Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The Effects of Personal Data Management on Competition and Welfare

Jiajia Cong, Noriaki Matsushima

informs
https://pubsonline.informs.org/journal/mnsc
MANAGEMENT SCIENCE
Articles in Advance, pp. 1–19
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

The Effects of Personal Data Management on Competition and Welfare

Jiajia Cong,[a] **Noriaki Matsushima**[b,*]

[a] College of Business, City University of Hong Kong, Kowloon Tong, Hong Kong; [b] Osaka School of International Public Policy, University of Osaka, Toyonaka, Osaka 560-0043, Japan

*Corresponding author

Contact: jiajia.cong@cityu.edu.hk, https://orcid.org/0009-0004-0967-130X (JC); matsushima@osipp.osaka-u.ac.jp, https://orcid.org/0000-0002-7865-3875 (NM)

Received: October 6, 2024
Revised: August 4, 2025
Accepted: November 15, 2025
Published Online in Articles in Advance: April 1, 2026

https://doi.org/10.1287/mnsc.2024.08024

Abstract. This study examines the impact of consumers' personal data management on firm competition in the data collection and application markets, as well as on welfare outcomes. Consumers purchase products from differentiated firms in these two markets. Initially, in the data collection market, firms compete to collect consumer data to predict preferences in the data application market, enabling them to offer personalized prices to their targeted customers. Before firms set prices in the data application market, targeted consumers can erase their data at a fixed cost, thereby becoming untargeted. We show that personal data management leads to higher prices, lower firm profits, negative externalities among consumers, and reduced consumer surplus in the data application market. In the data collection market, personal data management *intensifies competition* and improves consumer surplus. Our analysis further explores extended scenarios of personal data management—including the rights to opt out of data collection, data portability, and data ownership—and reveals that firms' two-market profits and total consumer surplus change nonmonotonically with consumers' data rights. Specifically, granting consumers data rights beyond simple data erasure can reverse the effects of erasure alone, benefiting firms while harming consumers. Moreover, data management influences cross-market competition quite differently depending on which market consumers exercise their data rights. Finally, we discuss proactive strategies firms can employ to maintain profitability under personal data laws.

History: Accepted by Dmitri Kuksov, marketing.

Funding: J. Cong acknowledges the financial support from National Natural Science Foundation of China (Projects 72473024, 72192845) and support from the Shenzhen Research Institute, City University of Hong Kong. The authors gratefully acknowledge the financial support from the Japan Society for the Promotion of Science (JSPS), KAKENHI [Grants JP19H01483, JP20H05631, JP21H00702, JP21K01452, JP21K18430, JP23K20593, and JP23K25515], Nomura Foundation, the International Joint Research Promotion Program at University of Osaka, and the program of the Joint Usage/Research Center for "Behavioral Economics" at ISER, University of Osaka.

Supplemental Material: The online appendix is available at https://doi.org/10.1287/mnsc.2024.08024.

Keywords: personal data management • data collection • data application • price discrimination • personal data laws

1. Introduction

Personal data has become a pivotal asset in the modern economy. Data gathered from customers' product usage holds significant value across a firm's multiple business domains, especially in revealing consumer preferences. This data-centric approach fosters strong integration among business units through coordinated data collection and utilization strategies. Consequently, firms are increasingly dedicating substantial resources to acquire and harness personal data, applying these insights across diverse industries to secure competitive advantages.

Numerous examples illustrate firms' efforts to collect and utilize consumer data across markets.[1] Major auto manufacturers are developing smart car systems to collect granular driving data for usage-based insurance, such as pay-how-you-drive programs. For instance, Tesla, Ford, and GM offer usage-based insurance that provides tailored premiums based on customer characteristics and collected driving data.

Similarly, e-commerce giants like Alibaba and JD.com, competing to serve millions of customers, leverage their collected e-commerce data to assess creditworthiness and offer personalized, in-house financial services with tailored interest rates and prices. Additionally, the ride-hailing platform Bolt collects comprehensive travel data from its users, leveraging this information to personalize its food delivery offers.

In digital health, Google has partnered with various medical institutions and controversially acquired Fitbit in 2021, gaining access to millions of consumers' health data. Simultaneously, Google is expanding its health and insurance businesses with data-driven personalized offers (Chen et al. 2022, Ozalp et al. 2022). Amazon gathers consumer data through its e-commerce platform and smart devices (e.g., smart speakers) and has made significant investments in health and medical services via a series of acquisitions (Ozalp et al. 2022).

The extensive collection and utilization of consumer data have significantly enhanced firms' ability to use personalized offers/pricing to extract consumer surplus. Recent empirical studies have demonstrated the efficacy of such data-driven price discrimination (Shiller 2020, Smith et al. 2022, Dubé and Misra 2023), prompting extensive policy debates (Goldfarb and Tucker 2019, Wagner and Eidenmüller 2019).[2]

However, mounting concerns about personal data have spurred a global wave of personal data legislation, such as the EU General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA). These laws empower consumers with greater control over their personal data, aiming to enhance individual welfare. Among these rights, data erasure is the most prominent, followed by data portability. In response, firms across various industries have implemented privacy policies that acknowledge and support consumers' ability to manage their personal data. This paper examines the implications of these personal data management rights on market competition, firm profitability, and welfare across both data collection and application markets.

This study considers a duopoly model in which consumers purchase products in two *independent* markets: market *B* (e.g., wearable devices and cars) is for data collection, while market *A* (e.g., healthcare and insurance) is for data application. Consumers' product preferences in the two markets are independent. Two multimarket firms compete in both markets, initially attempting to attract consumers in market *B* with uniform pricing. The personal data collected by each firm reveals its customers' product preferences in market *A*. For instance, granular biometric data collected by smartwatches can help uncover consumers' preferences for healthcare and insurance products. Thus, in our model, purchase behavior in market *B* does not reveal consumer preferences in market *A*; however, collected data can. After competition in market *B*, each firm in market *A* offers personalized prices to its targeted consumers, for whom it has collected data, and a uniform price to untargeted consumers, for whom it does not have data.

Before firms decide on prices in market *A*, each firm's targeted consumers can erase their data from the firm's database to become untargeted consumers at a fixed cost, an option referred to as *data management*. Consumers who erase their data (i.e., opt-out consumers) can escape personalized prices and choose the uniform prices offered by firms. In our benchmark scenario, no consumer can manage their personal data.

We find that consumer data management leads to higher prices, lower firm profits, negative externalities among consumers, and reduced consumer surplus in market *A*. The explanation for the higher prices begins with the types of consumers who self-select to erase their data: targeted consumers with a strong preference for the firm's product in market *A* erase their data to avoid being charged high personalized prices. Consequently, each firm's uniform price rises because this price is applied to these additional opt-out consumers, who have a high willingness to pay for the firm's product. This price increase incentivizes firms to set higher personalized prices for consumers who remain opt-in, as their outside option—purchasing at the rival's higher uniform price—becomes less attractive. In other words, opt-in consumers bear negative externalities generated by opt-out consumers' data management. As each opt-out consumer downplays the pricing externalities, some marginal opt-out consumers may end up paying a higher uniform price than their personalized prices in the benchmark scenario without data management.

We now explain the impacts on firm profits and consumer surplus in market *A*. A firm's profits decline because uniform pricing fails to efficiently extract surplus from opt-out consumers with high reservation values, even though the firm can more efficiently extract surplus from opt-in consumers via higher personalized prices. Opt-in consumers become worse off due to the higher personalized prices. However, opt-out consumers are barely better off, as they pay an inflated uniform price along with data management costs, resulting in lower consumer surplus in market *A*. Although the data management cost represents a deadweight loss, increasing this cost can benefit both firms and consumers in market *A* by reducing the number of opt-out consumers and mitigating the aforementioned equilibrium changes.

Personal data management *intensifies competition* in market *B* by increasing firms' incentives to attract consumers and collect their data, although it decreases profits in market *A*. This occurs because a firm with a larger customer base in market *B* can inflate its uniform price more in market *A* due to a larger number of opt-out consumers. This price inflation reduces the incentive for marginal consumers to erase their data because the benefit of escaping personalized prices diminishes. Consequently, the firm can impose elevated personalized prices on more targeted consumers, particularly those with high willingness to pay, thereby enhancing profitability in market *A*. This dynamic makes a firm's market *A* profit more sensitive to its market share in market *B*, leading to fiercer competition in market *B*. As a result, firms earn lower profits in market *B*, while consumers benefit from intensified competition. Integrating both markets, personal data management reduces a firm's profits and increase total consumer surplus.

Building on the baseline model, we explore several extended scenarios related to personal data management, including: (i) consumers can opt out of data collection in market *B*, which incorporates the baseline model's no collection opt-out as a special case, (ii) consumers have additional data portability in market *A*, and (iii) consumers have an additional vertical characteristic of servicing costs. In extension (i), collection opt-out harms firm profits and benefits consumers if the number of collection opt-out consumers remains below a certain threshold. Beyond that threshold, the effects are reversed. Therefore, *collection opt-out, as an additional aspect of data management, has a nonmonotonic impact on firm profits and consumer welfare.* In extension (ii), data portability intensifies competition in market *A* but mitigates competition in market *B*. In extension (iii), an increase in heterogeneity of servicing costs accelerates competition in market *B*. We next provide the intuition behind the results of extensions (i) and (ii).

Introducing a data collection opt-out creates two consumer groups in market *B*: those who opt in and those who opt out. In market *A*, the presence of collection opt-out consumers *reduces* price competition. This is because competition for these opt-out consumers, which involves uniform pricing by both firms, is less intense than for opt-in consumers, where one firm employs uniform pricing while the other uses personalized pricing. Consequently, firms in market *A* obtain higher profits, while consumer surplus decreases. In market *B*, the effect of collection opt-out consumers depends on their prevalence. A small number of opt-outs initially intensifies price competition due to enhanced profitability of personalized pricing in market *A*. However, as the number of collection opt-outs grows beyond a certain threshold, this effect reverses, reducing competition in market *B* and resulting in higher firm profits and lower consumer surplus.

When data portability (GDPR Article 20) is available alongside data erasure in the data application market, the majority of consumers share their data with firms through portability, triggering Bertrand competition in personalized pricing for those consumers. This new effect diminishes the firm's profit in market *A*. Moreover, the lower profitability of personalized pricing reduces the sensitivity of profits in market *A* to market shares in market *B*, thereby weakening competition in market *B*. Overall, firms achieve higher profits across both markets.

Finally, we demonstrate that firms can proactively mitigate the negative impacts of data management on profits through two simple pricing strategies: (i) imposing a price cap and (ii) committing to avoid data-driven price discrimination. While these strategies limit firms' ability to fully leverage price discrimination, they discourage consumers' data erasure, a key source of negative impacts on profits. Furthermore, such constraints in market *A* can boost profits by reducing competition in the data collection market *B*.

2. Literature Review

Firstly, our study contributes to the literature of behavior-based price discrimination (BBPD). Traditional BBPD models (e.g., Caminal and Matutes 1990, Bester and Petrakis 1996, Chen 1997, Villas-Boas 1999, Fudenberg and Tirole 2000) assume that firms use purchase history for third-degree price discrimination, often resulting in intensified competition and reduced profits. Later research has explored profitable conditions for BBPD. For instance, Pazgal and Soberman (2008) highlight loyalty benefits, while Chen and Zhang (2009) show that limiting sales to loyal customers can soften competition. Rhee and Thomadsen (2017) find that BBPD can be profitable among vertically differentiated firms. Shin and Sudhir (2010) explore the effects of taste changes and demand heterogeneity, finding that taste changes facilitate customer poaching and higher period-two prices in many parameter sets. Additionally, demand heterogeneity among consumers makes poaching less aggressive, leading to profitable BBPD.[3]

Recent works consider BBPD with data-driven personalized pricing (e.g., Choe et al. 2018, 2022; Chen et al. 2022; Laussel and Resende 2022). Choe et al. (2018) demonstrate asymmetric outcomes with personalized pricing for returning customers, while Choe et al. (2022) explore the impact of data sharing. Laussel and Resende (2022) find personalized products can benefit firms that have fixed, differentiated products.[4] Chen et al. (2022) analyze a multimarket firm leveraging cross-market data for personalized pricing, highlighting potential consumer welfare implications.[5] In summary, typical BBPD models generally treat consumer data as a passive input, limiting consumers' ability to influence how firms collect and utilize their data.

