

Research Repository

Data Linkages and Privacy Regulation

Accepted for publication in the American Economic Journal: Microeconomics

Research Repository link: <https://repository.essex.ac.uk/41981/>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the published version if you wish to cite this paper.

<http://doi.org/10.1257/mic.20250190>

Data Linkages and Privacy Regulation*

Rossella Argenziano[†] Alessandro Bonatti[‡]

November 13, 2025

Abstract

We assess the efficacy of privacy regulation when consumers are privacy conscious. We develop a model of data linkages where a consumer interacts sequentially with two firms: one firm collects data on consumer behavior, and the other firm leverages the data to set a quality level and a price. A data linkage benefits the consumer in equilibrium when the recipient firm is sufficiently similar to the collecting firm. We endogenize linkage formation under several existing privacy regulations. Voluntary consent requirements benefit consumers in equilibrium but blanket bans on discriminatory price and quality offers harm them.

KEYWORDS: consumer privacy; privacy regulation; consumer consent; data linkages; data rights; price discrimination; transparency; ratchet effect.

JEL CLASSIFICATION: D44, D82, D83.

*Argenziano acknowledges financial support through British Academy Leverhulme Research Grant SRG18R1–180705. Bonatti acknowledges financial support through NSF Grant SES–1948692. We thank Charles Angelucci, Heski Bar-Isaac, Dirk Bergemann, Gonzalo Cisternas, Jacques Crémer, Bob Gibbons, Martin Peitz, Salvatore Piccolo, Miguel Villas-Boas, Alex Waddington, Juanjuan Zhang, and Jidong Zhou for helpful comments, and seminar participants at Bank of Canada, Bergamo, City, Essex, Mannheim, MIT, THEMA, Toulouse, and Warwick for useful discussions, and Roi Orzach for stellar research assistance.

[†]Economics Department, University of Essex, Colchester CO4 3SQ, UK, rargenz@essex.ac.uk.

[‡]MIT Sloan School of Management, Cambridge, MA 02142, bonatti@mit.edu.

1 Introduction

From browsing and search histories to geolocation and health data to social media activity, individual firms and large digital platforms gather vast amounts of information about their users. These data are rarely acquired directly from consumers. Instead, most “big” datasets consist of information that firms “deduce from consumer behavior such as data about transactions on a platform (observed data)” (Vestager, 2020). That is, firms collect data about *behavior* from their interaction with the consumer, such as the sale of a good or the provision of a service. Other firms can then purchase these data to infer information about individual *preferences* and use it to target advertising, product offers, and even personalized prices across a myriad of subsequent transactions. Thus, data trades can create *linkages* across seemingly unrelated activities.¹

Until recently, most data sharing occurred unbeknownst to consumers or without their explicit consent. In response to such practices, regulatory interventions such as the European Union’s General Data Protection Regulation (GDPR) and the California Privacy Rights Act (CPRA) imposed transparency and consent requirements for data sharing. More recently, several U.S. states have enacted similar comprehensive privacy laws. At the federal level, legislative efforts such as the American Data Privacy and Protection Act (ADPPA) did not move forward for a vote in the full House or the Senate.

In this paper, we provide a framework to assess the efficacy of privacy regulation in the presence of strategic consumers. We find that the specific type of protection guaranteed to consumers when they exercise their rights determines whether transparency and consent requirements can limit the formation of data linkages that harm consumers and foster the creation of those that benefit them.

When consumers are *privacy conscious*, the welfare implications of privacy regulation depend not only on the allocation of control rights over data but also on the impact of exercising those rights on the terms of trade in the transactions where the data are first collected. Therefore, our approach consists of developing an equilibrium theory of data linkages with two key elements: (a) firms and consumers decide which linkages are formed under the prevailing transparency and consent requirements; and (b) both firms and consumers modify their behavior in the presence of a data linkage.

To this end, we introduce firm heterogeneity into a model of behavior-based price (and quality) discrimination. A single consumer interacts sequentially with two firms. The active

¹Marketing data from different transactions are linked through customer lifetime value (CLV) scores. These are measures of individual-consumer profitability that merchants use to personalize prices, quality levels, and other perks (Bonatti and Cisternas, 2020). Another prominent example is the Chinese Social Credit system, which determines access to credit, housing, and travel (Tirole, 2021).

firm in each period sets a quality level and a price (the *terms of trade*), and the consumer chooses how much to consume. The consumer’s preferences are correlated across the two periods. Therefore, if a data linkage exists, the second-period firm can observe the outcome of the first-period interaction and use this information to match its quality level and price to the consumer’s inferred willingness to pay. Within this framework, we derive the conditions under which data linkages increase consumer surplus. We then analyze the performance of recent privacy regulations against this welfare benchmark.

In the second period, a data linkage impacts both value creation (through quality discrimination) and value appropriation (through inefficient price discrimination) by the firm. To explain heterogeneity across firms in terms of the relative salience of price and quality discrimination, we introduce a firm-level *efficiency* index, which captures both the cost of offering a higher-quality product and the value it generates for the consumer.

In the first period, the consumer distorts her consumption level away from the static optimum to manipulate the second firm’s beliefs (Proposition 2). Specifically, if the consumer anticipates the data to be used mostly for price discrimination, she understates her willingness to pay in the first period to receive a less expensive (but lower quality) product in the second period—the canonical *ratchet effect* of Laffont and Tirole (1988). Conversely, if data are mostly used to target the quality of future products, the consumer overstates her willingness to pay to receive a higher-quality (though more expensive) product—the *niche envy* effect introduced by Turow (2008). The consumer’s behavior then impacts the first-period equilibrium terms of trade in an intuitive manner: if the consumer distorts her demand downward (upward), the first-period firm lowers (raises) both its price and its quality level.

The overall effect of a data linkage on consumer surplus across both periods reflects three forces: second-period discrimination, first-period demand distortion, and changes in the terms of trade. We show that a data linkage increases expected consumer surplus in the second period when the data-acquiring firm has a sufficiently high efficiency index (Proposition 3). In the first period, consumers benefit from improved terms of trade when the recipient firm is sufficiently similar to the collecting firm. By contrast, the behavioral distortion induced by the linkage is always costly, regardless of direction. Hence, these distortion costs are minimized when the second-period firm has a moderate efficiency index, which limits the incentive to deviate from myopic first-period behavior (Proposition 4).

To identify which linkages form in equilibrium, we extend the model by allowing the two firms to contract over the creation of a data linkage. We assume contracting occurs at the *ex ante* stage (i.e., before the consumer learns her type) and that bargaining between the two firms is efficient. We then ask which linkages form under a set of policies that assign progressively stronger control rights to the consumer.

We begin our analysis with the benchmark case of no regulation. In this case, the two firms always form a linkage because the first-period firm cannot commit to maintaining the consumer’s privacy. The consumer correctly anticipates this and modifies her behavior accordingly (Proposition 6). Relative to this case, mandatory transparency forces the firms to announce the formation of a linkage to the consumer. However, if the consumer has no veto rights, transparency regulation only grants valuable commitment power to the firms, which internalize both the impact of a linkage on the consumer’s behavior and the value of trading the data. Consequently, a linkage forms in equilibrium if and only if it improves the producers’ joint profits (Proposition 7).

We then analyze the effects of required consumer consent. Because the formation of a linkage is initiated by firms, requiring consumers’ consent to form a linkage is equivalent to endowing them with formal veto power. However, since the data are traded contextually to the first-period transaction, the strength of the consumer’s rights can be undermined by the ability of the first-period firm to impose a penalty on the consumer for not sharing her data. This penalty determines the equilibrium price the consumer must pay for anonymity. Motivated by recently approved legislation in the EU, USA, Canada, Brazil and Japan, we examine the impact of consent regulation in three settings that differ in this critical aspect.²

The weakest form of consent requirement, which we label “required consent,” is imposed in Brazil and Japan and does not constrain the firms’ response in any way. Therefore, consumer consent can be “extorted” by the firm: the consumer can choose to remain anonymous, but the first-period firm can refuse to trade so that the price of anonymity is the entire surplus from the first transaction. In this case, a subset of the producer-optimal linkages forms because, anticipating the consumer would forego trade in exchange for anonymity, the first-period firm sometimes finds it more profitable to commit to anonymity (Proposition 8).

