaged and thus less likely to have missed the disclosure campaign. Columns (1)-(3) in Table 5 show the results continue to hold and that disclosure decreases with distance to the threshold, consistent with the main findings. We then further restrict the sample to consumers who visited the platform in the seven days leading up to the disclosure campaign, capturing an even more engaged set of the users. As shown in Columns (4)-(6) in Table 5, our results continue to hold. Users who are closer to the upgrade threshold are more likely to disclose their date of birth, mobile number, and credit or debit card details. This suggests that the patterns are unlikely to be an artefact of inattention, but instead reflect users' response to incentives.

Table 5 Robustness using active users

	Active in last 15 days			Active in last 7 days		
	(1)	(2)	(3)	(4)	(5)	(6)
	Birthday	Mobile no.	Credit/debit card	Birthday	Mobile no.	Credit/debit card
Distance (in points)	-1.224680***	-0.1238909***	-0.0614995***	-1.197559***	-0.1282870***	-0.064310***
	(0.053882)	(0.0091068)	(0.0069716)	(0.054625)	(0.0097695)	(0.007379)
User F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Observations	82772	184976	141694	75680	168970	129478

Notes: Clustered standard errors in parentheses. Distance in points divided by 10,000.

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

Third, another concern could be that the distance to the upgrade threshold may be confounded with the amount of time that remains to reach it. recollect that, consumers only have one year to upgrade to the next higher tier after which the number of tier points resets to zero. If users who are farther from the threshold also tend to have more time remaining to upgrade, our results might partially reflect differences in time left rather than responsiveness to incentives for disclosure. To account for this, we control for the number of days that remain for the opportunity to upgrade in equation 1. Table 6 shows that the results remain robust both in magnitude and significance even after we account for the time left for upgrade.

5. Quantifying Consumer Value of Privacy

While our results show that consumers respond to incentives while deciding to share personal information, it is unclear how much is their cost of disclosing personal information

Table 6 Robustness to including time left to upgrade

	(1)	(2)	(3)
	Birthday	Mobile no.	Credit/debit card
Distance (in points)	-1.179021***	-0.13888939***	-0.062068***
	(0.050416)	(0.00909865)	(0.0066462)
Time left to Upgrade	-0.163161**	0.00130533***	0.00056104***
	(0.069144)	(0.00011743)	(0.00006465)
User F.E.	Yes	Yes	Yes
Observations	110708	233398	173443

Notes: Clustered standard errors in parentheses. Distance in points divided by 10,000.

Pending Time for Upgrade (in days) divided by 100

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

in economic terms. Quantifying the monetary cost of disclosure of different types of personal information provides a more interpretable measure of privacy preferences and allows comparisons across data categories, consumer segments, and policy interventions. It also offers insight into the privacy paradox, in which consumers express strong privacy concerns yet frequently disclose personal data easily. A dollar-denominated estimate also enables researchers and practitioners to benchmark consumers' valuations against the implicit or explicit prices that firms assign to personal data through loyalty benefits, targeted offers, or data monetization practices. Hence, we next translate observed disclosure decisions into revealed monetary costs of sharing personal information, allowing us to assess the economic significance of privacy concerns and to examine how valuations vary across information types and consumers.

Our empirical setting allows us to bound the value of privacy of consumers. Each consumer finds themselves at a distance d_i from the upgrade threshold. Revealing their personal information results in some payoff that depends on this distance. The natural variation in d_i means that each consumer is awarded randomized payoff $\{y^m\}_{m=0}^{M}$ to reveal their personal information with $y^m < y^n$ for $m < n$. We observe a consumer i not disclosing personal information at a certain payoff, say y^0, only to disclose when awarded y^1. This indicates that the cost of disclosure $c_i \in (y^0, y^1]$. Aggregating such choice by consumers i belonging to a certain group I, allows us to back out the proportion of individuals in group I with the cost of disclosure between $(y^0, y^1]$. Performing this exercise for all such

payoff intervals yields the approximate empirical distribution of cost of disclosure. First, we provide the connection between the notional payoffs $\{y^m\}_{m=0}^M$ and the expected payoff gain at distance d_i. Later, we describe the procedure to back out this distribution based on the consumers' disclosure decision.