Our paper introduces an intermarket model with independent preferences across markets and the ascendant concept of personal data management—consumer actions affecting data collection, application, or both. We use BBPD with personalized pricing as the benchmark and examine how such endogenous data management, by disrupting firms' ability to collect, track, and price discriminate, reshapes competition and welfare across data-connected markets.

Our analysis departs from the typical BBPD literature in three key ways. First, in typical BBPD studies (e.g., Shin and Sudhir 2010), completely independent consumer preferences (or locations) across periods mean competition in each period is independent of the other, yielding BBPD outcomes identical to those without it. Our model, however, links competition across the two markets through firms' data collection and application (i.e., preference prediction and personalized pricing), thereby capturing the essence of firms' data-driven practices. Second, we demonstrate that consumers' data management yields richer results than a mere reduction in BBPD. Notably, data management generates negative externalities among consumers, disrupts the typical seesaw pattern of firm profits described in BBPD literature, and can cause both firm profits and total consumer surplus to vary nonmonotonically with consumers' data rights. Moreover, it can influence cross-market competition in significantly different ways depending on which market consumers exercise their data rights. Third, we explore firm strategies to proactively offset the negative impacts of consumers' data management (Section 6). While our work relates to Choe et al. (2018) and Chen et al. (2022), our focus on consumers' data management and its cross-market competitive effects is novel and absent in prior work. Additionally, our model incorporates richer consumer heterogeneity.

Within the burgeoning literature on consumer data management with BBPD, our study closely relates to Ke and Sudhir (2023). They examine how consumer data management affects competition among *ex-ante* homogeneous firms, finding that data erasure and portability create a seesaw effect on pricing. While these policies limit price discrimination after data application, they also discourage firms from initially lowering prices to incentivize consumer opt-in. Ke and Sudhir (2023) conclude that such policies benefit consumers in a competitive setting.

Our model, however, reveals more nuanced effects of data erasure and portability on cross-market price competition, potentially producing positive and negative welfare consequences for consumers. These consequences are a direct result of data management's dual impact. It raises uniform prices in the data application market, creating negative externalities for opt-in consumers. Furthermore, the reduction in opt-in consumers fuels competition for market share in data collection because firms seek to acquire more data and discourage consumers from opting out.

Secondly, our study contributes to the literature on imperfect customer recognition in competitive environments (e.g., Chen et al. 2001, Liu and Serfes 2004, Esteves 2014, Colombo 2016, Belleflamme et al. 2020), all of which assume that firms operate with an exogenously given level of imperfect recognition capability and examine how this affects market competition. A common finding is that imperfect customer recognition can mitigate price competition. For instance, in Chen et al. (2001), a firm's loyal customer may receive the low price intended for switchers, discouraging firms from lowering their prices.[6]

In our model, imperfect consumer recognition arises from consumers' endogenous data erasure, meaning that only high willingness-to-pay consumers become unidentifiable. Consequently, firms increase their prices. We also demonstrate that some consumers enhance recognition through data portability, intensifying price competition. Furthermore, our study highlights how consumers' endogenous changes in recognition within the data application market influence firms' competition to acquire consumer data, which in turn affects consumers' incentives to either decrease or increase their recognition.

Finally, our paper contributes to the growing literature on the impact of consumers' endogenous privacy choices and data management on firms' pricing strategies. A pioneering work by Acquisti and Varian (2005) considers a two-period monopoly model where consumers can anonymize themselves in period two. They show that intertemporal price discrimination is unprofitable if product qualities are identical across periods or if forward-looking consumers can costlessly anonymize themselves.[7] Subsequent studies by Belleflamme and Vergote (2016) and Koh et al. (2017), also in monopoly settings, focused on data application, demonstrate that consumers' data management choices signal their willingness to pay, prompting monopolists to adjust discriminatory prices for those who manage their data and those who do not.[8]

While we acknowledge this signaling effect, our study departs by examining a competitive environment where both discriminatory and uniform prices are influenced by competition. This difference leads to significant variations in how, and to what extent, personal data management creates pricing externalities for other consumers.

In a competitive setting, Anderson et al. (2023) consider a two-stage game involving competition in listed prices during the first stage and targeted discounts during the second. Unlike our model, they assume that consumers opt in or out simultaneously as firms set list prices and *before* they know their product preferences. As a result, consumers' decisions to opt in or out do not signal their preferences, and firms cannot adjust list prices based on signaled preferences. Rhodes and Zhou (2026) discuss consumers' excessive data sharing to firms and intraconsumer externalities in a general oligopoly model.[9] Generally, these papers focus solely on how data management affect the competition in the data application market. Our paper contribute to this strand of literature by connecting how consumers' endogenous data management affect both data collection and application markets. Furthermore, we discuss how firms can proactively offset the negative impacts of personal data management.

3. Baseline Model

Consider two markets, A and B. Market B is data rich, and firms can collect substantial consumer data in this market. Market A is lucrative, and firms can predict consumers' product preferences of this market using their data from market B. Our model is best understood through a concrete example. Specifically, we refer to market B as the market for wearable devices, such as Fitbit wristbands and Apple watches, and market A as the market for healthcare and insurance. Needless to say, our model applies to other markets where data are collected in one market and applied in another market.

Two firms serve each market—A_1 and A_2 in market A and B_1 and B_2 in market B—where firms A_1 and B_1 are two subsidiaries of firm 1 and firms A_2 and B_2 are two subsidiaries of firm 2. The two markets have the same group of consumers, and their mass is normalized to one. In each market, a consumer demands one unit of product. We normalize the marginal cost of production to zero and treat prices as profit margins.

In market B, consumers are indexed by their product preference y, which is uniformly distributed on [0, 1]. Firms B_1 and B_2 are located at 0 and 1, respectively, and compete on uniform prices. The price set by firm B_i is β_i $(i = 1, 2)$. Consumer $y \in [0, 1]$ obtains utility $u = v_B - ky - \beta_1$ from firm B_1 and utility $u = v_B - k(1 - y) - \beta_2$ from firm B_2, where the value of v_B is large such that the market is fully covered, and k is a positive constant that captures taste mismatch. If a consumer purchases from firm B_i, then the firm can collect her data through her usage of the product. We assume that customers must consent to data collection in the market B to utilize data-driven products, such as smartwatches and advanced driving systems, as well as platform services effectively.[10]

In market A, firms A_1 and A_2 are located at the endpoints 0 and 1, respectively, of the interval [0, 1], which is different from the interval in market B. Consumers are indexed by their product preference x, which is uniformly distributed on [0, 1]. Each consumer identifies her location in market A upon entering that market. Each consumer's *product preferences* in the two markets (i.e., her locations x and y) are independent of each other because the products serve different purposes (Matutes and Regibeau 1988, Armstrong and Vickers 2010).[11] For example, a consumer's preference for smartwatches is independent of her preference for healthcare and insurance plans. Thus, firms in market A cannot infer consumers' locations based on their locations in market B. Consumers privately know their exact realized locations in market A, which can be interpreted as the healthcare or insurance plan that perfectly matches their needs.[12] Firm A_i obtains all of firm B_i's collected consumer data. This data transfer means that if a consumer has purchased from firm B_i, then firm A_i can uncover her exact realized location x in market A with the help of consumer data and target her. For example, with the assistance of granular biometric data collected by smartwatches, firm A_i can identify the healthcare and insurance plans that align with a consumer's preferences (i.e., her location x). However, firm A_j $(j \neq i)$ knows only that her location x is uniformly distributed on [0, 1]. Firm A_i charges personalized prices $p_i(x)$ to its targeted consumers and a uniform price α_i to untargeted consumers.[13]

After realizing her product preference in market A, each consumer can choose whether to erase her data from firm B_i's datasets or object to firm B_i's transfer of her data to firm A_i. This choice is referred to as *data management*. We assume that the cost of data erasure is $\varepsilon \geqslant 0$ for every consumer.[14] Data erasure prevents firms from offering personalized prices to consumers who erased their data.

If a consumer does not erase her data, then she continues to be a targeted consumer of firm A_i. Firm A_i's targeted consumer has two choices: receiving utility $u = v_A - tx_i - p_i(x)$ from firm A_i or receiving utility $u = v_A - tx_j - \alpha_j$ from firm A_j, where $x_i = 1 - x_j$. If she erases her data, she cannot be targeted by either firm and thus has two choices: obtaining utility $u = v_A - tx_i - \alpha_i$ from A_i or obtaining utility $u = v_A - tx_j - \alpha_j$ from A_j.[15] The value of v_A is large such that the market is fully covered, and t is a positive constant that captures taste mismatch

The whole game proceeds as follows. The two firms in market B simultaneously decide their uniform prices, and consumers make their purchase decisions in market B after observing the prices. Subsequently, firms in market B collect data from customers based on their usage of the products. Following this, consumers identify their product preferences over A_1 and A_2 in market A. They then simultaneously decide whether to erase their data. Next, the two firms in market A simultaneously post uniform prices. Thereafter, firm A_i offers private personalized prices to its targeted consumers. After observing all available offers, consumers make purchasing decisions in market A. The sequential timing of price offers in market A is standard in the literature on personalized pricing (Thisse and Vives 1988, Shaffer and Zhang 2002, Choe et al. 2018), and reflects the flexibility in choosing personalized prices and allows us to solve for the subgame perfect Nash equilibrium in pure strategies.

4. Analysis of the Baseline Model

In this section, we first establish the benchmark equilibrium of no data management in Section 4.1. Next, we establish the equilibrium under personal data management and compare it with the benchmark equilibrium to reveal the impacts of data management across both markets in Section 4.2. Section 4.3 discusses how personal data management affects consumer surplus and social welfare in both markets.

4.1. Benchmark: No Data Management

We consider the benchmark in which consumers cannot erase their data from firm B_i's datasets. Thus, if a consumer is firm B_i's targeted consumer, she must be firm A_i's targeted consumer. Suppose that firm B_1 wins consumers on $[0, \delta]$ and that firm B_2 wins consumers on $[\delta, 1]$ in market B. As shown in Figure 1(a), at any point $x \in [0, 1]$ in market A, a segment, δ, of consumers is targeted by firm A_1, and the remaining $1-\delta$ segment is targeted by firm A_2. In other words, at any point x in market A, firm A_i's targeted consumers are firm A_j's untargeted consumers.

Firm A_i sets personalized prices $p_i(x)$ for targeted consumers and a uniform price α_i for untargeted consumers. A firm's optimal personalized prices $p_i(x)$ make its targeted consumers indifferent between purchasing from the firm itself and purchasing from the rival under the uniform price. As a result, the two firms' optimal personalized prices are $p_1(x) = \alpha_2 + t(1-2x)$ and $p_2(x) = \alpha_1 + t(2x-1)$. Both firms offer higher personalized prices to targeted consumers whose preferences align more closely with their products.

Expecting such personalized prices, firm A_1 wins the rival's targeted consumers on $[0, \bar{x}_2]$ by offering the uniform price α_1, where the consumer $\bar{x}_2 = 1/2 - \alpha_1/(2t)$ is indifferent between purchasing from A_1 under α_1 and purchasing from A_2 under $p_2(\bar{x}_2) = 0$. So, firm A_1 earns a profit of $\alpha_1(1-\delta)\bar{x}_2$ from α_1. Similarly, firm A_2 wins the rival's targeted consumers on $[\bar{x}_1, 1]$ with its uniform price α_2, where $\bar{x}_1 = 1/2 + \alpha_2/(2t)$, and earns a profit of $\alpha_2\delta(1-\bar{x}_1)$ from α_2. The equilibrium uniform prices and cutoffs are:

$$\alpha_1^n = \alpha_2^n = \frac{t}{2}, \quad \bar{x}_1^n = \frac{3}{4}, \quad \bar{x}_2^n = \frac{1}{4}, \tag{1}$$

where the superscript n indicates no data management. Consequently, the equilibrium personalized prices are:

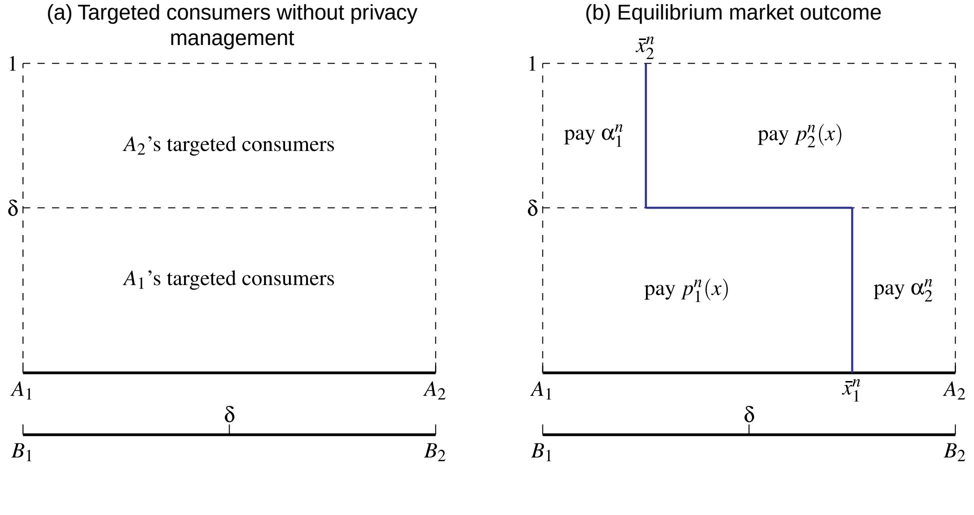
$$p_1^n(x) = \begin{cases} 3t/2 - 2tx & \text{if } x \leqslant \bar{x}_1^n, \\ 0 & \text{if } x \geqslant \bar{x}_1^n, \end{cases}$$

$$p_2^n(x) = \begin{cases} 0 & \text{if } x \leqslant \bar{x}_2^n, \\ 2tx - t/2 & \text{if } x \geqslant \bar{x}_2^n. \end{cases}$$

Figure 1(b) illustrates the equilibrium outcome in market A. Therefore, given $\delta \in [0, 1]$ in market B, the two firms' profits in market A are $\pi_{A_1}^n = t(2+7\delta)/16$ and $\pi_{A_2}^n = t(9-7\delta)/16$.