A strengthening of the consumer’s control rights, which we label “voluntary consent,” is imposed in Canada and California, as well as other U.S. states. Under this policy, the firm cannot refuse service and cannot threaten the consumer with off-equilibrium terms of trade, i.e., the price of anonymity is given by the total impact of a linkage on consumer surplus. Since voluntary consent is equivalent to imposing mutual veto rights on linkage formation, a data linkage forms in this case only when it is Pareto improving (Proposition 9). Finally, we consider even more stringent laws, such as the GDPR, which mandate zero-price anonymity: the firms must offer the same terms of trade regardless of the consumer’s consent decision.

The effects of any policy on consumer welfare are twofold: first, privacy regulation directly constrains the formation of data linkages across different transactions; second, for data

²Appendix A contains a list of privacy laws and the various constraints they impose on how firms can respond when a consumer denies consent.

linkages that do form, the consumer’s privacy concerns lead to behavior distortions and modified terms of trade. Requiring transparency by the firms allows them to commit to sharing data only when a linkage increases producer surplus. Because the sets of producer-optimal and consumer-optimal linkages are different, transparency policies have an ambiguous effect on consumer welfare. Furthermore, voluntary consent requirements improve consumer welfare, but blanket bans on price discrimination unambiguously reduce it, relative to the case of pure voluntary consent. Under the right to equal service and price, not all linkages that benefit consumers form, and consumers consent to harmful linkages they would have vetoed under the equilibrium discriminatory offers (Proposition 10).

Applied Relevance Our model assumes that consumers who are aware of a data linkage distort their behavior at the time of data collection. Furthermore, under legal requirements of transparency and consent, both firms and consumers make rational decisions that lead to vetoing some linkages. The first assumption is supported by a growing body of evidence. For example, Kostka and Antoine (2020) find that, in a large survey of Chinese citizens enrolled in Social Credit System initiatives, 94% report that they have changed their behavior at least once to influence an SCS assessment, and 50% specifically report that they have changed their shopping behavior. In the credit industry, there is increasing attention on credit-building products, a set of financial products that individuals with a low credit score or no credit history purchase only for the purpose of creating or improving their credit score (Burke et al., 2023). Relatedly, Tang (2019) shows that disclosure requirements for loan applicants in an online peer-to-peer lending platform in China significantly affect application rates.

In addition to these examples, anecdotal evidence shows that consumers display a range of strategic behaviors to manipulate, rather than prevent, data collection by firms to obtain lower prices. For example, they use tools such as VPNs and User Agent Switchers during their purchases³ or add products to their online shopping carts without completing the purchase to obtain “retargeting” discount offers.⁴

The incentive implications of linked transactions are even stronger in business-to-business interactions. For example, Paes Leme, Pal and Vassilvitskii (2016) describe the prevalence of personalized, adaptive reserve prices in Google-sponsored search keyword auctions. In this case, the linkage between past bids and future reserve prices motivates advertisers to understate their willingness to pay in an attempt to manipulate future reserve prices, as in

³See, in particular, <https://bit.ly/cheaperinflightWiFi> and <https://bit.ly/AirfareVPN>

⁴AdRoll, a digital marketing platform specialized in retargeting of “abandoners,” advises firms to “... start by reminding shoppers that they have abandoned their cart, then a friendly call to finish the purchase. If they’re still not budging, provide a compelling offer. (But you’ll want to make sure that you don’t encourage cart abandonment so shoppers can unlock the deal.)” See <https://bit.ly/RetargetingCarts>.

Golrezaei, Javanmard and Mirrokni (2020).

Finally, the result that firms and consumers veto some linkages as a result of privacy legislation is supported by recent empirical papers on the effects of the GDPR and of the California Consumer Privacy Act (CCPA, the precursor of the CPRA) (e.g., Bae, Mayya and Nian (2023), Aridor, Che and Salz (2023), Peukert et al. (2022), Johnson, Shriver and Goldberg (2023), and Canayaz, Kantorovitch and Mihet (2022)). These papers show a drop in traffic and website connectivity following the introduction of these regulations. In particular, Bae, Mayya and Nian (2023) find that, after the introduction of the CCPA, firms reduced their use of personalized advertising and personalized recommendations. These findings suggest a reduction in data linkages. Moreover, Aridor, Che and Salz (2023) document the ability of some consumers to avail themselves of privacy-preserving technology (e.g., clearing cookies, anonymous browsing) as an imperfect substitute for vetoing linkages even before the introduction of the GDPR.

Related Literature Our paper joins a vast body of work on the economics of privacy and markets for information surveyed, e.g., by Acquisti, Taylor and Wagman (2016) and Bergemann and Bonatti (2019). Our model is most directly related to the behavior-based price discrimination literature (Fudenberg and Villas-Boas, 2006, 2012), with seminal contributions by Villas-Boas (1999), Taylor (2004), Acquisti and Varian (2005), Calzolari and Pavan (2006), Zhang (2011), and most recently Baye and Sappington (2020). Relative to these papers, our model allows for heterogeneous sources and heterogeneous uses of data, introduces a protocol for endogenous linkage formation, and shows how it operates under different regulatory regimes.

Several recent contributions, including Cummings et al. (2016), Frankel and Kartik (2022), Ball (2025), Bonatti and Cisternas (2020), and Jann and Schottmüller (2020), study how consumers’ manipulation incentives reduce the amount of information transmitted *in equilibrium*. Closest to our work, Shen and Villas-Boas (2018) study a model where advertising messages are targeted on the basis of a consumer’s past purchases. The prospective utility gains from targeted advertising for the consumer affect the prices of the goods she buys and the amount of information conveyed by her purchases.

A growing body of literature—including Conitzer, Taylor and Wagman (2012), Belleflamme and Vergote (2016), Montes, Sand-Zantman and Valletti (2019), Ke and Sudhir (2023), Dengler and Prüfer (2021), Ali, Lewis and Vasserman (2023), and Chen (2022)—examines the amount of information available to firms when consumers can *actively* protect their privacy by remaining anonymous, at some direct cost or with signaling implications. Krähmer and Strausz (2023) study nonlinear pricing when consumers have heterogeneous

preferences over how much information to reveal, and show that offering a “privacy option” can be part of the optimal menu. Ichihashi (2020) analyzes *ex ante* information disclosure about a consumer’s preferences to a firm that tailors both product characteristics and price.

Doval and Skreta (2024) study product personalization in a two-period price-discrimination problem, where the buyer’s tastes are correlated over time. They show that, by offering an appropriately restricted menu of quality levels, the seller effectively commits to gaining less information about the consumer’s first-period type. This alleviates the ratchet effect, relative to offering the statically optimal mechanism, where the allocation is strictly monotone in the consumer’s type and hence fully informative. Relative to Doval and Skreta (2024), our model abstracts from product-line design but considers two distinct firms that negotiate over the formation of a data linkage under different regulatory regimes. This allows us to focus on the impact of formal privacy laws on the equilibrium level of consumer privacy in heterogeneous transactions.

Motivated by privacy protection laws, Fainmesser, Galeotti and Momot (2023) examine the distinction between data collection and data protection from adversaries. Jullien, Lefouili and Riordan (2020) study the related phenomenon of data *leakages*. These leakages are modeled as reduced-form negative consequences of information diffusion. Likewise, Dosis and Sand-Zantman (2023) study a dynamic model of price discrimination in which a monopolist faces a trade-off between processing and monetizing its consumers’ data by selling it to adverse third parties. In our paper, we instead focus on a specific microfoundation for privacy preferences and examine how it shapes the outcome of regulation.