5.1. Expected Payoff for Consumers

Assume that consumers churn out of the platform at a rate δ per loyalty point[6] and $\mathbf{Z}_i$ denotes consumer characteristics. If a consumer i is at x_i points, the distance from the upgrade threshold D_0 is,

$$d_i = D_0 - x_i$$

Before the threshold, consumers get a normalized value of 0, but beyond the threshold, they get $R > 0$ per loyalty point, i.e., their utility from p points.

$$u(p) = \begin{cases} 0 & d_i \geq p \\ R(p - d_i) & d_i < p \end{cases}$$

This does not include the lifetime continuation value after crossing the threshold, R/δ.

The action space is $a_i^j \in \{0, 1\}$ with $a_i^j = 1$ indicating disclosure of personal information j. The cost of disclosing personal information of type j is some function of consumer characteristics, $c_i^j = c^j(\mathbf{Z}_i; \theta)$. The reward for disclosing personal information is p^j points.

For ease of notation, we drop j, the type of personal information. The expected utility of the consumer i from the loyalty points is

$$\begin{aligned} \mathbb{E}[u(d_i, a_i = 0; \delta)] &= 0 \cdot Pr(\text{Churns before } d_i) + \frac{R}{\delta} \cdot Pr(\text{Doesn't churn before } d_i) \\ &= 0 \cdot \left(1 - e^{-\delta d_i}\right) + \frac{R}{\delta} e^{-\delta d_i} \end{aligned}$$

The expected utility from disclosing personal information is,

$$\mathbb{E}[u(d_i, a_i = 1; \delta)] = \begin{cases} \frac{R}{\delta} e^{-\delta(d_i - p)} - c_i & d_i \geq p \\ \frac{R}{\delta} + R(p - d_i) - c_i & d_i < p \end{cases} \tag{2}$$

ASSUMPTION 1 **(Identifying Assumption)**. *The distance from the upgrade threshold is independent of consumer characteristics before the disclosure of personal information.*

$$d_i \perp\!\!\!\perp \mathbf{Z}_i$$

[6] δ captures the rate of churn per loyalty point. This assumption can be micro-founded from minimal assumptions: Suppose $\tilde{\delta}$ per unit time is the rate of churn (death), and there is deterministic accumulation of loyalty points at τ intervals at an exponential rate of $\frac{\eta}{\tau}$ (so that per unit time accumulation rate is η). As $\tau \to 0$ in the continuous limit, the total points at time t approaches $\frac{t}{\eta}$. Naturally, the death time $T \sim \exp(\tilde{\delta})$ implies that the probability that the consumer churns before accumulating points p is $F_P(p) = \left(1 - \exp(\tilde{\delta}\eta p)\right)$. In our model, we assume $\delta = \tilde{\delta}\eta$.

5.2. Consumer Choice

The consumers disclose their personal information if,

$$\mathbb{E}[u(d_i, a_i = 1; \delta) - u(d_i, a_i = 0; \delta)] \geq 0$$

In other words, the cost of disclosing personal information is low enough

$$c_i \leq \begin{cases} \frac{R}{\delta} e^{-\delta d_i}(e^{\delta p} - 1) & d_i \geq p \\ \frac{R}{\delta}\left(1 - e^{-\delta d_i}\right) + R(p - d_i) & d_i < p \end{cases}$$

We can simplify the decision to disclose as $a_i = 1$ if $c_i \leq Rf(d_i; \delta)$ and $a_i = 0$ otherwise. Here, $f(d; \delta)$ is the net expected payoff gain from disclosure scaled by the benefit R. This expected payoff gain is decreasing in the distance d_i.

$$f(d_i) = \begin{cases} \frac{1}{\delta} e^{-\delta d_i}(e^{\delta p} - 1) & d_i \geq p \\ \frac{1}{\delta}\left(1 - e^{-\delta d_i}\right) + (p - d_i) & d_i < p \end{cases}$$

PROPOSITION 1. *The consumers are more likely to disclose their personal information the closer they are to the threshold.*

$$\frac{\partial}{\partial d_i} Pr(a_i = 1 | d_i) < 0$$

Intuitively, the consumers value being on the other side of the threshold. Their expected continuation value is greater if they are closer to the threshold. Therefore, we can rationalize the empirical findings in the previous sections, that consumers closer to the threshold are more likely to disclose by simply incorporating a churn rate δ.