We now focus on the equilibrium analysis of market B. In market B, a consumer considers not only her utility $v_B - ky_i - \beta_i$, where $y_i = 1 - y_j$, offered by firm B_i, but also her expected utility in the subsequent market A after purchasing from firm B_i. Specifically, the expected surplus of firm B_1's consumers in market A,

Figure 1. (Color online) Market Structure and Equilibrium Under No Data Management



denoted as $E[CS_{B_1}]$, is:

$$E[CS_{B_1}] = \int_0^{\bar{x}_1^n} (v_A - p_1^n(x) - tx)dx + \int_{\bar{x}_1^n}^1 (v_A - \alpha_2^n - t(1-x))dx = v_A - t. \quad (2)$$

Similarly, $E[CS_{B_2}] = v_A - t$. The indifferent consumer δ in market B is determined by

$$v_B - \beta_1 - k\delta + E[CS_{B_1}] = v_B - \beta_2 - k(1-\delta) + E[CS_{B_2}], \quad (3)$$

resulting in $\delta = (k + \beta_2 - \beta_1)/(2k)$, consistent with the standard Hotelling model. The profits of firms B_1 and B_2 are $\pi_{B_1}^n = \beta_1\delta$ and $\pi_{B_2}^n = \beta_2(1-\delta)$, respectively. Each firm i decides its uniform price β_i to maximize two-market profits: $\Pi_i^n = \pi_{A_i}^n + \pi_{B_i}^n$ $(i = 1, 2)$.

The equilibrium uniform prices in market B are $\beta_i^n = (16k - 7t)/16$ $(i = 1, 2)$, leading to $\delta^n = 1/2$. We have that $\pi_{B_i}^n = (16k - 7t)/32$ and $\pi_{A_i}^n = 11t/32$. Their equilibrium profits are $\Pi_i^n = (t + 4k)/8$. Increased product differentiation in market B (i.e., a larger k) consistently enhances firms' profits. In contrast, increased differentiation in market A (i.e., a larger t) has opposing effects: it increases firm A_i's profit but decreases firm B_i's. A higher t intensifies competition in market B because firms compete more aggressively for consumers, which can, in turn, translate into higher firm profits in market A. Overall, a larger t results in greater total profit across both markets.

Remark 1. Comparison with BBPD literature.

Section 4.1 extends Choe et al. (2018) by introducing independent consumer preferences across markets, following the approach of Shin and Sudhir (2010). Compared with Choe et al. (2018), our model yields higher firm profits and lower consumer surplus in each market due to easier poaching and higher prices. Moreover, these results differ from those of Shin and Sudhir (2010), who find that introducing independent preferences into Fudenberg and Tirole (2000) intensifies competition in the first period but weakens it in the second period.

4.2. Equilibrium Under Data Management

4.2.1. Equilibrium Analysis in Market *A*. Consumers can decide whether to erase their data from firm B_i's datasets. These decisions are based on their anticipated prices in market A, specifically α_1^a, α_2^a, $p_1^a(x)$, and $p_2^a(x)$, where the superscript a indicates anticipation. We focus on the case where firms in market A employ pure strategies in uniform pricing. Lemma 1 characterizes the types of consumers who choose to erase their data.

Lemma 1. *Given consumers' price anticipations in market A, firm B_1's consumers erase data if and only if their locations in market A are on $[0, \tilde{x}_1]$, and firm B_2's consumers erase data if and only if their locations in market A are on $[\tilde{x}_2, 1]$, in which*

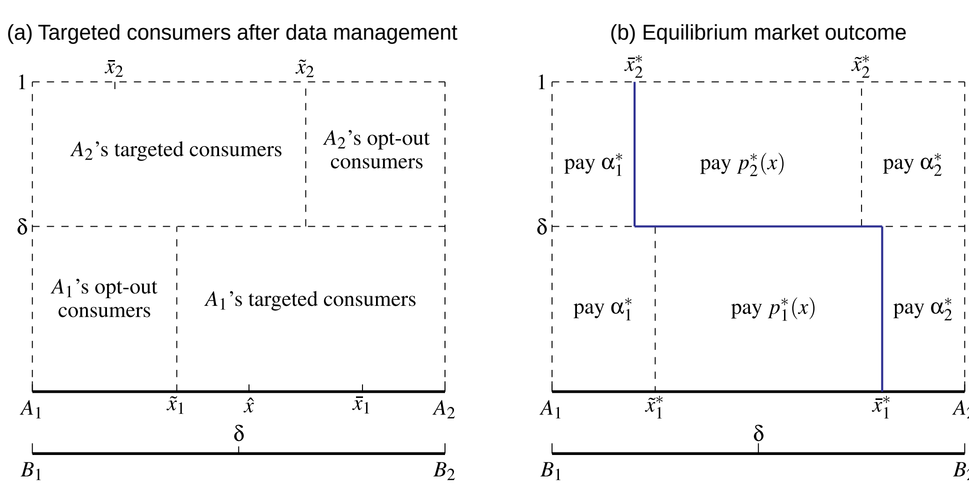
$$\tilde{x}_1 = \frac{1}{2} + \frac{\alpha_2^a - \alpha_1^a}{2t} - \frac{\varepsilon}{2t} \text{ and } \tilde{x}_2 = \frac{1}{2} + \frac{\alpha_2^a - \alpha_1^a}{2t} + \frac{\varepsilon}{2t}. \quad (4)$$

Lemma 1 states that consumers who match well with the firm that has collected their data choose to erase data in market A. They do so because they anticipate very high personalized prices in the case of no data erasure, preferring to enjoy lower uniform prices instead. Figure 2(a) shows the locations of these opt-out consumers.[16]

After data erasure, firm A_i's untargeted consumers are those who have purchased from firm B_j and opt-out consumers, who have purchased from firm B_i and have erased data. Firm A_i's targeted consumers are firm B_i's consumers who choose to opt in.

We analyze the equilibrium prices in market A after data erasure (i.e., $\tilde{x}_1$ and $\tilde{x}_2$ have been determined).

Figure 2. (Color online) Market Structure and Equilibrium Under Data Management



Firm A_1's opt-out consumers purchase from firm A_1 if and only if $x < \hat{x} = \frac{1}{2} + \frac{\alpha_2 - \alpha_1}{2t}$. Because opt-out consumers rationally anticipate equilibrium prices (i.e., α_i^a equals the equilibrium uniform price), inequality $\tilde{x}_1 < \hat{x} < \tilde{x}_2$ should hold.

We formulate the objectives of firms A_1 and A_2 to derive their optimal uniform prices. Firm A_1 wins the rival's targeted consumers $x < \max\{\bar{x}_2, 0\}$ with uniform price α_1, in which $\bar{x}_2 = 1/2 - \alpha_1/(2t)$. When $\alpha_1 \geqslant t$, firm A_1 forgoes poaching the rival's targeted consumers. Therefore, firm A_1's profit from its uniform price α_1 is

$$\begin{cases} \alpha_1\left[(1-\delta)\left(\dfrac{1}{2}-\dfrac{\alpha_1}{2t}\right)+\delta\tilde{x}_1\right] & \text{when } \alpha_1 \leqslant t, \\ \alpha_1\delta\tilde{x}_1 & \text{when } \alpha_1 \geqslant t \text{ and } \hat{x} \geqslant \tilde{x}_1. \end{cases}$$

Given the determined $\tilde{x}_1$, firm A_1's local optimal price for the second case of the above profit function is α_1 such that $\hat{x} = \tilde{x}_1$, which violates the definition of $\tilde{x}_1$ in (4) because $\tilde{x}_1 = \hat{x} - \varepsilon/(2t)$. Therefore, only the first case can be sustainable in equilibrium.

Similarly, firm A_2 applies the uniform price α_2 to its opt-out consumers and the rival's targeted consumers on $[\bar{x}_1, 1]$, where $\bar{x}_1 = 1/2 + \alpha_2/(2t)$. Firm A_2's profit from its uniform price α_2 is

$$\begin{cases} \alpha_2\left[\delta\left(\dfrac{1}{2}-\dfrac{\alpha_2}{2t}\right)+(1-\delta)(1-\tilde{x}_2)\right] & \text{when } \alpha_2 \leqslant t, \\ \alpha_2(1-\delta)(1-\tilde{x}_2) & \text{when } \alpha_2 \geqslant t \text{ and } \hat{x} \leqslant \tilde{x}_2. \end{cases}$$

Only the first case can be sustainable in equilibrium.

Given $\tilde{x}_1$ and $\tilde{x}_2$, the two firms' optimal uniform prices are $\alpha_1 = t/2 + t\tilde{x}_1\delta/(1-\delta)$ and $\alpha_2 = t/2 + t(1-\tilde{x}_2)(1-\delta)/\delta$. Since consumers rationally anticipate equilibrium prices (i.e., $\alpha_1^a = \alpha_1$ and $\alpha_2^a = \alpha_2$), based on (4), we have that

$$\alpha_1^* = \frac{t}{2} + \delta(t-\varepsilon), \quad \alpha_2^* = \frac{t}{2} + (1-\delta)(t-\varepsilon),$$
$$\tilde{x}_1^* = \frac{(1-\delta)(t-\varepsilon)}{t}, \quad \tilde{x}_2^* = 1 - \frac{\delta(t-\varepsilon)}{t}. \tag{5}$$

Compared with uniform prices in (1) under no data management, the second terms of α_i^* in (5) represent the increases in uniform prices.

Each firm's optimal personalized price at x is the rival's uniform price plus its relative advantage. Thus,

$$p_1^*(x) = \alpha_2^* + t(1-2x) \quad \text{and} \quad p_2^*(x) = \alpha_1^* + t(2x-1). \tag{6}$$

Note that a firm cannot win its targeted consumers who strongly prefer the rival's product, even if its personalized price is zero. Concretely, firm A_1 loses its targeted consumers on $[\bar{x}_1^*, 1]$, while firm A_2 loses its targeted consumers on $[0, \bar{x}_2^*]$, where

$$\bar{x}_1^* = \frac{3}{4} + \frac{(1-\delta)(t-\varepsilon)}{2t} \quad \text{and} \quad \bar{x}_2^* = \frac{1}{4} - \frac{\delta(t-\varepsilon)}{2t}.$$

Figure 2(b) illustrates the equilibrium outcome in market A. Therefore, given $\delta \in [0,1]$ in market B, the two firms' profits in market A are

$$\pi_{A1}^* = \delta\left[\alpha_1^*\tilde{x}_1^* + \int_{\tilde{x}_1^*}^{\bar{x}_1^*} p_1^*(x)dx\right] + (1-\delta)\alpha_1^*\bar{x}_2^*,$$
$$\pi_{A2}^* = (1-\delta)\left[\alpha_2^*(1-\tilde{x}_2^*) + \int_{\bar{x}_2^*}^{\tilde{x}_2^*} p_2^*(x)dx\right] + \delta\alpha_2^*(1-\bar{x}_1^*).$$

Compared with no data management, we have $\bar{x}_1^* > \bar{x}_1^n$ and $\bar{x}_2^* < \bar{x}_1^n$, implying that more consumers purchase from their less preferred firm. The underlying intuition is that higher uniform prices make poaching less likely to occur. Consequently, the number of poached consumers decreases.