A recent set of contributions, including Choi, Jeon and Kim (2019), Ichihashi (2021), Acemoglu et al. (2022), and Bergemann, Bonatti and Gan (2022), studies the data externalities associated with collecting information from multiple consumers with correlated preferences. This literature largely abstracts from distortions in behavior arising from the collection of personal data. A notable exception is Liang and Madsen (2024), who analyze the incentive effects of correlated types in a model of career concerns with multiple agents and data linkages.

Organization The rest of this paper is organized as follows. In Section 2, we introduce our framework and discuss the main assumptions. In Section 3, we analyze the linear equilibria of a static game without data linkages and of a dynamic game with data linkages. In Section 4, we evaluate which data linkages benefit consumers and firms. In Section 5, we endogenize the formation of data linkages as a function of the decision rights allocated to firms and consumers, and we map the outcome to existing privacy regulations. A comprehensive list of approved privacy regulations is in Appendix A, and all proofs are in Appendix B.

2 Model

Players, Actions and Payoffs Consider a single consumer who lives for two periods and interacts with two firms sequentially: a collecting firm in the first period and a receiving firm in the second period. The consumer's interaction with the collecting firm can inform the strategy of the receiving firm because her preferences are correlated across the two periods.

The active firm in each period $t = 1, 2$ sets a quality level y_t and charges a unit price p_t for its product. The consumer, in turn, chooses a quantity level q_t . The consumer's per-period surplus is given by

$$U(\theta_t, p_t, y_t, q_t) = (\theta_t + b_t y_t - p_t) q_t - \frac{q_t^2}{2}. \quad (1)$$

The consumer's period- t type θ_t captures her "baseline" willingness to pay per unit of the product, i.e., the intercept of her demand curve before accounting for the firm's quality y_t . The parameter b_t is a firm-level characteristic that represents the sensitivity of the consumer's willingness to pay to the quality of the good produced by firm t . Because the sensitivity b_t is independent of the consumer's type, it succinctly captures the nature of the interaction between any consumer and firm t . In particular, the case $b_t = 0$ corresponds to a pure price-setting firm. We refer to the price-adjusted value of quality $b_t y_t - p_t$ as the *terms of trade* that firm t offers to the consumer.

In each period, the consumer privately observes her type θ_t . The consumer's period- t type is drawn from a commonly known distribution on a positive interval. In particular, the first-period type $\theta_1 > 0$ is given by

$$\theta_1 = \mu + \varepsilon_1,$$

where ε_1 has mean 0 and variance $\sigma_1^2 > 0$. The second-period type $\theta_2 > 0$ is given by

$$\theta_2 = \mu + \varphi(\theta_1 - \mu) + \varepsilon_2, \quad (2)$$

where $\varphi \in [0, 1]$. The shock ε_2 is uncorrelated with ε_1 and has mean 0 and variance $\sigma_2^2 \geq 0$. We further assume that the lower bound of the type support is sufficiently high to yield positive demand in both periods.

Each firm t faces a marginal cost of producing quantity $q_t \geq 0$ that increases with the quality level $y_t \geq 0$. The firm also incurs a fixed cost of offering quality y_t to each consumer.

Firm t 's realized profits are given by

$$\Pi_t = p_t q_t - k_t y_t q_t - \frac{c_t y_t^2}{2}. \quad (3)$$

The parameters (b_t, c_t, k_t) are commonly known at the onset of the game.

For each period $t = 1, 2$, we further assume

$$b_t - k_t > 0 \quad \text{and} \quad b_t - k_t < \sqrt{2c_t}. \quad (4)$$

These two assumptions ensure the existence of an interior optimum for the firm's problem of choosing (p_t, y_t) . Because quality y_t raises willingness to pay at rate b_t but also unit costs at rate k_t , the former assumption ensures that the returns from increasing quality are non-negative and rules out the corner solution $y_t = 0$. Because a higher quality y_t allows the firm to both sell more and to charge a higher price, the latter assumption ensures that the returns from increasing quality are decreasing, i.e., that the firm's profits do not grow without bound as $y_t \rightarrow \infty$.

Firm 1 sets (p_1, y_1) on the basis of prior information on the consumer's type only. Firm 2's information depends on the presence of a *data linkage*. We define a data linkage as the ability of the second-period firm to observe the outcome of the first-period transaction (p_1, y_1, q_1) before interacting with the consumer. In the absence of a data linkage, firm 2 sets (p_2, y_2) on the basis of the common prior distribution. Figure 1 illustrates our model.

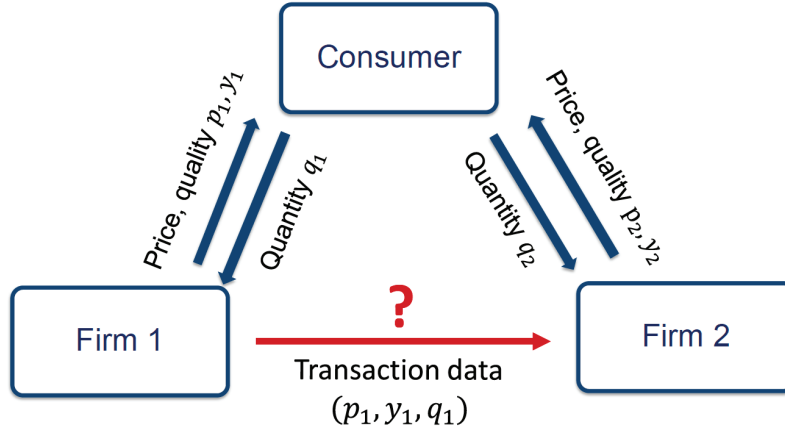


Figure 1: Model Depiction

Timing and Equilibrium The timing of our game is as follows:

1. Firm 1 offers a price p_1 and a quality level y_1 to the consumer.
2. The consumer learns her type θ_1 and selects a quantity q_1 .
3. In the presence of a data linkage, firm 2 observes the first-period outcome (p_1, y_1, q_1) before setting (p_2, y_2) . Otherwise, firm 2 sets (p_2, y_2) under the prior information only.
4. The consumer learns her type θ_2 and selects a quantity q_2 .

We focus on *linear equilibria*, as defined in Ball (2025). These are fully separating, Bayesian Nash equilibria in which the second-period firm’s strategy is linear in the first-period outcome variables (p_1, y_1, q_1) both on and off the path of play.⁵

Discussion of Assumptions An important modeling choice is the per-customer fixed cost of quality. This assumption is most plausible for service industries where a higher tier entails a lump-sum expenditure for each user: the staff needed to maintain dedicated call-center customer service lines for premium customers increases in the number of such customers; bespoke travel-planning firms spend time crafting itineraries regardless of trip length; and document-preparation platforms (e.g., tax software with expert review) pay a fixed reviewer fee regardless of the complexity of a tax return. In addition to the fixed costs, firms bear a marginal cost which also scales with quality. Of course, the relative importance of these fixed and marginal components depends on the application at hand.

Our model restricts firms to charging linear prices for their products. A more general approach would allow nonlinear pricing in both periods. For example, firm 1 could offer a menu of price-quality-quantity bundles (p_1, y_1, q_1) , and the consumer chooses which bundle to buy. With a data linkage, firm 2 then observes the consumer’s first-period choice from the menu and offers another (possibly degenerate) menu. A well-known result for this class of models (Calzolari and Pavan, 2006) is that joint producer surplus is larger if the second-period firm operates under prior information only, i.e., without a data linkage. In our setting, restrictions on the firms’ pricing instruments restore the profitability of data trades.

Appendix D shows that our results are robust when the second-period firm can employ nonlinear pricing in the form of a two-part tariff. Under both linear and nonlinear pricing, the consumer retains some private information about her second-period type θ_2 , and the associated information rent gives her an incentive to distort her first-period quantity choice to manipulate beliefs. We provide conditions under which the game with two-part tariffs admits a linear equilibrium, and we prove that, under those conditions, allowing nonlinear pricing does not qualitatively alter the strategic impact of a data linkage.

In our model, a data linkage allows the second-period firm to perfectly observe the outcome of the first-period interaction (p_1, y_1, q_1) . Appendix E relaxes this assumption and shows that all the qualitative properties of our game with a linkage extend to the case of noisy observations of the consumer’s first-period purchases.