5.3. Empirical Distribution of Cost of Disclosure

The identifying assumption can also be re-stated as follows – the distance is independent of cost of disclosure, i.e.,

$$\{d_i, f(d_i; \delta)\} \perp\!\!\!\perp c_i \text{ for all } i$$

We estimate the churn rate δ using data for a period of 75 days preceding the disclosure campaign by estimating a Cox-hazard model. We consider consumers that did not make a purchase for a period of 30 consecutive days as consumers who have churned. This churn rate is group-specific.

We construct a discrete grid of $(M + 1)$ values of the scaled expected payoff values $\{y^m\}_{m=0}^{M}$ with $y^m < y^n$ for any $m < n$. Because $f(d; \delta)$ is monotonic, we invert it to arrive at $(M + 1)$ discrete values of distance from threshold, $\{d^m\}_{m=0}^{M}$,

$$d^{M-m} = f^{-1}(y^m; \delta^I)$$

Consider a consumer i belonging to group I. Suppose the consumer discloses at distance $d_i \in [d^{M-m-1}, d^{M-m})$, while does not disclose at an observed distance $d_i \geq d^{M-m}$. Based on the consumer's choice we infer that,

$$y^m = f(d^{M-m}; \delta^I) < c_i^I \leq f(d^{M-m-1}; \delta^I) = y^{m+1}$$

Following our identifying assumption, d_i is independent across consumers. This is akin to an experiment where we reward $\{y^m\}_{m=0}^M$, randomly to consumers and observe whether they disclose their personal information. The proportion of such consumers who disclose their personal information when offered y^m or below is the proportion of consumers with cost of disclosure below y^m. We derive such proportions conditional on consumers who disclose personal information. That is, we focus on the estimator of the distribution conditional on disclosure, $D = \{i : a_i = 1\}$ and for all $m \in \{0, 1, \ldots, M\}$.[7]

$$\begin{aligned} F_C^I(y^m) \approx \hat{F}_C^I(y^m) &= \hat{Pr}(C \leq y^m = f(d^{M-m}; \delta^I)|D) \\ &= \frac{\hat{Pr}(C \leq f(d^{M-m}, \delta^I), i \in D)}{\hat{Pr}(i \in D)} \\ &= \frac{\sum_{i=1}^N \mathbb{1}\{a_i = 1, d_i \geq d^{M-m}, i \in D\}}{\sum_{i=1}^N \mathbb{1}\{i \in D\}} \end{aligned}$$

The above approximation becomes sharper the finer the partition is for $\{d^m\}_{m=0}^M$ (or $\{y^m\}_{m=0}^M$). The final expression is the estimate of the approximate empirical distribution. It captures the proportion of consumers belonging to category I with privacy valuation below y^m, and is computed by calculating the proportion of individuals who disclosed their private information when at a distance greater than or equal to d^{M-m}.

5.4. Cost of Disclosure of Personal Information

In the previous subsection, we described the method to obtain the approximate empirical distribution. However, this distribution is over the scaled cost of disclosure. We need to rescale it by R to arrive at the real value in USD. Since we define R as the extra dollar amount each loyalty point is worth in the higher tier relative to the lower tier, to calculate

[7] Note that we focus on the estimates conditional on disclosure. If the consumers are randomly sampled, $Pr(C < c) = Pr(C < c|D)$. Alternatively, we recover only a part of the distribution. We refrain from making any claims regarding the cost of disclosure for consumers who never disclose, as such claims are likely confounded by consumers' lack of attention on the platform.