In comparison with benchmark prices, Proposition 1 summarizes how prices in market A change under data management.

Proposition 1. *Under data management, firms in market A charge higher uniform and personalized prices compared with no data management: $\alpha_i^* > \alpha_i^n$ and $p_i^*(x) > p_i^n(x)$, for $i \in \{1,2\}$. Furthermore, the increases in a firm's uniform price and the rival's personalized prices are more significant*

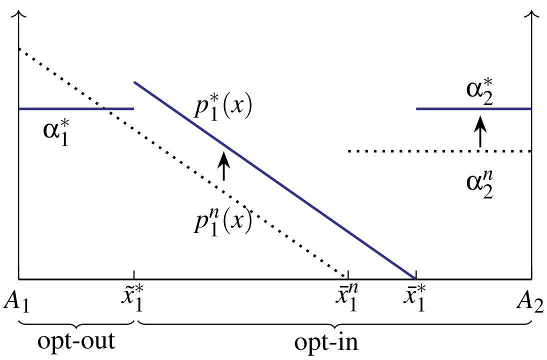
(i) *when the data erasure cost ε declines, or*

(ii) *when the firm's market share in market B (i.e., δ for firm 1 and $1-\delta$ for firm 2) expands.*

The intuition behind Proposition 1 is as follows. Firm A_i uses its uniform price to serve its own opt-out consumers—those who strongly prefer the firm—and to poach the rival's targeted consumers. Its uniform price needs to compete with the rival's uniform price α_j for the first group of consumers and compete with the rival's personalized prices $p_j(x)$ for the second group. Since the rival's uniform price α_j is higher than the flexible personalized price $p_j(x)$ for marginal consumers, opt-out consumers exhibit lower price sensitivity for the firm's uniform price. These consumers also have a higher willingness to pay, which boosts demand for the firm's uniform price. The combination of low price sensitivity and high demand allows firm A_i to be a "fat cat" in uniform price competition (Fudenberg and Tirole 1984, Chen et al. 2001). The fat cat effect is stronger for firm A_i when the cost of data erasure is lower or when its sister firm captures a larger market share in market B. As illustrated in Figure 3, firm A_2's uniform price rises from α_2^n under no data management to α_2^* under data management.[17]

Since each firm's personalized prices compete with the rival's uniform price, firms' personalized prices also increase under data management (see $p_1^*(x)$ and $p_1^n(x)$ in Figure 3). Corollary 1 summarizes the discussion of Proposition 1.

Figure 3. (Color online) Comparison of Equilibrium Prices Paid by A_1's Consumers



Corollary 1. *Opt-in consumers face negative externalities resulting from opt-out consumers through higher market prices, α_i^* and $p_i^*(x)$.*

We now examine price changes for opt-out consumers. Since each opt-out consumer downplays her impact on driving up market prices, those who benefit marginally from data management may find that the inflated uniform price α_i^* exceeds their benchmark personalized prices $p_i^n(x)$. For instance, as illustrated in Figure 3, firm A_1's opt-out consumers who are close to $\tilde{x}_1^*$ face a price α_1^* that is higher than $p_1^n(x)$.

Finally, to elucidate the impact of data management on a firm's profitability, we decompose $\pi_{A_1}^* - \pi_{A_1}^n$ into three distinct components:

$$\pi_{A_1}^* - \pi_{A_1}^n = \underbrace{\delta\left(\alpha_1^*\tilde{x}_1^* - \int_0^{\tilde{x}_1^*} p_1^n(x)dx\right)}_{\text{opt-out consumers}} + \underbrace{\delta\left(\int_{\tilde{x}_1^*}^{\bar{x}_1^*} p_1^*(x)dx - \int_{\tilde{x}_1^*}^{\bar{x}_1^n} p_1^n(x)dx\right)}_{\text{opt-in consumers}} + \underbrace{(1-\delta)(\alpha_1^*\bar{x}_2^* - \alpha_1^n\bar{x}_2^n)}_{\text{poached consumers}}. \quad (7)$$

Under a fixed value of δ, we can discuss the property of (7). The first component in (7) is negative because of the firm's inability to efficiently extract surplus from high-willingness-to-pay consumers. The second component in (7) is positive. Proposition 1 explains the reason: as the rival charges a higher uniform price under data management, firm A_i increases its personalized prices and more of its targeted consumers purchase under the increased personalized prices. The third component in (7) is negative because under data management, firm A_i increases its uniform price due to opt-out consumers, resulting in inefficient consumer poaching compared with the benchmark scenario where the uniform price is solely designed for consumer poaching.

We mention the validity of the equilibrium outcomes in market A. First, the equilibrium cutoffs $\tilde{x}_i^*$ and $\bar{x}_i^*$, as depicted in Figure 2(b), are confined within the range of zero to one, which holds under $t/2 \leqslant \varepsilon \leqslant t$. Second, given the rival's price, neither firm should benefit by deviating to a higher uniform price for its opt-out consumers and forego poaching the rival's targeted consumers, nor by lowering the uniform price to attract the rival's opt-out consumers. This translates to the condition $\sqrt{3}t/2 \leqslant \varepsilon \leqslant t$. Therefore, we impose the assumption $\sqrt{3}t/2 \leqslant \varepsilon \leqslant t$ to ensure that the outcomes in (5) hold for any $\delta \in [0,1]$.[18]

4.2.2. Equilibrium Analysis in Market *B*. We turn our attention to the equilibrium analysis of market B, the data collection market. The indifferent consumer δ in market B is determined in the same way as in (3), in which $E[CS_{B_i}]$ is the ex-ante expected surplus of firm B_i's consumer in market A. We have

$$E[CS_{B_1}] = \int_0^{\tilde{x}_1^*} (v_A - \alpha_1^* - tx - \varepsilon)dx + \int_{\tilde{x}_1^*}^{\bar{x}_1^*} (v_A - p_1^*(x) - tx)dx + \int_{\bar{x}_1^*}^{1} (v_A - \alpha_2^* - t(1-x))dx. \quad (8)$$

The expression for $E[CS_{B_2}]$ can be derived in a similar manner. We calculate $E[CS_{B_1}] - E[CS_{B_2}]$ and obtain:

$$E[CS_{B_1}] - E[CS_{B_2}] = \frac{(\delta - (1-\delta))(t-\varepsilon)\varepsilon}{t} > 0 \quad \text{if and only if } \delta > 1-\delta. \quad (9)$$

Equation (9) highlights an important implication: firm B_i's consumers will experience relatively higher expected surpluses if its market share in market B exceeds the rival's market share. This relationship between $E[CS_{B_i}]$ and its market share stems from the results described in Proposition 1. For example, as firm B_1's market share, δ, increases, firm A_1's uniform price increases, while its personalized prices decrease. Since consumers are more likely to remain as opt-in consumers, they can expect to retain a higher expected surplus from firm A_1. In contrast, firm A_2's personalized prices increase, implying lower expected consumer surplus from firm A_2. Consequently, a larger δ leads to a higher $E[CS_{B_1}]$ than $E[CS_{B_2}]$.

We derive equilibrium prices in market B. The indifferent consumer in market B is $\delta = 1/2 + t(\beta_2 - \beta_1)/[2(tk + \varepsilon^2 - t\varepsilon)]$. Compared with the indifferent consumer $\delta = 1/2 + (\beta_2 - \beta_1)/(2k)$ in Section 4.1, firm B_i's demand is more price elastic here because $t/[2(tk + \varepsilon^2 - t\varepsilon)] > 1/(2k)$. Firms' profits in market B are $\pi_{B_1} = \beta_1\delta$ and $\pi_{B_2} = \beta_2(1-\delta)$. Firm i sets its price β_i to maximize two-market profits: $\Pi_i = \pi_{A_i} + \pi_{B_i}$. Equilibrium prices in market B are

$$\beta_1^* = \beta_2^* = \frac{16tk - 8t^2 + 15\varepsilon^2 - 14t\varepsilon}{16t}, \quad (10)$$

implying that $\delta^* = 1/2$. Firm B_i's equilibrium profits are equal to $\beta_i^*/2$.

The equilibrium outcomes in market A are determined by replacing δ with $1/2$:

$$\tilde{x}_1^* = \frac{1}{2} - \frac{\varepsilon}{2t}, \quad \tilde{x}_2^* = \frac{1}{2} + \frac{\varepsilon}{2t}, \quad \bar{x}_1^* = 1 - \frac{\varepsilon}{4t},$$

$$\bar{x}_2^* = \frac{\varepsilon}{4t}, \quad \alpha_1^* = \alpha_2^* = t - \frac{\varepsilon}{2},$$

$$p_1^*(x) = \begin{cases} 2t(1-x) - \varepsilon/2 & \text{when } \tilde{x}_1^* < x < \bar{x}_1^*, \\ 0 & \text{when } \bar{x}_1^* \leqslant x, \end{cases}$$

$$p_2^*(x) = \begin{cases} 0 & \text{when } x \leqslant \bar{x}_2^*, \\ 2tx - \varepsilon/2 & \text{when } \bar{x}_2^* < x < \tilde{x}_2^*. \end{cases}$$

The equilibrium profits of firms A_1 and A_2 are $\pi_{A_1}^* = \pi_{A_2}^* = (12t^2 + 3\varepsilon^2 - 4t\varepsilon)/(32t)$. The equilibrium profits of firms 1 and 2 are $\Pi_1^* = \Pi_2^* = (8tk + 2t^2 + 9\varepsilon^2 - 9t\varepsilon)/(16t)$.

Proposition 2. *Compared with the benchmark, in the equilibrium under data management,*

(i) *price competition in market B intensifies:* $\beta_i^* < \beta_i^n$, *for* $i \in \{1, 2\}$;

(ii) *both firm A_i and firm B_i earn lower profits, resulting in reduced overall profits for firm i across both markets.*

The first result of Proposition 2 suggests that *firms compete more aggressively* to win consumers and collect their data in market B when consumers can erase their data. The rationale behind this proposition is as follows. Consider firm 1 as an example. An increase in δ has two major effects: (i) a decrease in the ratio of opt-out consumers among firm A_1's consumers, $\tilde{x}_1^*$ (see (5)), thus enlarging the range to which A_1 applies the elevated personalized prices; and (ii) an increase in its uniform price, thus partially offsetting the profit loss from opt-out consumers. Due to these two effects, firm i's profit in market A becomes more sensitive to its market share in market B, enhancing its incentives to increase market share in market B relative to the benchmark. Furthermore, as shown in (9), firm B_i with a larger market share is more likely to attract consumers, which further intensifies price competition in market B for market share.

The second result of Proposition 2 shows that data management decreases firms' profits in both markets. Firm A_i experiences lower profits due to its inability to effectively extract surplus from high-willingness-to-pay consumers and its inefficiency in poaching the rival firm's consumers, even though it earns higher profits from remaining opt-in consumers (see the discussion right after (7)). Firm B_i's lower profit is driven by the intensified price competition in the market.

Increased product differentiation in market B (i.e., a larger k) consistently enhances firms' profits. In contrast, increased differentiation in market A (i.e., a larger t) has opposing effects, increasing firm A_i's profit while decreasing firm B_i's. In contrast to the benchmark scenario, firms' overall two-market profit declines with t. This occurs because, as t rises, the link between market B market share and firm A_i's profit strengthens, leading to a substantial increase in competition within market B.

Remark 2. The role of $E[CS_{B_i}]$.

The difference in expected surplus $E[CS_{B_1}] - E[CS_{B_2}]$ affects the location of the indifferent consumer in market B. This effect vanishes if $E[CS_{B_1}] = E[CS_{B_2}]$ regardless of δ, as seen in Section 4.1, or if consumers do not take into account the expected surplus in market A (i.e., myopic consumers). In such scenarios, consumers in market B consider only their utility in that market. Nevertheless, even in the case of myopic consumers, Proposition 2(i) still holds because the two major underlying effects are independent of $E[CS_{B_i}]$, which acts as a third effect that intensifies price competition in market B. Consequently, Proposition 2(ii) continues to hold.

4.3. Impacts of Data Management on Welfare

Proposition 3 summarizes how consumer surplus changes across the two markets under data management.