Finally, the consumer knows the characteristics (b_t, c_t, k_t) of the firm she will face in each

⁵Linear equilibria are more tractable in our setting than Perfect Bayesian Equilibria. Weng, Wu and Yin (2023) study the Riley Perfect Bayesian equilibria of our model and find similar qualitative properties. In Appendix C, we show that, as the support of first-period types θ_1 grows without bound, the Riley equilibrium strategy for each consumer type converges pointwise to our linear equilibrium (Proposition 11).

period. Argenziano and Bonatti (2023) introduce a distribution of heterogeneous second-period firms, which captures the consumer’s uncertainty in period 1 over her future interactions. The analysis of privacy regulation in Section 5 yields more easily interpretable results with a single firm in each period but can be applied to the richer setting as well.

3 Equilibrium Strategies

The equilibrium strategies critically depend on the presence of a data linkage across the two periods. We begin with the game without a linkage in Section 3.1. In Section 3.2, we uncover the strategic impact of data linkages in a dynamic model where firm 2 can infer the consumer’s first-period type from observing the first-period outcome.

3.1 The Game without a Data Linkage

In the absence of a data linkage, the two stages of the game are separate. Each firm $t = 1, 2$ sets (p_t, y_t) based on the prior distribution of θ_t . The consumer observes the current firm’s (p_t, y_t) and maximizes her current-period surplus (1) by choosing the quantity

$$q(\theta_t, p_t, y_t) = \theta_t + b_t y_t - p_t. \quad (5)$$

The period- t firm maximizes the expected value of its profits (3) given the demand function obtained by integrating (5) over the consumer’s types. Because $\mathbb{E}[\theta_1] = \mathbb{E}[\theta_2] = \mu$, for any choice of quality y_t , the resulting firm-level demand curve is given by

$$\mathbb{E}[q(\theta_t, p_t, y_t)] = \mu + b_t y_t - p_t.$$

Thus, firm t ’s choice of quality y_t is akin to a costly investment that shifts out the demand curve. For any level of y_t , firm t then chooses the monopoly price

$$p_t(y_t) = \frac{1}{2}(\mu + b_t y_t + k_t y_t).$$

The firm’s problem then consists of identifying the optimal y_t^* given the distribution of consumer types and considering that monopoly pricing enables the firm to appropriate only a fraction of the surplus it generates.

Proposition 1 (Equilibrium without a Linkage).

Without a data linkage, the equilibrium price and quality levels in each period are given by

$$\hat{p}(\mu) = \frac{c_t + (b_t - k_t) k_t}{2c_t - (b_t - k_t)^2} \mu, \quad (6)$$

$$\hat{y}(\mu) = \frac{b_t - k_t}{2c_t - (b_t - k_t)^2} \mu. \quad (7)$$

Both firms' equilibrium price and quality increase in μ . Intuitively, a firm offers a higher quality at a higher price when it anticipates that the consumer will purchase more units. The rates at which price and quality increase are crucial to understand the impact of the firms' beliefs on the consumer's surplus, and the incentives for the consumer to manipulate those beliefs. We then ask, does the value of quality $b_t y_t$ respond more or less than price p_t to a change in the expected consumer type? In other words, do the terms of trade improve or worsen for a consumer who is believed to be of a higher type? And how does this affect her surplus? To this end, we first define *firm t 's efficiency* as

$$\lambda_t \triangleq \frac{(b_t - k_t)^2 - c_t}{2c_t - (b_t - k_t)^2}.$$

The function λ_t captures the efficiency of an improvement in quality by firm t because it is increasing in $(b_t - k_t)^2$ and decreasing in c_t ; it is also positive if the former is larger than the latter, i.e., if the additional surplus per unit of the product created by an increase in quality exceeds its marginal cost. Under the assumptions in (4), the efficiency index λ_t takes values in $(-1/2, \infty)$.

We can now examine the total impact of the firm's beliefs on the equilibrium terms of trade using the firm's efficiency index. The expression for the terms of trade, $b_t y_t - p_t$, can be rewritten as

$$b_t \hat{y}(\mu) - \hat{p}(\mu) = \lambda_t \cdot \mu. \quad (8)$$

Furthermore, substituting (8) into the demand function (5), we can write the realized quantity level as

$$\hat{q}(\theta_t, \mu, \lambda_t) = \theta_t + \lambda_t \mu, \quad (9)$$

the realized consumer surplus for consumer type θ_t as

$$\hat{U}(\theta_t, \mu, \lambda_t) = \frac{1}{2} \hat{q}(\theta_t, \mu, \lambda_t)^2 = \frac{1}{2} (\theta_t + \lambda_t \mu)^2, \quad (10)$$

and the realized profits as

$$\hat{\Pi}(\theta_t, \mu, \lambda_t) = \frac{1}{2} \mu (2\theta_t - \mu) (1 + \lambda_t). \quad (11)$$

We can now conclude that when a firm can offer a higher quality efficiently, such that $\lambda_t > 0$, the value of the firm's optimal quality $b_t \hat{y}$ increases faster than the monopoly price $\hat{p}$ as a function of the expected consumer type. Thus, the consumer obtains better terms of trade and a higher realized surplus when the firm's prior beliefs $\mu := \mathbb{E}[\theta_t]$ improve. The opposite holds when $\lambda_t < 0$, i.e., when the surplus created by an improvement in quality is small compared to its marginal cost.

3.2 The Game with a Data Linkage

We now turn to the game played by the consumer with firms λ_1 and λ_2 in the presence of a data linkage. Thus, firm 2 observes the terms of trade offered by firm 1 and the quantity purchased by the consumer in the first period.

Relative to the game without a linkage, the consumer knows that her first-period choice of quantity q_1 conveys information about her type θ_1 to firm 2. Since the consumer knows θ_1 and λ_2 , she solves the following problem:

$$\max_{q_1} [U(\theta_1, p_1, y_1, q_1) + \mathbb{E}[\hat{U}(\theta_2, m_2(q_1, p_1, y_1), \lambda_2) \mid \theta_1]], \quad (12)$$

where $\hat{U}(\cdot)$ is defined in (10), θ_2 is defined in (2), and $m_t(q_1, p_1, y_1)$ denotes firm 2's posterior mean beliefs over the consumer's period- t type, i.e.,

$$\begin{aligned} m_2(q_1, p_1, y_1) &\triangleq \mathbb{E}[\theta_2 \mid q_1, p_1, y_1] = (1 - \varphi) \mu + \varphi m_1(q_1, p_1, y_1), \\ m_1(q_1, p_1, y_1) &\triangleq \mathbb{E}[\theta_1 \mid q_1, p_1, y_1]. \end{aligned}$$

In a linear equilibrium, firm 2 holds (degenerate) posterior beliefs over θ_1 that are captured by the function $m_1(q_1, p_1, y_1)$. Firm 2 then updates its belief over θ_2 from the prior μ to a weighted average of μ and θ_1 . As the posterior beliefs in (12) play the same role as the prior beliefs in (10) above, the direction of the consumer's manipulation incentives depends on the sign of λ_2 . In turn, the strength of such incentives depends on both the magnitude of λ_2 and the sensitivity of firm 2's posterior $m_2(q_1, p_1, y_1)$ to the first-period quantity.

Proposition 2 establishes the existence and uniqueness of linear equilibria in our dynamic game and characterizes the equilibrium strategies.

Proposition 2 (Equilibrium with a Linkage).