R, we focus on the main benefits consumers derive from upgrading to the higher tier in the loyalty program. First, consumers receive a one time vouchers worth INR 200 when they upgrade. Second, being in the higher tier of the loyalty program gives consumers exclusive access to an additional 5% discount for the last day in every month. On average, a consumer spends INR 1,216.63 in the last week of every month. Assuming, consumers get a 5% additional discount on this sum, it translates into a total discount of 729.98 INR in a year. Hence, the total savings is 929.98 INR in a year and given an average consumer spends INR 34,348.39 in a given year and hence earns 3,435 loyalty points, the extra amount in INR that each loyalty point is worth in the higher tier is $929.98/3435 = 0.27$INR. To convert this into US dollars, we divide this figure by 20.3, the purchasing power parity in 2018-19.[8] This translates to $R = 0.27/20.3 = 0.013$.

Figure 2 shows the empirical CDF of cost of disclosure, scaled by R, for different types of information. Figure 2b shows the CDF for cost of disclosure of date of birth.[9] Among the consumers who disclosure, the mean valuation of date of birth is approximately \$0.10 ($p < 0.01$).[10] Similarly, Figures 2b and 2c show the CDF for cost of disclosure of mobile number and credit or debit card details, respectively. The cost of disclosure is higher for more sensitive data. The mean cost of disclosure of mobile number is \$0.51 ($p < 0.01$), and the mean cost of disclosure of credit or debit card details is \$0.52 ($p < 0.01$). These figures demonstrate that privacy has a real and heterogeneous economic value to consumers, and that contact and financial data are valued more than demographic data (Cranor et al. 2000).

6. Heterogeneity in Valuation Across Consumer Segments

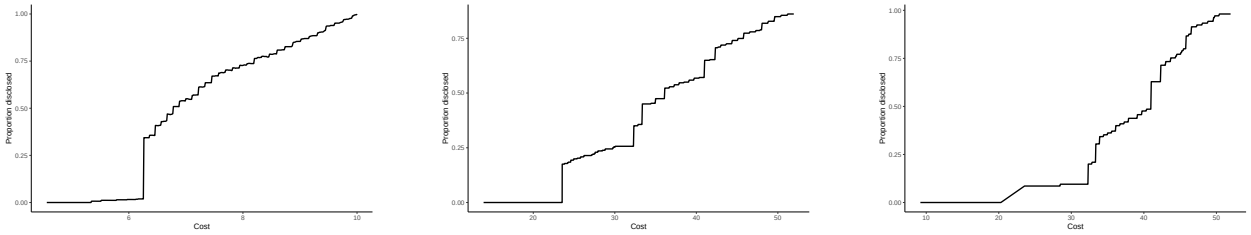
While our results show that consumers value personal information, the cost if disclosure is unlikely to be uniform across individuals. Consumers differ in their engagement levels, accumulated benefits, and paths to progress within the loyalty program, all of which may shape how costly it feels to disclose personal data. Understanding who values privacy more, and under what circumstances, is essential for both theory and managerial decisions regarding incentive design. Hence, we explore heterogeneity in disclosure costs based on consumers' ability to earn loyalty points through other means.

[8] Source: https://data.un.org/Data.aspx?d=WDI&f=Indicator_Code%3APA.NUS.PPP (accessed on November 25, 2025).

[9] scaled by R

[10] We compute standard errors using bootstrapping.

Figure 2 Empirical CDF of valuations (scaled by R)



(a) Date of birth **(b) Mobile number** **(c) Credit or debit card details**

In this loyalty program, consumers earn loyalty two points through both, purchases and participating in engagement activities (e.g., games, campaigns). These alternative loyalty points-earning channels shape how much incremental value consumers derive from receiving loyalty points in exchange for data disclosure.