Proposition 3. *Compared with no data management, the availability of data management leads to seesaw changes in consumer surplus across market A and market B. In market A, the consumer surplus of opt-out consumers remains unchanged, while opt-in consumers experience a decrease in their surplus, leading to a reduced overall consumer surplus in market A. In market B, consumer surplus increases.*

This seesaw change in consumer surplus results from intensified price competition in market B and increased prices for the majority of consumers in market A (see Figure 3). Aggregated consumer surplus across both markets is higher under data management.

Compared with no data management, social welfare in market A declines under data management. This is because a greater number of consumers mismatch with the less preferred firm, as indicated by $\bar{x}_1^n < \bar{x}_1^*$ and $\bar{x}_2^n > \bar{x}_2^*$. Additionally, the costs of data management cause a deadweight loss, further reducing social welfare. In market B, social welfare remains unchanged due to the equal market split. In summary, from Propositions 2 and 3, the availability of data management fosters a redistribution of surplus among market participants and across different markets.

Corollary 2. *In the equilibrium under data management, consumer surplus in market A increases with* ε, *while consumer surplus in market B decreases with* ε.

Corollary 2 summarizes how the data management cost, which represents a deadweight loss, impacts consumer surplus. An increase in ε discourages consumers from erasing data, thereby weakening the effects described in Section 4.2 that adversely affect consumer surplus in market A. Consequently, price competition in market B is mitigated, negatively affecting consumer surplus.

5. Enhanced Models: The Portfolio of Data Management Rights

Sections 5.1, 5.2, and 5.4 discuss three key data management policies: consent to data collection, data portability, and default opt-in/opt-out. Section 5.3 enhances the baseline model by incorporating consumers' multidimensional characteristics. For simplicity, we assume that the transportation costs in both markets are equal, denoted as $t = k$, thereby reducing the number of parameters.

5.1. Consumers Can Opt Out of Data Collection

We consider the scenario in which consumers can simultaneously decide whether to opt out of data collection in market B. If a consumer opts out, she receives utility $u = v_B - \phi - ky_i - \beta_i$ (where $y_i = 1 - y_j$) from firm B_i. The parameter $\phi \geqslant 0$ represents the disutility from malfunctioning data-driven features, which can vary among consumers.[19]

Given a consumer's opt-out disutility ϕ, we explain how she chooses between opting in and opting out in market B. For each option, she evaluates the joint expected utility across the two markets. Given the anticipated equilibrium in market A, consumer y_i of firm B_i opts in to data collection if

$$\underbrace{(v_B - ky_i - \beta_i) + E[CS^{in}_{B_i}]}_{\text{expected opt-in utility in two markets}} \geqslant \underbrace{(v_B - ky_i - \beta_i - \phi) + E[CS^{out}_{B_i}]}_{\text{expected opt-out utility in two markets}}. \quad (11)$$

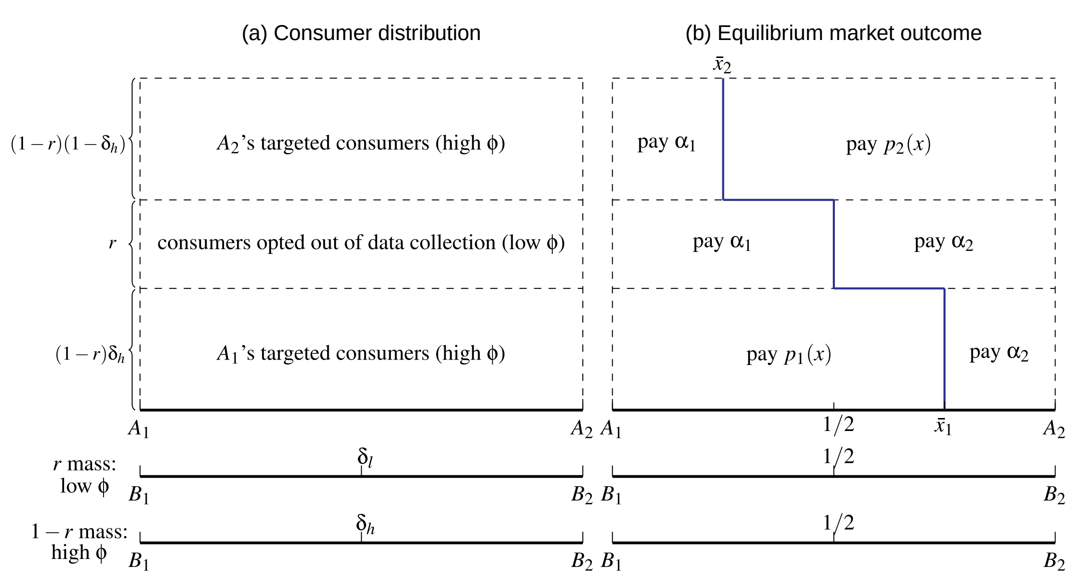
Otherwise, she opts out of data collection. Here, $E[CS^{in}_{B_i}]$ and $E[CS^{out}_{B_i}]$ denote her expected opt-in and opt-out utilities in market A, respectively.

Our analysis then proceeds under two scenarios, depending on whether opt-in consumers can erase data in market A. First, data cannot be erased before its application, as in Section 4.1. We refer to this as the "*collection opt-out*" scenario. Second, data can be erased prior to application, as in Section 4.2. We call this the "*full control*" scenario.

To analyze consumers' opt-out behavior, we initially consider two polar cases. First, if consumers' heterogeneous ϕ values are uniformly high, such that (11) holds for all consumers, they all choose to opt in to data collection, resulting in an analysis identical to Section 4. Second, if their heterogeneous ϕ values are uniformly low, such that (11) fails for all consumers, they all opt out, leading to the standard Hotelling results in both markets: $\alpha^H_i = t$, $\beta^H_i = k$, for $i \in \{1,2\}$, where the superscript H indicates the Hotelling outcome.[20] The Hotelling outcome yields the highest uniform prices in both markets, resulting in the highest firm profits and lowest consumer surplus compared with the equilibrium outcomes in Sections 4.1 and 4.2.

We now consider a scenario in which consumers incur highly heterogeneous opt-out disutilities in market B, leading to the coexistence of both opt-in and opt-out behaviors. Specifically, an exogenous fraction $r \in [0,1]$ of consumers face negligible disutility and choose to opt out of data collection, while the remaining $1 - r$ consumers face substantially higher disutility and therefore opt in.[21] Figure 4(a) illustrates the distribution of consumers in the scenario of collection opt-out: consumers can opt-out in market B but opt-in consumers cannot erase data before its application.

Figure 4. (Color online) Market Structure and Equilibrium Under Collection Opt-Out



Firm A_1 can target consumers with a mass of $(1-r)\delta_h$, where the subscript h indicates high ϕ, while firm A_2 can target consumers with a mass of $(1-r)(1-\delta_h)$.

The presence of collection opt-out consumers influences uniform and personalized prices (see Figure 4(b)). A firm's uniform price not only aims to poach the rival's targeted consumers, as described in Section 4.1, but also involves a weaker form of Hotelling competition (see the regions labeled "pay α_i" in Figure 4(b)). Consequently, the equilibrium uniform prices lie between $\alpha_i^H = t$ and $\alpha_i^n = t/2$. An increase in the number of opt-out consumers (i.e., a higher r) pushes the equilibrium uniform prices closer to the Hotelling price α_i^H. Consequently, firm A_i's profit consistently increases with r, as Hotelling competition reflects the weakest competitive intensity. Meanwhile, consumer surplus always decreases as r grows.

In market B, competition depends on the sensitivity of firm A_i's profit to firm B_i's market share. Both "higher personalized prices" and "a larger opt-in population" in market A increase this sensitivity, intensifying competition in market B. Increasing the number of opt-out consumers (r) creates a trade-off between these two effects. For small r, increasing r raises personalized prices (through higher uniform prices) without significantly reducing the number of opt-in consumers. This heightened sensitivity intensifies competition in market B. Conversely, a large r implies a small opt-in population, which reduces sensitivity and weakens competition in market B. Overall, the equilibrium prices in market B exhibit a U-shaped pattern in r.

In the full control scenario, the equilibrium properties established under collection opt-out continue to hold. An additional feature is that collection opt-in consumers can erase data in market A, as characterized in Lemma 1. The effects of these application opt-out consumers align closely with the observations in Section 4.2: they lead to higher prices in market A and intensified competition in market B. Proposition 4 summarizes the discussions when consumers can opt out of data collection.

Proposition 4. *When a fraction $r \in (0,1)$ of consumers face negligible disutility and choose to opt out of data collection, the equilibrium outcomes under collection opt-out (resp., full control) blend the Hotelling outcome with the results presented in Section 4.1 (resp., Section 4.2). In both cases:*

(i) *firm A_i's profit increases monotonically with the number of consumers opting out of data collection (i.e., r), whereas firm B_i's profit is U-shaped in r, exceeding its profit under no opting out of data collection (i.e., $r = 0$) once r exceeds a certain threshold $\hat{r}$ (which differs between collection opt-out and full control scenarios);*

(ii) *consumer surplus in market A decreases monotonically with r, while in market B it is inversely U-shaped in r, falling below the surplus under no opting out of data collection (i.e., $r = 0$) when $r > \hat{r}$.*

Compared with the scenario without data collection opt-out in Section 4, Proposition 4 states that allowing consumers to opt out of data collection can benefit firms and hurt consumers, particularly when the number of opting-out consumers (i.e., r) is not small.

Combining Propositions 2 and 4, we find that data management disrupts the typical seesaw pattern of firm profits observed in BBPD studies (e.g., Fudenberg and Tirole 2000, Shin and Sudhir 2010). Specifically, with data erasure in the application market, firms' profits decline in both markets; however, when consumers can opt out of data collection, firms' profits increase in both markets if $r > \hat{r}$.

Moreover, we find that where consumers manage their personal data affects competition across markets quite differently. An opt-out option in the data collection market raises prices in both markets whenever $r > \hat{r}$, benefiting firms but harming consumers. In contrast, an opt-out option in the data application market (data erasure) tends to increase prices there while lowering prices in the data collection market, which harms firms but can enhance overall consumer surplus across both markets.

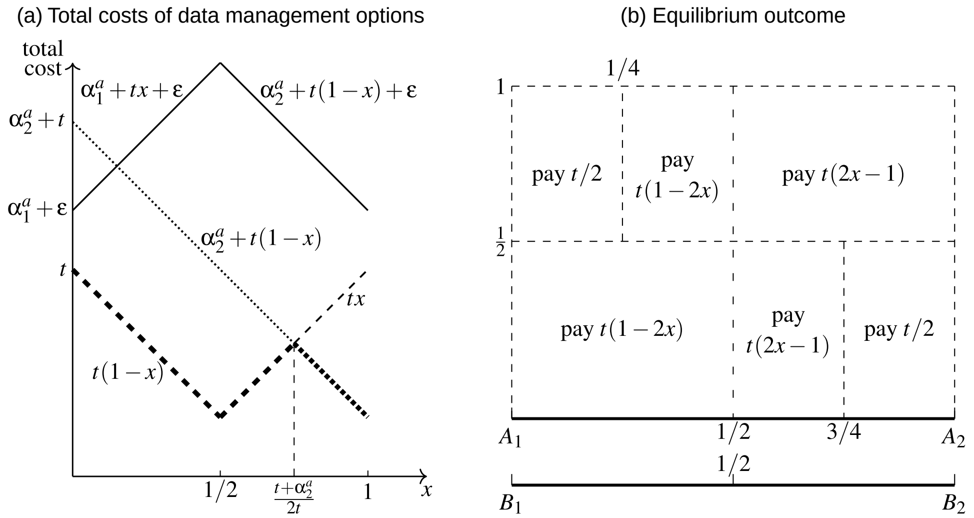
5.2. Data Portability

We explore an extended scenario where consumers have the options of data erasure and data portability for data management. Data portability allows targeted consumers of firm B_i to request the transfer of their data to the rival firm B_j at no cost.[22] The equilibrium without data management in this extension is the same as that in Section 4.1.