There exists a unique linear equilibrium of the game with a data linkage. In this equilibrium:

1. *The consumer's first-period demand function is given by*

$$q_1^*(\theta_1, p_1, y_1) = \frac{1 + \sqrt{1 + 4\lambda_2(\lambda_2 + 1)\varphi^2}}{2} \theta_1 + b_1 y_1 - p_1 + \frac{2\lambda_2(\lambda_2 + 1)(1 - \varphi)\varphi\mu}{1 + \sqrt{1 + 4\lambda_2(\lambda_2 + 1)\varphi^2}}. \quad (13)$$

2. *The first-period terms of trade $b_1 y_1^* - p_1^*$ are given by*

$$\frac{\lambda_1(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 2\varphi - 1)}{2\varphi} \mu. \quad (14)$$

3. *The consumer's second-period demand function and the equilibrium terms of trade are given by $q(\theta_2, p_2, y_2)$ and $b_2 \hat{y}((1 - \varphi)\mu + \varphi\theta_1) - \hat{p}((1 - \varphi)\mu + \varphi\theta_1)$, as defined in (5) and (8), respectively.*

In the first period, the consumer distorts her equilibrium behavior away from the naive benchmark in (5) whenever $\lambda_2 \neq 0$. In particular, the last term in the equilibrium demand function (13) indicates that the intercept of the consumer demand function increases if $\lambda_2 > 0$ and decreases if $\lambda_2 < 0$. Moreover, the coefficient on θ_1 in (13) indicates that the consumer places more weight on her type than in the myopic benchmark if $\lambda_2 > 0$ and less weight if $\lambda_2 < 0$. From the perspective of firm 1, the consumer's behavior translates into a further shift of the demand curve (upwards if $\lambda_2 > 0$ and downwards if $\lambda_2 < 0$).

Finally, the magnitude of the equilibrium distortion in q_1 is increasing in $|\lambda_2|$. To gain intuition, fix a linear conjecture by firm 2 for the first-period consumer's strategy,

$$\hat{q}_1 = \hat{\alpha}\theta_1 + \hat{\beta}y_1 + \hat{\gamma}p_1 + \hat{\delta}.$$

Now consider the weight α that the consumer's best reply places on θ_1 as a function of firm 2's conjecture $\hat{\alpha}$. If firm 2 expects the consumer to maximize her period-1 surplus (i.e., $\hat{\alpha} = 1$), the consumer would place more (less) weight on her type in the first period depending on whether $\lambda_2 > 0$ (to raise firm 2's beliefs) or $\lambda_2 < 0$ (to depress them). Figure 2 illustrates the consumer's best replies to firm 2's conjecture for different values of $\lambda_2 < 0$ (left) and $\lambda_2 > 0$ (right), respectively.

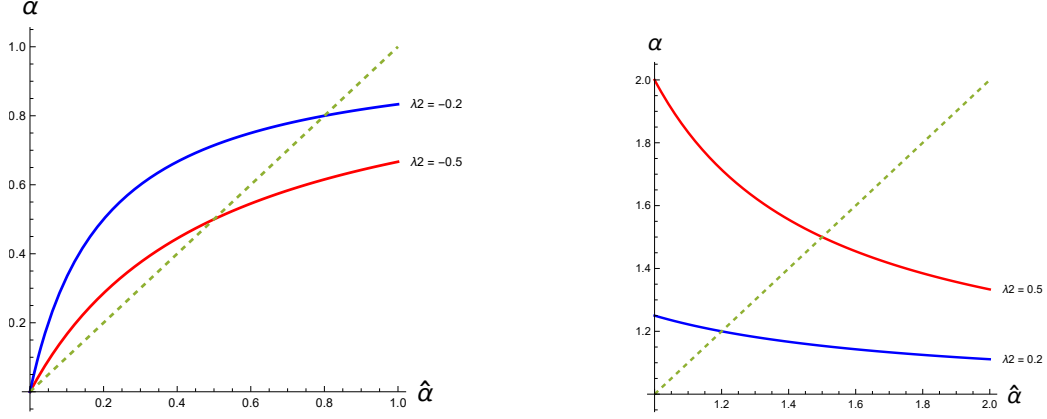


Figure 2: Consumer Best Responses

4 Welfare Effects of Data Linkages

To be able to evaluate the welfare effects of the linkages that endogenously emerge under different types of privacy legislation, we explore which linkages increase the consumer's and the firms' *ex-ante* payoffs in equilibrium, respectively. We start by considering the effects in the second period, when the linkage affects the transaction in which information is used, then proceed to consider the first period, when information is acquired, and finally, analyze the overall intertemporal effects.

4.1 Welfare Effects: Information Use

To quantify the welfare effect of a linkage on the second-period payoffs, we start by observing that the second-period game is analogous to the game of Section 3.1. In particular, if a linkage does not form, the second-period game is equal to the static game played under prior information and the equilibrium strategies are given by (5)-(7). If instead a linkage forms, the linear equilibrium of Section 3.2 is played, and the second-period game is played under complete information on θ_1 . We therefore define

$$\Delta U_2 = \mathbb{E}[\hat{U}(\theta_2, (1 - \varphi) \mu + \varphi \theta_1, \lambda_2)] - \mathbb{E}[\hat{U}(\theta_2, \mu, \lambda_2)] \quad (15)$$

as the expected value of a linkage in period 2 for the consumer, and

$$\Delta \Pi_2 = \mathbb{E}[\hat{\Pi}(\theta_2, (1 - \varphi) \mu + \varphi \theta_1, \lambda_2)] - \mathbb{E}[\hat{\Pi}(\theta_2, \mu, \lambda_2)]$$

as the expected value of a linkage for firm 2.

Proposition 3 (Second-Period Welfare Effects).

1. $\Delta\Pi_2$ is positive and increasing in σ_1^2 and φ .
2. ΔU_2 is positive if $\lambda_2 > 0$ and negative if $\lambda_2 < 0$. Moreover, its absolute value is increasing in σ_1^2 and φ .
3. There exists a unique $\underline{\lambda}_2 \in (-1/2, 0)$ such that a data linkage in period 2 increases total surplus $\Delta U_2 + \Delta\Pi_2$ if and only if $\lambda_2 > \underline{\lambda}_2$.

A linkage improves firm 2's information about the consumer's period-2 type θ_2 by creating a positive correlation between the firm's beliefs m_2 and θ_2 . Firm 2 benefits from this superior information regardless of λ_2 because information allows the firm to tailor its price and quality levels to the consumer's type. Discriminatory prices and qualities do not, however, unambiguously harm consumers. Indeed, such discrimination benefits consumers on average if and only if $\lambda_2 > 0$. Intuitively, for $\lambda_2 > 0$ positive correlation between beliefs and type implies better terms of trade for the consumer when she consumes more units, because her consumption is increasing in her type. Vice versa, for $\lambda_2 < 0$ the correlation implies *worse* terms of trade when the consumption is high. Finally, the total expected value of a linkage in period 2 increases with information for $\lambda_2 > 0$ as well as for some moderately negative λ_2 .

Given that the value of a linkage is strictly connected to the correlation it creates between the firm's beliefs m_2 and the type θ_2 , it is not surprising that its absolute value is increasing in σ_1^2 and φ , both for firm 2 and for the consumer. Intuitively, a larger variance of the first-period type and/or a larger type persistence imply that the correlation created by the linkage is stronger.

4.2 Welfare Effects: Information Acquisition

We now analyze the welfare effect of a linkage on the first-period payoffs. In the first period, a data linkage introduces behavior distortions and impacts the terms of trade, as described in Proposition 2. While the distortion in consumer behavior is a function of firm 2's efficiency λ_2 only, the effect of this demand shift on the prevailing terms of trade in period 1 depends critically on both firms' efficiency levels. We begin with two examples in Figure 3. In both examples, we consider the limit case of a pure price-setting firm 1, with $b_1 = k_1 = 0$, hence $\lambda_1 = -1/2$. The two examples, however, differ in the sign of λ_2 .