To do this, after deriving the approximate empirical distribution $\hat{F}_C^I(v)$, we can also compare whether the average cost of disclosure differs across groups $I \in \{0,1\}$ based on $\mathbf{Z} \in \{\mathbf{Z}_0, \mathbf{Z}_1\}$. To do so, we parameterize the distribution to a gamma distribution $\Gamma(\alpha, \lambda)$ where α is the shape and λ the scale with the mean given by α/λ

$$F_C^I(c) = \begin{cases} \Gamma\left(\mu^0, \frac{\mu_0}{\theta^0}\right) & I = 0 \\ \Gamma\left(\mu^1, \frac{\mu_1}{\theta^0+\theta^1}\right) & I = 1 \end{cases}$$

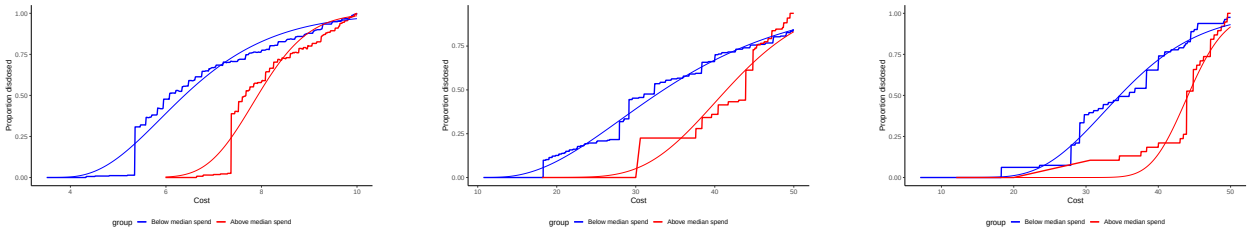
Given this parameterization , $E[C^I] = \theta^0 + \mathbb{1}\{I=1\}\theta^1$. Therefore, $E[C^1] - E[C^0] = \theta^1$. We estimate the parameters $(\mu^0, \mu^1, \theta^0, \theta^1)$ using non-linear least squares to match the distribution, $F_C^I(c)$ to $\hat{F}_C^I(c)$. We conduct bootstrapping to repeatedly draw and construct samples to obtain the confidence intervals for θ^1. This allows us to conduct a test of means over the cost of disclosure across different groups.

We begin with purchases, the primary channel for accumulating loyalty points. Consumers who spend more can progress toward tier upgrades through routine buying alone. For them, points earned via disclosure provide lower marginal benefit, because advancement is achievable without relinquishing personal information. This logic implies a higher cost of disclosure of personal information among high spenders, which we test empirically. Using historical purchase data, we classify consumers with spending at or above the median as high spenders and the remainder as low spenders.

Figure 3 shows the empirical CDF for the cost of disclosure of different types of information among these two groups of consumers. Table 7 summarizes the results. Columns

(1)–(2) show average valuations by group across the three information types, and Column (3) shows that high spenders have a significantly higher cost of disclosure across all types of information ($p < 0.01$). The gap is especially pronounced for more sensitive information such as mobile number and payment card details. The cost of disclosure of date of birth is \$0.02 higher for high spenders relative to low spenders. Similarly, their cost of disclosure of mobile number and payment card details is higher by \$0.08 and \$0.12.

Figure 3 Empirical CDF of valuations based on past spending (scaled by R)



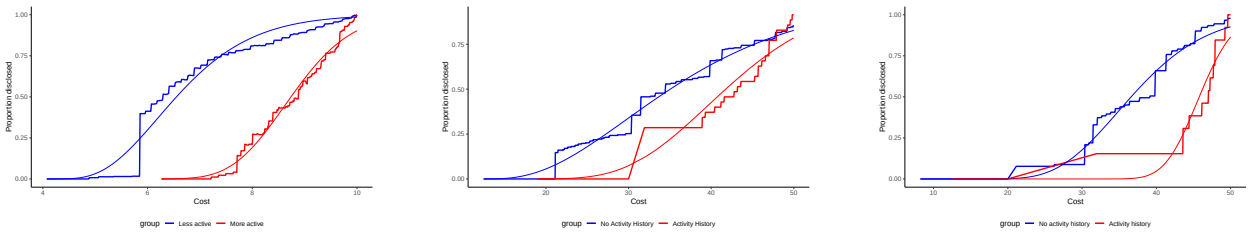
(a) Date of birth **(b) Mobile number** **(c) Credit or debit card details**

Second, we focus on activities, such as games and platform challenges, another mechanism for accumulating loyalty points. Consumers who regularly participate in activities are more likely to advance without sharing data. Thus, loyalty points in exchange for personal information again carry lower marginal value for them, implying higher cost of disclosure. We classify consumers with at least one prior activity prior to disclosure campaign as more active and others as less active.