We illustrate personal data management strategies using firm A_1's consumer located at x. A similar discussion applies to firm A_2's consumers. Firm A_1's consumer at location x has four data management options to minimize her anticipated total cost, which includes price, transportation costs, and the data erasure cost ε (if applicable). First, if the consumer erases her data, her total cost is either $\alpha_1^a + tx + \varepsilon$ or $\alpha_2^a + t(1-x) + \varepsilon$ (solid lines in Figure 5(a)). Second, if consumer x chooses data portability and does not erase her data from A_1's datasets, she will be targeted by both firms with personalized prices. Her total cost is $p_1^a(x) + tx = t(1-x)$ when $x \leqslant 1/2$ and $p_2^a(x) + t(1-x) = tx$ when $x \geqslant 1/2$ (dashed lines). Third, if consumer x chooses data portability and erases her data from firm A_1's datasets, she will be targeted only by firm A_2 and will incur a total cost of $\alpha_1^a + tx + \varepsilon$ (the upward-sloping solid line).[23] Fourth, if she does not manage her data, her total cost is $\alpha_2^a + t(1-x)$ (the dotted line).

Figure 5(b) illustrates the equilibrium data management choices made by firm A_i's targeted consumers.

Figure 5. Total Costs of Data Management Options and Equilibrium Outcome



Firm A_1's targeted consumers in the interval $[0,3/4]$ utilize data portability to make their data accessible to both firms, triggering intense competition between personalized prices. Consequently, consumers in $[0,1/2]$ purchase from firm A_1 under personalized prices $p_1(x) = t(1-2x)$, while those in $[1/2,3/4]$ purchase from firm A_2 under personalized prices $p_2(x) = t(2x-1)$. Consumers in $[3/4,1]$ do not manage personal data and purchase from A_2 at the uniform price $\alpha_2 = t/2$. Firm A_2's targeted consumers exhibit symmetrical behavior.

The equilibrium shown in Figure 5(b) requires $\varepsilon \geqslant t/2$, which covers the range of ε in the baseline model. Due to the high data erasure cost, even consumers who strongly prefer their targeted firm do not choose data erasure, as the lower personalized prices resulting from data portability are more beneficial. However, if $\varepsilon < t/2$, these consumers prefer to erase data rather than opt for data portability.[24] The key results and underlying intuition remain unchanged.

Proposition 5. *Suppose consumers in market A have access to data erasure and data portability for data management. When $\varepsilon \in [t/2, t]$, compared with no data management,*

(i) *uniform prices in market A remain unchanged, but personalized price competition intensifies;*

(ii) *firms in market A earn lower profits, and those in market B earn higher profits, resulting in higher two-market profits for firms 1 and 2;*

(iii) *consumer surplus increases in market A and decreases in market B.*

We explain the intuition behind Proposition 5 point to point. First, firm A_i's uniform price is used solely to poach the rival's targeted consumers, as in the benchmark of no data management in Section 4.1. Second, lower profitability of personalized pricing reduces the sensitivity of profits in market A to market shares in market B, thereby weakening competition in market B. As a result, profits increase. Third, intensified competition in A leads to higher consumer surplus in market A. Conversely, consumers become worse off in market B due to reduced competition. Aggregated consumer surplus across both markets is lower under data management. Additionally, data portability maximizes social welfare by enabling consumers to consistently purchase from their preferred firms in each market.

Table 1 shows that firms' two-market profits and total consumer surplus can vary nonmonotonically with the scope of consumers' data rights. The benchmark for the first and second columns is the scenario without data management (Section 4.1), and that for the third and fourth columns is Section 4.2 (E in the first column). The additional data rights in the second and fourth columns—data portability and opt-out in market B ($r > \hat{r}$)—have a nonmonotonic influence. In contrast, the additional data right in the third

Table 1. The Effects of Data Rights Scope on Firm Profits and Consumer Welfare

	Data erasure in market A (denote it as E)	E *and* data portability	Given E, opt-out in market B (if $\underline{r < \hat{r}}$)	Given E, opt-out in market B (if $\underline{r > \hat{r}}$)
Firm's two-market profits	⇓	⇑	⇓	⇑
Total consumer surplus	⇑	⇓	⇑	⇓

column—opt-out in market B ($r < \hat{r}$)—further amplifies the effect of E, strengthening benefits to consumers and harms to firms. Therefore, firms need to carefully consider the specific types of data rights consumers have because the impact can vary significantly.

5.3. Consumers with Horizontal and Vertical Characteristics

The model setting of this extension aligns with Section 4.2, with the notable exception of incorporating heterogeneous servicing costs for consumers across the two markets. Specifically, consumers have differing servicing costs: $c_B \in [0, \bar{c}_B]$ in market B and $c_A \in [0, \bar{c}_A]$ in market A, which are their private information.[25] At each point on the Hotelling line of market l ($l \in \{A, B\}$), service costs c_l are uniformly distributed over the interval $[0, \bar{c}_l]$. Thus, consumers are uniformly distributed across the space $[0,1] \times [0, \bar{c}_l]$ in market $l \in \{A, B\}$. Figure 6 illustrates the consumer distributions in both markets.

For targeted consumers, firms in market A accurately identify both product preferences x and servicing costs c_A, enabling them to charge personalized prices $p_i(x, c_A)$. For untargeted consumers, firms charge uniform prices α_i.

5.3.1. Benchmark: No Data Management. The equilibrium analysis in market A is analogous to that in Section 4.1. However, firm A_i's marginal consumer poached by the rival, $\bar{x}_i$, varies with c_A, as in Figure 6. This is because the firm's lowest $p_i^n(x, c_A)$ to retain a consumer is her servicing cost c_A, making high-cost consumers easier to be poached by the rival's uniform price. Then, the rival firm A_j adjusts its uniform price to $\alpha_j^n = (t + \bar{c}_A)/2$, accounting for the average cost of the poached consumers $\bar{c}_A/2$. Since firm A_i's personalized prices $p_i^n(x, c_A)$ depend only on location x and the rival's uniform price, its optimal personalized prices are independent of c_A. In contrast, the range it can offer $p_i^n(x, c_A)$ depends on c_A: $\overline{x}_1^n = 3/4 + (\overline{c}_A - 2c_A)/(4t)$ and $\overline{x}_2^n = 1/4 - (\overline{c}_A - 2c_A)/(4t)$.

In market B, the indifferent consumer δ is determined in the same manner as in (3), independent of c_B. The subsequent analysis mirrors Section 4.1 and can be found in Section 6 of Online Appendix A.

5.3.2. Data Management. As shown in Figure 7, the marginal consumers who choose to erase data, denoted as $\tilde{x}_i$, are consistent with Lemma 1, independent of c_A. Consequently, the equilibrium uniform prices in market A rise relative to the benchmark: $\alpha_1^* - \alpha_1^n = \delta(t - \varepsilon)$ and $\alpha_2^* - \alpha_2^n = (1 - \delta)(t - \varepsilon)$. Thus, firms' personalized prices also increase. As each firm's profit in market A becomes more sensitive to its sister firm's market share in market B under data management, price competition in market B intensifies.

Our analysis demonstrates that the impacts of data management align closely with Section 4.2, even with the introduction of heterogeneous servicing costs as a vertical dimension. Proposition 6 summarizes the result.

Proposition 6. *Regardless of whether consumers can manage data, the profit of firm A_i increases with the upper bound of the servicing cost $\bar{c}_A$ if its customer base (i.e., δ for firm A_1 and $1 - \delta$ for firm A_2) exceeds $2/3$. Consequently, a higher $\bar{c}_A$ motivates firms to compete more aggressively in market B. In equilibrium, firms evenly divide customers (i.e., $\delta = 1/2$), leading to all firms' profits decrease with servicing costs.*

The intuition behind Proposition 6 is as follows. Servicing costs in market A motivate firms to charge higher uniform and personalized prices. These higher prices in market A have contrasting effects on the firm's profit. On one hand, the increased uniform price implies that firm A_i cannot efficiently poach the rival's targeted consumers, resulting in profit losses. On the other hand, firm A_i can impose higher personalized prices on more of its targeted consumers,

Figure 6. (Color online) Benchmark Equilibrium

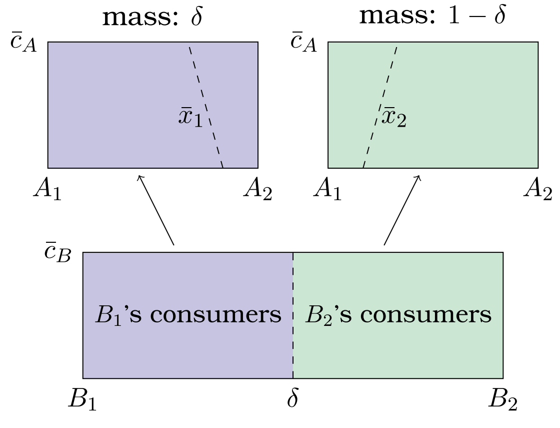
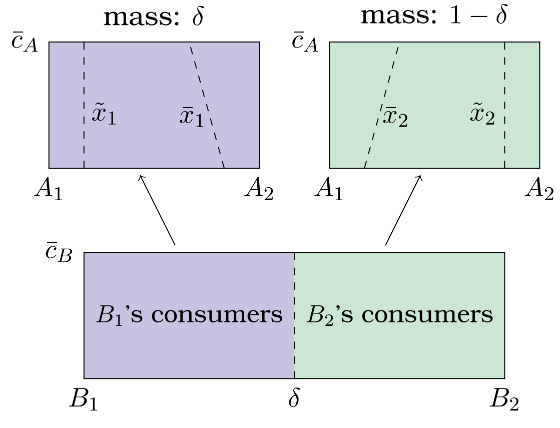


Figure 7. (Color online) Data Management Equilibrium



leading to profit gains. Therefore, firm A_i's profit increases with servicing costs whenever the beneficial effect is strong enough, which occurs when firm A_i's customer base is large. This mechanism directly motivates firms to compete more aggressively in market B for market shares as servicing costs rise. However, in equilibrium, firms equally split market B, implying that neither firm gains sufficient market share. Consequently, all firms' profits decrease with servicing costs.

5.4. Data Ownership: Opt-In or Opt-Out as Default?

The baseline model implicitly assumes that firms own user data, as consumers are automatically opted in unless they actively choose otherwise. We now shift data ownership to consumers, establishing an opt-out default where consumers can choose to opt in to data application within market A; otherwise, firms cannot utilize consumer data, even if collected. In this scenario, both firms in market A offer a benefit $b \in [0, t]$ to incentivize consumers to consent to data utilization.[26] The game in market B remains the same as in the baseline model, while the game in market A proceeds as follows: (i) consumers simultaneously choose whether to opt in; (ii) the two firms simultaneously post their uniform prices, followed by offering personalized prices and delivering the promised benefit b to their opt-in consumers; (iii) after observing all available offers, consumers make their purchasing decisions.

In market A, firm A_1's consumers with a realized location $x > \tilde{x}_1 = 1/2 + (\alpha_2^a - \alpha_1^a)/(2t)$-$b/(2t)$ choose to opt in to enjoy low personalized prices and the opt-in benefit b (if positive). Similarly, firm A_2's consumers with a realized location $x < \tilde{x}_2 = 1/2 + (\alpha_2^a - \alpha_1^a)/(2t) + b/(2t)$ also choose to opt in. However, firm A_i's targeted consumers with nearby locations do not opt in due to concerns about high personalized prices. Consequently, firm A_i's segments of opt-in and opt-out consumers resemble those depicted in Figure 2(a), and the equilibrium analysis is quite similar. Nonetheless, firm A_i obtains lower profits compared with the baseline model in Section 4.2, as firms must pay per-consumer opt-in benefit b. The equilibrium prices in market B are given by

$$\beta_1^\star = \beta_2^\star = \frac{8t^2 + 15b^2 - 14tb}{16t} + b,$$

which is similar to expression (10) in the baseline model (under $t = k$), but with an additional term b. In other words, as firm A_i experiences reduced profits in market A due to the opt-in benefit b, firm B_i is less motivated to compete for consumers. Proposition 7 compares the equilibrium when $b = \varepsilon$ in this section with the equilibrium in Section 4.2.

Proposition 7. *Compared with the scenario where firms own data and consumers can erase it, shifting data property rights to consumers leads to diminished profits for firms in market A. However, it results in increased profits in market B and, consequently, higher two-market profits. Consumer surplus becomes higher in market A and lower in market B.*

Both the opt-in benefit b here and opt-out cost ε in the baseline model influence competition within the two markets. A crucial difference between them is that b represents a transfer to an opt-in consumer and does not constitute a social cost, whereas ε is a social cost incurred by an opt-out consumer. Consequently, shifting data property rights to consumers enhances consumer surplus in market A through accrued benefits and saved data management costs ε. However, it lowers consumer surplus in market B by dampening price competition.