Figure 3(a) captures the *ratchet* effect (Laffont and Tirole, 1988): because $\lambda_2 < 0$, the consumer knows that a higher posterior belief by this firm leads to worse terms of trade. Firm

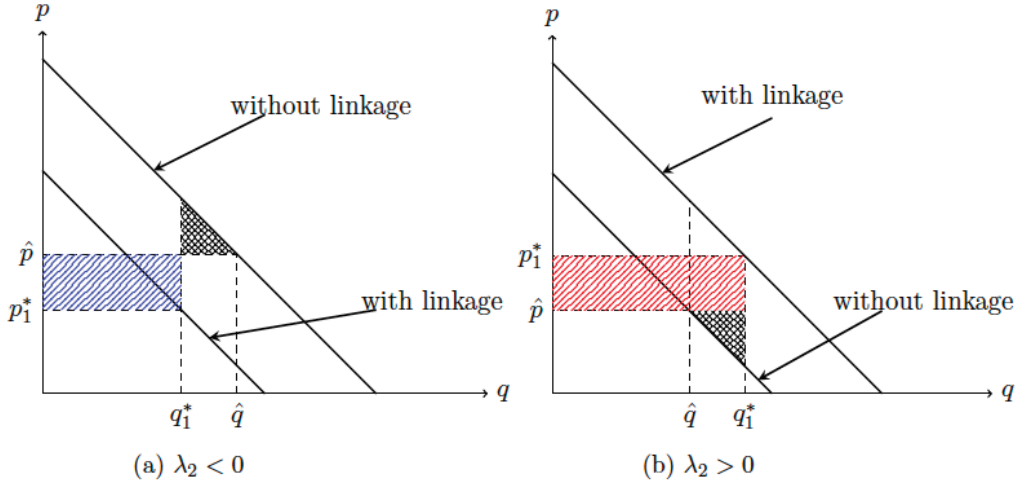


Figure 3: Demand and Price Shifts when $\lambda_1 = -1/2$

1 anticipates the consumer's concern over the second-period price, expects a lower demand curve, and charges a lower price p_1 . In equilibrium, the consumer purchases a smaller quantity q_1^* at a lower price p_1^* relative to the outcome $(\hat{q}, \hat{p})$ of a game with no linkage. Importantly, the consumer may benefit from this outcome, as the inframarginal discount $\hat{p} - p_1^*$ on q_1^* units (i.e., the blue rectangle) can compensate for the loss from foregoing consumption of the marginal units (i.e., the black triangle).

Figure 3(b) captures the *niche envy* effect described by Turow (2008): firm 2's efficiency is $\lambda_2 > 0$, which means the consumer wishes to manipulate the firm's beliefs upward to improve her terms of trade. Thus, firm 1 expects a higher demand curve with a linkage than without, and charges a higher price $p_1^* > \hat{p}$. Nonetheless, the consumer purchases more units than is myopically optimal, $q_1^* > \hat{q}$. Therefore, in the first period, the consumer buys “too much” (the black triangle) at a higher price (the red rectangle).

For any (λ_1, λ_2) , the welfare effects of a data linkage in the first period operate through the following two channels. First, the terms of trade offered by firm 1 change to reflect the shift in consumer demand. Comparing the equilibrium terms of trade with and without a linkage, given in (8) and (14), we obtain

$$(b_1 y_1^* - p_1^*) - (b_1 \hat{y} - \hat{p}) = \frac{\mu}{2\varphi} \lambda_1 (\sqrt{1 + 4\lambda_2(\lambda_2 + 1)\varphi^2} - 1). \quad (16)$$

This difference is proportional to λ_1 and has the same sign as $\lambda_1 \lambda_2$. Thus, the consumer obtains better first-period terms of trade if the two firms are similar, in the sense that they both can (or cannot) offer a higher quality efficiently. This effect is best captured by the blue vs. red rectangles in Figure 3: in panel (a), λ_1 and λ_2 have the same sign; in panel

(b), they have opposite signs. Likewise, if $\lambda_1 > 0$ and $\lambda_2 > 0$, then an upward shift in the consumer's first-period demand leads to more generous terms of trade in the first period.

Second, consumers engage in costly signaling in the first period. Even though firm 1 fully anticipates their behavior, signaling still leads consumers to distort their first-period quantity away from the optimal amount. Because, for each type θ_1 , demand is distorted by an amount that is linear in λ_2 , the magnitude of this loss (the area of the black triangles in Figure 3) is proportional to $(\lambda_2)^2$.

Proposition 4 compares the expected consumer and producer surplus in the first period when a data linkage is active between firms λ_1 and λ_2 to the expected surplus levels in the absence of linkage. In what follows, let $\tilde{\sigma}_1 \triangleq \sigma_1/\mu$ denote the coefficient of variation of the consumer's first-period type. Further, let

$$U_1^*(\theta_1, \lambda_1, \lambda_2) = (\theta_1 + b_1 y_1^* - p_1^*) q_1^*(\theta_1, p_1^*, y_1^*) - q_1^*(\theta_1, p_1^*, y_1^*)^2 / 2 \quad (17)$$

denote the equilibrium consumer surplus in period 1 in the game with a linkage, and

$$\Delta U_1 = \mathbb{E}[U_1^*(\theta_1, \lambda_1, \lambda_2)] - \mathbb{E}[\hat{U}(\theta_1, \mu, \lambda_1)] \quad (18)$$

denote the value of a linkage for the consumer in period 1. Similarly, let

$$\mathbb{E}[\Pi^*(\theta_1, \lambda_1, \lambda_2)] = \mathbb{E}[p_1^* q_1^*(\theta_1, p_1^*, y_1^*) - k_1 y_1^* q_1^*(\theta_1, p_1^*, y_1^*) - c_1 y_1^{*2} / 2] \quad (19)$$

denote firm 1's equilibrium payoff in the presence of a linkage and

$$\Delta \Pi_1 = \mathbb{E}[\Pi^*(\theta_1, \lambda_1, \lambda_2)] - \mathbb{E}[\hat{\Pi}(\theta_1, \mu, \lambda_1)] \quad (20)$$

the value of a linkage for firm 1. We then show the following results.

Proposition 4 (First-Period Welfare Effects).

1. $\Delta U_1 > 0$ if and only if $\lambda_2 (\lambda_1 - h_{CS}(\lambda_2, \tilde{\sigma}_1)) \geq 0$, where $h_{CS}(\lambda_2, \tilde{\sigma}_1)$ is a continuous function that has the same sign as λ_2 for any $\tilde{\sigma}_1$. Moreover, ΔU_1 is decreasing in $\tilde{\sigma}_1$.
2. $\Delta \Pi_1 > 0$ if and only if $\lambda_2 > 0$.

Intuitively, firm 1 benefits from a data linkage if and only if the resulting change in consumer behavior increases the demand for its product. The effect on consumer surplus is slightly more involved.

Consistent with the above trade-offs, a linkage benefits consumers in the first period only if two conditions are met: first, the firms' efficiency levels λ_t must have the same sign so that

the terms of trade improve. Second, λ_1 must be sufficiently large in magnitude compared to λ_2 to ensure that the improvement in the terms of trade given by (16), which is proportional to λ_1 , offsets the negative effect of the distortion in consumption driven by λ_2 .

Finally, greater prior uncertainty $\tilde{\sigma}_1$ does not affect the amount of distortion in behavior, but it unambiguously worsens its impact on expected consumer surplus, so that the region where a linkage improves consumer surplus shrinks as $\tilde{\sigma}_1$ increases. To get an intuition for this, notice that the demand distortion, i.e. the difference between the quantities purchased with and without a linkage, is linear in the consumer type, because the consumer places a different weight on her type. Hence, the loss in consumer surplus due to this distortion is quadratic in the type. (This can be easily seen in Figure 3 where in both panels, the change in consumer surplus due to the purchase of a myopically sub-optimal quantity is represented by the area of the black triangle.) Since the loss is quadratic, it is a convex function of the type θ_1 , hence its expected value increases in the variance of θ_1 .⁶

Figure 4 illustrates the set of pairs (λ_1, λ_2) whose linkage is beneficial to consumers in the first period for $\varphi = 1$ and two different levels of $\tilde{\sigma}_1$.