Figure 4 shows the empirical CDF for the cost of disclosure of different types of information among more and less active consumers. Columns (4)–(6) of Table 7 indicate that more active consumers have significantly higher cost of disclosure than less active consumers ($p < 0.01$), with larger differences for mobile and payment card information. The cost of disclosure of date of birth is \$0.03 higher for more active consumers relative to less active consumers. Similarly, their cost of disclosure of mobile number and payment card details is higher by \$0.07 and \$0.11.

Taken together, these patterns show that cost of disclosure depends not only on information sensitivity and incentive strength, but also on the availability and perceived appropriateness of alternate pathways to benefits. When consumers have viable alternatives to earn points, they value privacy more. On the other hand, when alternatives are limited, they

Figure 4 **Empirical CDF of valuations based on past activity (scaled by R)**



(a) Date of birth **(b) Mobile number** **(c) Credit or debit card details**

Table 7 **Difference in cost of disclosure based on behavioral factors**

	Past spend			Past activity		
	(1) Low Spenders	(2) High Spenders	(3) Δ	(4) Less Active	(5) More Active	(6) Δ
Birth date	0.09	0.11	0.02***	0.09	0.12	0.03***
Mobile number	0.48	0.56	0.08***	0.50	0.57	0.07***
Credit or debit card	0.47	0.59	0.12***	0.50	0.61	0.11***

Notes: Bootstrapped confidence intervals are in parentheses.

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

are more willing to trade information. This is consistent with boundary-regulation theory Altman (1975) such that when valued benefits are accessible through other behaviors, such as purchasing and participating in activities, consumers maintain tighter informational boundaries. Managerially, this suggests that data requests are likely to be more effective among consumers without strong alternative paths to tier progress.

7. Discussion & Conclusion

As regulatory initiatives such as the GDPR and CCPA expand consumers' control over the use of personal data, firms increasingly rely on voluntary data sharing to guide personalization, targeting, and customer analytics. In these environments, consumer privacy preferences become central to the success of data collection efforts. Yet, consumers face a complex trade-off - sharing data may yield convenience, personalization, or direct economic benefits, while simultaneously exposing them to risks such as data breaches, unwanted contact, or discriminatory pricing. Understanding how consumers evaluate these trade-offs in real economic environments is essential for both researchers and practitioners.

Using revealed preference data from a large online retailer, this paper examines whether and how much consumers value their privacy. Our setting leverages quasi-experimental variation in the value of incentives due to the tiered structure of the loyalty program. This allows us to observe consumers' disclosure decisions where real financial benefits are at stake, overcoming concerns that stated intentions may not reflect actual behavior.

Three key insights emerge from our analysis. First, we find strong evidence that consumers value privacy. The likelihood of disclosure increases as the effective payoff grows, that is, as consumers approach a tier upgrade threshold. This provides causal evidence of a privacy calculus such that consumers weigh rewards against the perceived cost of revealing personal information.

Second, by using this variation, we recover the economic cost of disclosure of personal information. We find that demographic data such as date of birth is valued much lower ($\approx$ \$0.10) than more sensitive information such as mobile number and payment card details ($\approx$ \$0.50), consistent with heterogeneity in privacy preferences based on data type.

Third, we identify heterogeneity in cost of disclosure across consumers. Consumers with strong alternative means of progressing in the loyalty program, through frequent purchasing or participation in activities, attach significantly higher cost to disclosing personal information. These patterns align with boundary-regulation theory such that when valued benefits can be achieved through norm-consistent behavior, consumers maintain stricter informational boundaries. On the other hand when alternatives are limited, they become more willing to trade privacy for rewards. Thus, differences in consumer-firm relationship strength and incentive salience jointly shape the privacy calculus.