5.5. Other Extended Models

This section provides a brief overview of several extended models, with detailed analyses presented in Online Appendix B.

5.5.1. The Interests of Firm A_2 and Firm B_2 Are Partially Aligned. We extend Section 4 with the following modification: firm B_2 is not a sister firm of A_2 within firm 2 but shares its collected consumer data with firm A_2 in exchange for a share $\lambda \in [0, 1]$ of firm A_2's profits as compensation, where λ captures the degree of alignment between firms A_2 and B_2. We find the impacts of data erasure are qualitatively similar to those in Section 4. Moreover, as λ increases, firm B_2 competes more aggressively to capture market share, resulting in intensified competition in market B and a lower value of δ. The effects of a lower δ on the outcomes in market A closely align with our observations in Section 4.

5.5.2. Perfect Preference Correlation Across Markets. This extension examines a scenario where some consumers exhibit perfectly correlated preferences across two markets, specifically where their locations satisfy $x = y$. Under data management, among these consumers, firm A_1's targeted customers with locations $x < \tilde{x}_1$ choose to erase their data, while firm A_2's customers with locations $x > \tilde{x}_2$ also erase their data to avoid high personalized prices. Consequently, compared with the no data management case, both uniform and personalized prices in market A increase. Moreover, due to the same intuition as Proposition 2(i), competition in market B intensifies. As a result, consumers become worse off in market A but better off in market B.

5.5.3. Data-Driven Product Personalization. Built on Section 4, this extension assumes that firm A_i can use data analytics to offer personalized products with a common additional value $\Delta v > 0$ to consumers who do not erase their data in market A. The effects of an increase in Δv on profits are mixed in markets A and B—basically positive and negative. In market A, an increase in Δv leads to higher personalized prices and a wider range of consumers purchasing personalized products, benefiting firms. However, in market B, anticipating such higher profits, an increase in Δv accelerates competition. Compared with the benchmark of no data management, the impacts of data management are highly consistent with those observed in Section 4.2.

5.5.4. Vertical Product Differentiation. Building on the works of Rhee and Thomadsen (2017) and Jing (2017), we examines a scenario in which A_1 and A_2 are vertically differentiated and consumers differ in their tastes $\theta \in [0, 1]$ for quality. Firm A_i accurately identifies the quality preferences of its targeted consumers and charge personalized prices $p_i(\theta)$, while the firm charges a uniform price α_i to untargeted consumers due to the lack of their preference information. When comparing the equilibrium under no data management to that under data management, the results are highly consistent with those of the baseline model presented in Section 4.2.

6. Firms' Proactive Strategies and Managerial Implications

This section explores marketing strategies that firms can utilize to mitigate the negative impacts of consumers' data management and highlight more managerial implications.

6.1. Committing to Avoid Personalized Pricing

Suppose firms in market A independently commit to using only uniform pricing. In this scenario, firms in both markets compete in the same manner as in the standard Hotelling model. Consequently, such individual commitments can yield the highest profits for firms in each market. However, these enhanced profits come at the expense of consumer surplus.[27]

6.2. Price Cap to Achieve Higher Profits

In market A, consumers choose to erase their data to avoid high personalized prices. Suppose each firm independently establishes a price cap for market A, denoted as $\bar{p}$. An appropriately designed $\bar{p}$ by each firm can render costly data erasure unbeneficial for consumers. Consequently, all consumers in market A may choose to remain opt-in, reversing the changes described in Section 4.2. In other words, market outcomes will revert to those observed under no data management, allowing firms to regain their higher profit levels.[28]

6.3. More Managerial Implications

Besides the above two self-constrained price discrimination, our study reveals the following additional managerial implications.

1. *Facilitate Opt-Out for Data Collection to Increase Profits* Consumers who choose to opt out of data collection can act as a competitive buffer in the data collection and application markets, enabling firms to achieve higher profits (see Proposition 4). Thus, giving consumers the option to opt out for data collection—despite a potential reduction in the volume of collected data—can be a strategically profitable decision for firms.

2. *Leverage Data Portability for Higher Profits* Data portability can lead to increased two-market profits by weakening competitive pressures in data collection. Consequently, firms should be more willing to accommodate consumer requests for data portability, recognizing its potential to drive profitability (see Proposition 5).

3. *Data Property Rights to Consumers: Opportunity in Disguise* Granting data property rights to consumers may seem like relinquishing valuable data assets. However, this strategy can actually benefit firms by diminishing competition in the data collection market, thereby enhancing profitability (see Proposition 7).

7. Conclusions

We analyze a duopoly model where consumers purchase products in two markets: one for data collection and one for data application. In the data collection market, firms collect consumer data. In the data application market, they compete by setting personalized prices for their own consumers and uniform prices for other consumers. Before firms set their prices, consumers can manage their data. Specifically, costly data erasure allows a consumer to avoid personalized pricing and opt for uniform prices. Additionally, we explore enhanced scenarios, including opting out of data collection, heterogeneous serving costs for consumers, data portability, and data ownership.

Contrary to expectations, we find that consumer data erasure results in higher prices and lower profits in data application markets. In the data collection market, firms compete more aggressively to attract consumers due to the higher sensitivity of their profits in the data application market to their market share. This heightened sensitivity arises because firms with larger customer bases can increase uniform prices more in the data application market. This makes data erasure less appealing and facilitates broader personalized pricing. The impact of data erasure on consumers is also twofold: it harms those who opt in while simultaneously benefiting the overall surplus in the data collection market.

Our analysis of the enhanced scenarios further reveals that a firm's two-market profits and total consumer surplus change nonmonotonically with consumers' data rights. Moreover, the impact of data management on cross-market competition varies significantly depending on the market in which consumers exercise their data rights. These findings underscore the need for careful consideration of consumer choice and market dynamics when designing policies related to personal data management.

Our theoretical results yield several testable empirical implications. First, when erasure becomes easier (smaller ε), application-market prices rise, collection-market prices fall, firm profits decline, and aggregate consumer surplus increases (Section 4.2). Second, partial opt-outs of data collection in market B generate a nonmonotonic profit effect, with competition first intensifying and then relaxing (Section 5.1). Third, portability mandates sharpen personalized-price rivalry in market A but soften competition in market B, thereby raising overall firm profits (Section 5.2). Finally, credible public commitments to uniform pricing reduce erasure rates and temper acquisition-market aggressiveness (Section 6). These mechanisms suggest that observable policy changes, such as APIs enabling portability, automated deletion requests, or consent-default modifications, should yield observable and falsifiable patterns in firm pricing strategies, customer acquisition costs, and opt-out rates, guiding future empirical research and policy evaluation.

Acknowledgments

The authors thank the Department Editor (Dmitri Kuksov), the Associate Editor, and anonymous reviewers for many valuable comments and suggestions. The authors thank Elias Carroni, Yi-Ling Cheng, Chongwoo Choe, Jay Pil Choi, Gwen-Jirō Clochard, Vincenzo Denicolò, Massimo de Francesco, Dingwei Gu, Bruno Jullien, Jianpei Li, Zhuoran Lu, Alireza J. Naghavi, Arina Nikandrova, Hiroyuki Odagiri, Flavio Pino, Marcel Preuss, Susumu Sato, Ke Shi, Naoki Wakamori, Chutian Wang, Zibin Xu, Zhendong Yin, Linqi Zhang, Wen Zhou, and the seminar participants at Competition Policy Research Center (CPRC) at the Japan Fair Trade Commission (JFTC), Fudan University, Hitotsubashi University, ShanghaiTech University, Shinshu University, University of Bologna, WHU, SHUFE, SJTU, CUFE, CityUHK, 2023 MaCCI Annual Conference, CEA 2023, AMES 2023, EARIE 2023, JEI 2023, and APIOC 2023 for their valuable discussions and comments. Although N. Matsushima serves as a member of the CPRC at the JFTC, the views expressed in this paper belong to the authors and should not be attributed to the JFTC. The usual disclaimer applies.

Endnotes

[1] Section 1 of Online Appendix B offers an in-depth examination of these real-world business cases.

[2] See Esteves and Resende (2016) and Ezrachi and Stucke (2016) about real-world cases of personalized offers.

[3] Jing (2016), Li and Jain (2016), and Colombo (2018) explore other variations of BBPD, considering post-consumption valuation discovery, consumer fairness concerns, and heterogeneous price elasticities, respectively, demonstrating how these factors influence competition and profitability.

[4] Laussel (2023) extends Chen et al. (2020) to two periods, finding BBPD less profitable.

[5] Related studies like Cordorelli and Padilla (2024), de Cornière and Taylor (2024), and Herresthal et al. (2025) also address the benefits and challenges of inter-market data usage.

[6] Moorthy and Shahrokhi Tehrani (2023) show that combining personalized pricing and advertising in a Hotelling model with ad-driven segment sizes intensifies pointwise asymmetric Bertrand competition.

[7] Li et al. (2025) analyze a two-period model of a storable good monopolist, showing that sharing data on first-period consumption increases profits and social welfare but harms consumer welfare. Ning et al. (2025) consider a monopolist's advertising strategy when both the firm and consumers have imperfect information about consumer preferences. They analyze how opt-in and opt-out data policies influence the effectiveness of advertising.

[8] Montes et al. (2019) and Valletti and Wu (2020) also examine how consumers' data management can be used to avoid personalized prices. Montes et al. (2019) focus on the data broker's optimal data sales strategy, while Valletti and Wu (2020) endogenize the quality of profiling technology in a monopoly setting.

[9] Ichihashi (2020), Ali et al. (2023), and Miklós-Thal et al. (2024) model privacy as strategic information disclosure by consumers, exploring optimal disclosure strategies and their influence on market outcomes. Braulin (2023) examines different privacy regimes when the preferences of consumers, who make single-item purchases, are two-dimensional.

[10] The scenario where consumers can opt out of data collection is analyzed in Section 5.1.

[11] Section 4 of Online Appendix B presents an analysis in which some consumers' product preferences are perfectly correlated across the two markets, confirming the robustness of our main insights.

[12] The Hotelling model is frequently used to analyze differentiation in healthcare and insurance plans (Gal-Or 1997, Ellis 1998, Biglaiser and Ma 2003, Chen et al. 2022). Such products are commonly believed to differentiate horizontally (Biglaiser and Ma 2003, Gaynor and Town 2011). For example, healthcare can differentiate in approved physician lists, diagnostic testing, real-time monitoring, and high-tech services, while insurance plans differ in coverage, care providers, reimbursement policies, and other characteristics. Section 5.3 explores the scenario where consumers also differ in their servicing costs.

[13] Using consumer data, firm A_i can offer tailored products to its targeted consumers, delivering enhanced value. Section 5.5 briefly discusses the scenario incorporating both data-enabled product personalization and personalized pricing, with a detailed analysis provided in Section 5 of Online Appendix B.

[14] Our main results do not change if consumers have heterogeneous data erasure costs. Refer to Section 8 of Online Appendix B for the corresponding analysis.

[15] The California Consumer Privacy Act prohibits businesses from discriminating against individuals who choose to erase their data, including practices like charging different prices or denying service. Although the GDPR lacks an explicit anti-discrimination provision, its principles of fairness, transparency, and the protection of individual rights could be leveraged to challenge businesses that

discriminate against individuals managing their data. Therefore, we assume that firm A_i charges the same uniform price to both its opt-out consumers and the rival's targeted consumers.

[16] Lemma 1 contrasts with the findings of Ke and Sudhir (2023) regarding consumer data management. In their study, consumers with a strong preference for a firm opt in because they can reap greater benefits from personalized products, while the firm cannot efficiently extract surplus due to uncertainty about their valuations. In our model, firms do not face such evaluation uncertainty with the assistance of data. Lemma 1 continues to hold even if we introduce benefits from personalized products, as demonstrated in Section 5 of Online Appendix B. This is because the firm continues to charge very high personalized prices on nearby targeted consumers to extract more surplus from the larger "pie."

[17] The inflated uniform prices align with the finding in the literature of imperfect customer recognition (e.g., Chen et al. 2001): a firm's loyal customer may receive the low price intended for switchers, discouraging firms from lowering their prices.