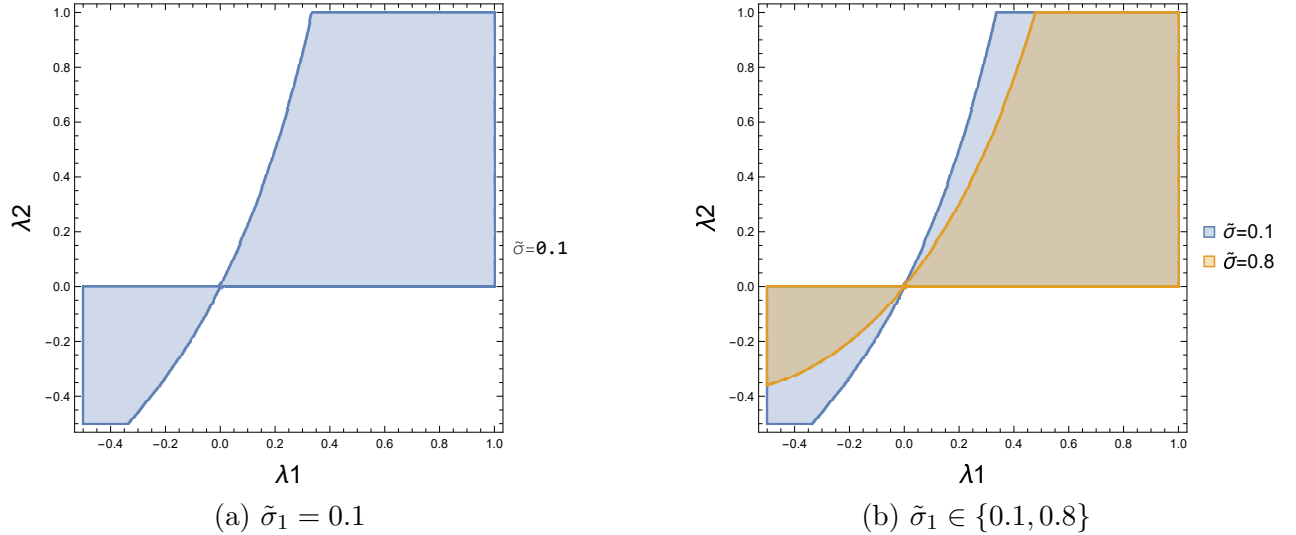


Figure 4: First-Period Consumer Surplus Improving Linkages

4.3 Intertemporal Welfare Effects of a Linkage

We now describe the total effects of a linkage on welfare. Proposition 5 characterizes the welfare impact of a data linkage across both periods. Figure 5 illustrates the sets of consumer- and firm-beneficial linkages (λ_1, λ_2) . We refer to these sets as Λ^{CS} and Λ^{PS} , respectively. In

⁶The social cost of data linkages also increases with $\tilde{\sigma}_1$ because the impact on profits is constant in $\tilde{\sigma}_1$.

what follows, let

$$\Delta U = \Delta U_1 + \Delta U_2 \quad (21)$$

denote the total value of a linkage for the consumer. Similarly, let

$$\Delta \Pi = \Delta \Pi_1 + \Delta \Pi_2 \quad (22)$$

denote the total value of a linkage for the two firms.

Proposition 5 (Intertemporal Welfare Effects).

1. $\Delta U > 0$ if and only if the following holds:

For $\lambda_2 > 0$, either $\lambda_1 - h_{CS}(\lambda_2, \tilde{\sigma}_1) \geq 0$, or $\lambda_1 - h_{CS}(\lambda_2, \tilde{\sigma}_1) \leq 0$ and $\tilde{\sigma}_1 \geq \bar{\sigma}_1(\lambda_1, \lambda_2, \varphi)$.

For $\lambda_2 < 0$, $\lambda_1 - h_{CS}(\lambda_2, \tilde{\sigma}_1) \leq 0$ and $\tilde{\sigma}_1 \leq \bar{\sigma}_1(\lambda_1, \lambda_2, \varphi)$.

2. A data linkage increases total firm profits if and only if either $\lambda_2 > 0$, or $\lambda_2 < 0$ and $\lambda_1 < h_{PS}(\lambda_2)$, where $h_{PS}(\lambda_2)$ is a continuous and increasing function of λ_2 .

For part 1, it is useful to compare Figure 5a to Figure 4a: the set of beneficial linkages appears to have shifted to the left. This has a simple intuition. First, consider positive values of λ_2 . We know from Proposition 3 that $\Delta U_2 > 0$. If λ_1 is sufficiently large that ΔU_1 is also positive, i.e. for the linkages in the region depicted in Figure 4a, the total value for the consumer is obviously positive. If instead λ_1 is moderately to the left of the threshold in Figure 4a, so that $\Delta U_1 < 0$, the total value of the linkage can still be positive, provided that the positive second-period effect dominates. This in turn requires a sufficiently large ex-ante uncertainty on the consumer's period-1 type. For negative values of λ_2 instead, $\Delta U_2 < 0$. Therefore, the total value of the linkage for the consumer is positive if and only if $\Delta U_1 > 0$ and the uncertainty is sufficiently small so that the negative second-period effect is offset. This explains why for $\lambda_2 < 0$ the region where the linkage value is positive in Figure 5a is a subset of the corresponding region in Figure 4a.

For part 2, recall that all firms benefit from linkages with $\lambda_2 > 0$. These linkages increase profits in the second period due to price discrimination and raise demand in the first period by means of the consumer's manipulation incentives. By continuity, for any λ_1 , there exists a threshold $\lambda_2 < 0$, above which even linkages that decrease first-period profits still increase total profits. Furthermore, recall that the distortion in a consumer's behavior is proportional to her average demand, which is an increasing function of λ_1 . Therefore, high- λ_1 firms suffer relatively more from linkages with $\lambda_2 < 0$ firms (i.e., their threshold λ_2 is higher) because the resulting downward distortion in consumer behavior is more costly for them.

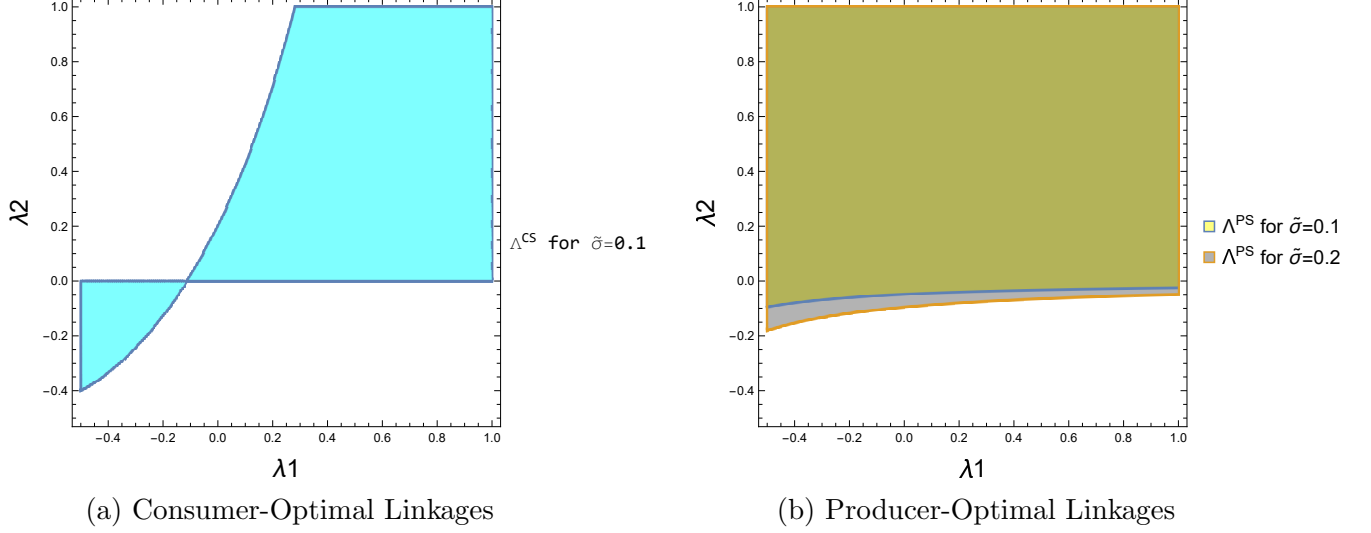


Figure 5: Consumer- and Producer-Optimal Linkages

5 The Impact of Privacy Regulation

Having characterized the welfare consequences of data linkages, we now turn to examine the impact of policies that regulate data governance. The last decade has seen a wave of new privacy laws, the most noticeable being the General Data Protection Regulation (GDPR) adopted by the EU in 2016 and the California Privacy Rights Act (CPRA) of 2020.