Our findings are important for managers and policymakers. For managers, our results offer clear guidance on designing data-collection strategies. Data requests are most successful when they occur near salient thresholds and incentives are commensurate with the sensitivity of the data requested. Further, because consumers with fewer alternative paths to rewards are more likely to disclose, firms must ensure that incentive structures do not disproportionately pressure disengaged customers in ways that may harm trust or fairness. For policymakers, our findings underscore the limits of notice-and-consent models. Even modest incentives can induce the sharing of sensitive data. Regulation may need to account not only for transparency of data use but also for contextual appropriateness and the influence of behavioral frictions on disclosure.

While our study provides valuable insights, there are limitations to our research. It focuses on a single retailer in India, and cost of disclosure might differ across cultural, regulatory, and market contexts. Additionally, we observe short-run disclosure decisions in response to one-time incentives and cannot speak to long-term evolution of privacy attitudes or potential learning effects. Future research could examine recurring disclosure requests, cross-cultural variation, and how data breaches or trust-building initiatives reshape the privacy calculus over time.

Notwithstanding these limitations, our findings reveal that consumers do care about privacy in the marketplace, although the average cost of disclosure is low. By quantifying the economic value of cost of disclosure using field-based revealed preferences data, this study bridges theoretical perspectives and managerial relevance, advancing our understanding of how consumers navigate the trade-offs inherent in digital data exchange.

References

Altman I (1975) The environment and social behavior: privacy, personal space, territory, and crowding. .

Athey S, Catalini C, Tucker C (2017) The digital privacy paradox: Small money, small costs, small talk. Technical report, National Bureau of Economic Research.

Auxier B, Rainie L, Anderson M, Perrin A, Kumar M, Turner E (2019) Americans and privacy: Concerned, confused and feeling lack of control over their personal information .

Bergemann D, Bonatti A (2019) The economics of social data: An introduction .

Cranor LF, Reagle J, Ackerman MS (2000) Beyond concern: Understanding net users' attitudes about online privacy .

Culnan MJ, Bies RJ (2003) Consumer privacy: Balancing economic and justice considerations. *Journal of social issues* 59(2):323–342.

Jerath K, Miller KM, Sokol DD (2025) Towards developing an understanding of consumers' perceived privacy violations in online advertising. *Available at SSRN 4736957* .

Kivetz R, Urminsky O, Zheng Y (2006) The goal-gradient hypothesis resurrected: Purchase acceleration, illusionary goal progress, and customer retention. *Journal of marketing research* 43(1):39–58.

Kopalle PK, Sun Y, Neslin SA, Sun B, Swaminathan V (2012) The joint sales impact of frequency reward and customer tier components of loyalty programs. *Marketing Science* 31(2):216–235.

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

Nissenbaum H (2004) Privacy as contextual integrity. *Wash. L. Rev.* 79:119.

Nissenbaum H (2011) Privacy in context: Technology, policy, and the integrity of social life. *Journal of Information Policy* 1:149–151.

Nunes JC, Dreze X (2006) The endowed progress effect: How artificial advancement increases effort. *Journal of Consumer Research* 32(4):504–512.

Nunes JC, Drèze X (2006) Your loyalty program is betraying you. *Harvard business review* 84(4):124–31.

Appendix A: Data

Table 8 Illustrative Example: Transaction History and Tier Progress for a Representative Consumer

Action Date	Credit Date	Source Type	Transaction Type	Points	Cumulative Points (Redemption)	Cumulative Points (Tier Progress)	Current Tier	Distance To Next Tier
12/1/18	12/1/18	Welcome Points	Credit	100	100	100	1	400
12/2/18	12/2/18	Activity	Credit	400	500	500	2	1500
12/14/18	12/14/18	Purchase	Due	21	500	500	2	1500
12/16/18	12/16/18	Activity	Credit	100	600	600	2	1400
12/16/18	12/16/18	Redemption	Debit	80	520	600	2	1400
12/22/18	12/22/18	Purchase	Due	97	520	600	2	1400
12/2/18	1/7/19	Purchase	Credit	28	548	628	2	1372
12/2/18	1/7/19	Purchase	Credit	30	578	658	2	1342
12/2/18	1/7/19	Purchase	Credit	28	606	686	2	1314
12/2/18	1/7/19	Purchase	Credit	68	674	754	2	1246