[18] Intuitively, a very small ε incentivizes many consumers to erase their data. This tempts a firm to charge a higher uniform price to its opt-out consumers, thereby forgoing the opportunity to poach the rival's consumers, or to lower the uniform price in order to attract a larger number of the rival's opt-out consumers. Section 13 of Online Appendix B demonstrates how this assumption is derived. Many studies (e.g., Worden 2025) have demonstrated the significant costs consumers face when attempting to erase their data. Under this assumption, the share of consumers erasing data is around 7% on the equilibrium path, moderately aligning with real-world consumer behaviors on data management. For instance, the GDPR decreases total recorded online searches by 10.7% (Aridor et al. 2023); Alcobendas et al. (2026) found that only 5% of users have opted out of behavioral targeting under GDPR and CCPA. The parameter range of ε can be significantly expanded when some consumers in market A cannot be targeted by either firm (see Section 2.2 of Online Appendix B).

[19] If consumers incur intrinsic privacy costs c when their data is collected in market B (Acquisti et al. 2016, Lin 2022, Ke and Sudhir 2023, Rhodes and Zhou 2026), the opt-out disutility can be expressed as the net value $\phi - c$. Hence, a lower opt-out disutility ϕ is equivalent to a higher privacy cost c.

[20] In Section 2 of Online Appendix B, we outline the specific conditions for consumers' heterogeneous ϕ being "uniformly high" and "uniformly low."

[21] Equivalently, this scenario can be interpreted as a fraction r of consumers whose data in market B are uninformative about their product preferences (i.e., locations) in market A, while the remaining consumers' data are informative.

[22] According to personal data laws, the responsibility for data portability primarily falls on the firm, and consumers do not incur significant costs. For instance, the guidelines in the GDPR specify that "the overall system implementation costs should not be charged to the data subjects." Similarly, the CCPA explicitly mandates that consumers should not be charged for data portability. Our results remain unchanged even if we assume that data portability entails a small cost for consumers.

[23] In Figure 5(a), we omit the upward-sloping solid line for $x \geqslant 1/2$. A consumer at x choosing the third option pays $p_2^a(x) + t(1-x) + \varepsilon = \alpha_1^a + tx + \varepsilon$, where the anticipated personalized prices $p_2^a(x)$ follow (6).

[24] When $\varepsilon < t/2$, the lines $\alpha_1^a + tx + \varepsilon$ and $\alpha_2^a + t(1-x) + \varepsilon$ in Figure 5(a) shift downward, causing part of $\alpha_1^a + tx + \varepsilon$ to lie below the line $t(1-x)$. A detailed analysis of this case is provided in Section 14 of Online Appendix B.

[25] For example, in market B of smart cars, c_B may represent costs related to customer communication, problem-solving, and after-sales services. In market A of auto insurance, c_A reflects the customer's driving risk, significantly impacting the firm's insurance costs. Refer to Shin et al. (2012) and Subramanian et al. (2014) for a more in-depth discussion on the variation in servicing costs among customers.

[26] This benefit may take the form of exclusive sales, coupons, discounts, or other data-related perks. The constraint $b \leqslant t$ ensures that the benefit is not so substantial as to fully negate the product differentiation between the firms. Technically, b plays a similar role to ε in the baseline model.

[27] Note that such commitments are sustainable if firms do not have incentives to deviate to using personalized pricing.

[28] Note that such price caps are sustainable if firms do not have incentives to deviate to removing such caps.

References

Acquisti A, Varian HR (2005) Conditioning prices on purchase history. *Marketing Sci.* 24(3):367–381.

Acquisti A, Taylor C, Wagman L (2016) The economics of privacy. *J. Econom. Lit.* 54(2):442–492.

Alcobendas M, Kobayashi SJ, Shi K, Shum M (2026) The impact of privacy measures on online advertising markets. *Rev. Econom. Stud.* Forthcoming.

Ali SN, Lewis G, Vasserman S (2023) Voluntary disclosure and personalized pricing. *Rev. Econom. Stud.* 90(2):538–571.

Anderson S, Baik A, Larson N (2023) Price discrimination in the information age: Prices, poaching, and privacy with personalized targeted discounts. *Rev. Econom. Stud.* 90(5):2085–2115.

Aridor G, Che YK, Salz T (2023) The effect of privacy regulation on the data industry: Empirical evidence from GDPR. *RAND J. Econom.* 54(4):695–730.

Armstrong M, Vickers J (2010) Competitive non-linear pricing and bundling. *Rev. Econom. Stud.* 77(1):30–60.

Belleflamme P, Vergote W (2016) Monopoly price discrimination and privacy: The hidden cost of hiding. *Econom. Lett.* 149:141–144.

Belleflamme P, Lam WMW, Vergote W (2020) Competitive imperfect price discrimination and market power. *Marketing Sci.* 39(5):960–1015.

Bester H, Petrakis E (1996) Coupons and oligopolistic price discrimination. *Internat. J. Indust. Organ.* 14(2):227–242.

Biglaiser G, Ma CA (2003) Price and quality competition under adverse selection: Market organization and efficiency. *RAND J. Econom.* 34(2):266–286.

Braulin FC (2023) The effects of personal information on competition: Consumer privacy and partial price discrimination. *Internat. J. Indust. Organ.* 87:102923.

Caminal R, Matutes C (1990) Endogenous switching costs in a duopoly model. *Internat. J. Indust. Organ.* 8(3):353–373.

Chen Y (1997) Paying customers to switch. *J. Econom. Management Strategy* 6(4):877–897.

Chen Y, Zhang ZJ (2009) Dynamic targeted pricing with strategic consumers. *Internat. J. Indust. Organ.* 27(1):43–50.

Chen Z, Choe C, Matsushima N (2020) Competitive personalized pricing. *Management Sci.* 66(9):4003–4023.

Chen Y, Narasimhan C, Zhang ZJ (2001) Individual marketing with imperfect targetability. *Marketing Sci.* 20(1):23–41.

Chen Z, Choe C, Cong J, Matsushima N (2022) Data-driven mergers and personalization. *RAND J. Econom.* 53(1):3–31.

Choe C, King S, Matsushima N (2018) Pricing with cookies: Behavior-based price discrimination and spatial competition. *Management Sci.* 64(12):5669–5687.

Choe C, Matsushima N, Tremblay MJ (2022) Behavior-based personalized pricing: When firms can share customer information. *Internat. J. Indust. Organ.* 82:102846.

Colombo S (2016) Imperfect behavior-based price discrimination. *J. Econom. Management Strategy* 25(3):563–583.

Colombo S (2018) Behavior- and characteristic-based price discrimination. *J. Econom. Management Strategy* 27(2):237–250.

Cordorelli D, Padilla J (2024) Data-driven envelopment with privacy-policy tying. *Econom. J.* 134(658):515–537.

de Cornière A, Taylor G (2024) Data-driven mergers. *Management Sci.* 70(9):6473–6482.

Dubé J-P, Misra S (2023) Personalized pricing and consumer welfare. *J. Political Econom.* 131(1):131–189.

Ellis RP (1998) Creaming, skimping and dumping: Provider competition on the intensive and extensive margins. *J. Health Econom.* 17(5):537–555.

Esteves R-B (2014) Price discrimination with private and imperfect information. *Scand. J. Econom.* 116(3):766–796.

Esteves R-B, Resende J (2016) Competitive targeted advertising with price discrimination. *Marketing Sci.* 35(4):576–587.

Ezrachi A, Stucke ME (2016) *Virtual Competition: The Promise and Perils of the Algorithm Driven Economy* (Harvard University Press, Cambridge, MA).

Fudenberg D, Tirole J (1984) The fat-cat effect, the puppy-dog ploy, and the lean and hungry look. *Amer. Econom. Rev. Papers Proc.* 74(2):361–366.

Fudenberg D, Tirole J (2000) Customer poaching and brand switching. *RAND J. Econom.* 31(4):634–657.

Gal-Or E (1997) Exclusionary equilibria in health-care markets. *J. Econom. Management Strategy* 6(1):5–43.

Gaynor M, Town RJ (2011) Competition in health care markets. Pauly MV, Mcguire TG, Barros PP, eds. Chapter 9, *Handbook of Health Economics*, vol. 2 (North Holland, Waltham, MA), 499–637.

Goldfarb A, Tucker C (2019) Digital economics. *J. Econom. Lit.* 57(1): 3–43.

Herresthal C, Mayskaya T, Nikandrova A (2025) Data linkage between markets: Hidden dangers and unexpected benefits. Preprint, submitted January 31, https://doi.org/10.2139/ssrn.4238784.

Ichihashi S (2020) Online privacy and information disclosure by consumers. *Amer. Econom. Rev.* 110(2):569–595.

Jing B (2016) Customer recognition in experience vs. inspection good markets. *Management Sci.* 62(1):216–224.

Jing B (2017) Behavior-based pricing, production efficiency, and quality differentiation. *Management Sci.* 63(7):2365–2376.

Ke T, Sudhir K (2023) Privacy rights and data security: GDPR and personal data driven markets. *Management Sci.* 69(8):4389–4412.

Koh B, Raghunathan S, Nault BR (2017) Is voluntary profiling welfare enhancing? *MIS Quart.* 41(1):23–41.

Laussel D (2023) Do firms always benefit from the presence of active customers? *Appl. Econom.* 55(20):2292–2307.

Laussel D, Resende J (2022) When is product personalization profit-enhancing? A behavior-based discrimination model. *Management Sci.* 68(12):8872–8888.

Li KJ, Jain S (2016) Behavior-based pricing: An analysis of the impact of peer-induced fairness. *Management Sci.* 62(9):2705–2721.

Li X, Li B, Yang Z (2025) The dark side of voluntary data sharing. *MIS Quart.* 49(1):155–178.

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

Liu Q, Serfes K (2004) Quality of information and oligopolistic price discrimination. *J. Econom. Management Strategy* 13(4):671–702.

Matutes C, Regibeau P (1988) "Mix and match": Product compatibility without network externalities. *RAND J. Econom.* 19(2):221–234.

Miklós-Thal J, Goldfarb A, Haviv AM, Tucker C (2024) Digital hermits. *Marketing Sci.* 43(4):697–708.

Montes R, Sand-Zantman W, Valletti T (2019) The value of personal information in online markets with endogenous privacy. *Management Sci.* 65(3):1342–1362.

Moorthy S, Shahrokhi Tehrani S (2023) Targeting advertising spending and price on the Hotelling line. *Marketing Sci.* 42(6):1057–1079.

Ning ZE, Shin J, Yu J (2025) Targeted advertising as implicit recommendation: Strategic mistargeting and personal data opt-out. *Marketing Sci.* 44(2):390–410.

Ozalp H, Ozcan P, Dinckol D, Zachariadis M, Gawer A (2022) "Digital colonization" of highly regulated industries: An analysis of Big Tech platforms' entry into health care and education. *Calif. Management Rev.* 64(4):78–107.

Pazgal A, Soberman D (2008) Behavior-based discrimination: Is it a winning play, and if so, when? *Marketing Sci.* 27(6):977–994.

Rhee KE, Thomadsen R (2017) Behavior-based pricing in vertically differentiated industries. *Management Sci.* 63(8):2729–2740.

Rhodes A, Zhou J (2026) Personalization and privacy choice. *RAND J. Econom.* Forthcoming.

Shaffer G, Zhang ZJ (2002) Competitive one-to-one promotions. *Management Sci.* 48(9):1143–1160.

Shiller BR (2020) Approximating purchase propensities and reservation prices from broad consumer tracking. *Internat. Econom. Rev. (Philadelphia)* 61(2):847–870.

Shin J, Sudhir K (2010) A customer management dilemma: When is it profitable to reward one's own customers? *Marketing Sci.* 29(4):671–689.

Shin J, Sudhir K, Yoon D (2012) When to "fire" customers: Customer cost-based pricing. *Management Sci.* 58(5):932–947.

Smith AN, Seiler S, Aggarwal I (2022) Optimal price targeting. *Marketing Sci.* 42(3):476–499.

Subramanian U, Raju J, Zhang J (2014) The strategic value of high-cost customers. *Management Sci.* 60(2):494–507.

Thisse J-F, Vives X (1988) On the strategic choice of spatial price policy. *Amer. Econom. Rev.* 78(1):127–137.

Valletti T, Wu J (2020) Consumer profiling with data requirements: Structure and policy implications. *Production Oper. Management* 29(2):309–329.

Villas-Boas JM (1999) Dynamic competition with customer recognition. *RAND J. Econom.* 30(4):604–631.

Wagner G, Eidenmüller H (2019) Down by algorithms: Siphoning rents, exploiting biases, and shaping preferences: Regulating the dark side of personalized transactions. *Univ. Chicago Law Rev.* 86(2):581–610.

Worden K (2025) Erasing illusions: A statutory framework for deletion in U.S. data privacy. *Texas A&M Law Rev.* 12(2):925–960.