These regulatory interventions focus on three principles: *transparency*, *consent*, and *limits to discrimination*. The transparency principle establishes that a consumer must be made aware of the existence of a linkage. The consent requirement gives the consumer the right to veto the formation of a linkage, i.e., to stop the transfer of her data from one firm to another. Finally, limits to service, price, and quality discrimination prevent firms from penalizing consumers for denying consent to the formation of a linkage. We examine the implications of each of these principles for consumer surplus in the context of our model.

Throughout this section, we maintain several assumptions. First, firms cannot commit to the terms of trade before a data linkage is formed, i.e., they cannot induce the consumer to agree to the transfer of her data from firm 1 to firm 2 with the promise of lower prices or better products. Second, the consumer makes any decisions related to linkage formation before learning her realized willingness to pay. With this assumption, we capture the idea that each consumer visits a specific firm's website repeatedly and agrees or disagrees with its "terms of use" independent of her current-day inclination to shop. Third, the two firms bargain efficiently over the transfer of period 1 outcome data. For ease of exposition, we assume that firm 1 holds all the bargaining power, for example, because firm 1 is a large

digital platform. However, our results do not rely on a specific bargaining protocol.⁷

Our main results are as follows. We begin with the benchmark case of a fully unregulated market, where firms can freely establish linkages and cannot commit to maintaining the consumer’s privacy: in such an environment, all possible linkages are formed. We then examine the impact of transparency policies. We find that transparency benefits firms (but not necessarily consumers), by allowing firm 1 to commit to not sharing data with firm 2 when doing so would decrease the total producer surplus. Next, we consider the additional requirement of explicit consumer consent to the formation of a linkage. Consent rules differ according to the limits they impose on the discriminatory treatment of consumers who deny their consent, and this affects the impact they have on consumer surplus.

We examine three specific forms of consent rules: *required consent*, whereby a firm can commit to refuse to trade with a consumer who denies her consent to data transfer; *voluntary consent*, whereby a firm cannot commit to refuse to trade nor can it commit to arbitrary terms of trade that condition on the consent decision; and the *right to equal service and price*, whereby the firm cannot condition the transaction or the terms of trade on the consent decision at all. Although one could imagine that the strongest constraints on discrimination best protects consumers’ interests, we will show that once we take the equilibrium behavior of both firms and consumers into account, voluntary consent is the best form of regulation.

5.1 Unregulated Linkage Formation

We first consider an unregulated environment, in which firms have full control over linkage formation but lack commitment power: consumers have no legal right to request that their data not be shared and firm 1 has no credible way to commit not to share them.

In this scenario, the two firms contract on the formation of a linkage at the onset of the game. The formation of a linkage is unobservable to the consumer, who must therefore infer the outcome of this negotiation. In Proposition 6, we show that, for every pair (λ_1, λ_2) , a linkage is formed, and the market equilibrium of Proposition 2 is played.

Proposition 6 (No Regulation).

In the absence of privacy regulation, a data linkage forms for every pair (λ_1, λ_2) .

To capture the intuition for this result, suppose that for a pair (λ_1, λ_2) the consumer expects that her data will not be shared. At the onset of the game, the firms then have an

⁷Our analysis rules out fixed monetary payments for consumer consent. These payments are not used in practice, likely because of moral hazard concerns on the consumers’ side. Nonetheless, we characterize this benchmark in Appendix F. Likewise, we assume that a linkage is formed (or blocked) before the consumer learns her type, so that the consent decision has no signaling role. Appendix G validates this assumption by analyzing pooling equilibria in a game where consent decisions are made by informed consumers.

incentive to form the linkage: firm 1's profits would be unaffected, because the consumer would remain unaware of the linkage, and firm 2 would benefit from receiving the information, as shown in Proposition 3.

In terms of consumer surplus, the total absence of privacy is clearly problematic. Consumers would like a linkage to form if and only if $(\lambda_1, \lambda_2) \in \Lambda^{CS}$, which we characterized in Proposition 5. In terms of total producer surplus, some privacy would also be desirable: the overall effect of a linkage on total profits is positive only if $(\lambda_1, \lambda_2) \in \Lambda^{PS}$, which we also characterized in Proposition 5. Data sharing benefits firm 2 for all values of λ_2 , but it decreases firm 1's overall profits if λ_2 is sufficiently negative, even if firm 1 can extract the entire value of information from firm 2.

5.2 Transparency

Let them know precisely what you're going to do with their data. (Steve Jobs, All Things Digital Conference, 2010)

The *transparency* principle—that the consumer should be informed of how her personal data will be used and shared—is not only attractive strategic advice, it is a pillar of all recent privacy legislation. For example, GDPR §2.13.1 establishes that “*Where personal data relating to a data subject are collected (...) the controller shall provide the data subject with (...) the recipients or categories of recipients of the personal data.*” Similarly, CPRA §1798.110 establishes that “*(a) A consumer shall have the right to request that a business that collects personal information about the consumer disclose to the consumer...the categories of third parties to whom the business discloses personal information.*” In this scenario, consumers have no legal right to request that their data not be shared, but firm 1 has a credible way to commit to privacy if it finds it convenient.

In the context of our model, a law imposing transparency requires firm 1 to announce to the consumer whether it formed a linkage with firm 2 *before* the first-period interaction. Once again, the two firms bargain efficiently at the onset of the game, knowing that the consumer's demand function in period 1 will depend on the presence of a data linkage. Therefore, firm 1 internalizes the cost of ratchet forces when negotiating with firm 2. As a result, firms λ_1 and λ_2 agree to form a linkage if and only if it increases their total surplus, $(\lambda_1, \lambda_2) \in \Lambda^{PS}$, as in Proposition 5 and Figure 5.

Proposition 7 (Transparency).

Under transparency, a data linkage forms if and only if $(\lambda_1, \lambda_2) \in \Lambda^{PS}$.

Transparency increases total producer surplus because it allows firm 1 to commit to privacy whenever the losses from the ratchet effect in the first period exceed the gains

from the sale of information. The effect on consumer surplus depends on the two firms' efficiency levels. Requiring transparency improves consumer surplus compared to the absence of regulation by preventing the formation of those linkages that are harmful both to the consumer and (jointly) to the producers, i.e. $(\lambda_1, \lambda_2) \notin \Lambda^{PS} \cup \Lambda^{CS}$. However, transparency also prevents the formation of a beneficial linkage whenever $(\lambda_1, \lambda_2) \in \Lambda^{CS} \setminus \Lambda^{PS}$. Therefore, the welfare effect of transparency in isolation is ambiguous.

Most regulations, however, do not only force firms to “tell” consumers what will happen to their data, but also force them to “ask” the consumer for permission. Next, we examine the role of consent requirements.

5.3 Consent

A second principle that is currently well established in privacy protection legislation is that consumer data processing and sharing require the explicit consent of the consumer. For example, CPRA §1798.120 establishes that *“A consumer shall have the right, at any time, to direct a business that sells or shares personal information about the consumer to third parties not to sell or share the consumer’s personal information.”*

In the context of our model, this type of regulation grants the consumer “veto power” over the formation of a linkage. Because transparency already offers the firms an opportunity to prevent unprofitable linkages, the addition of consent gives both parties *de facto* veto rights.

Under this regulation, consumers have a right to privacy of their transaction data.⁸ However, the impact of rights to privacy on consumer welfare is subtle. First, the consumer must consider the cost of exercising her rights, which is measured by the impact of denying consent on the ensuing terms of trade. Second, firm 1 anticipates the consumer’s response and optimally chooses whether to request consent to the formation of a linkage, or simply guarantee the privacy of the transaction data. Thus, the consumer’s right to privacy impacts the prevailing information structure even when it is not exercised on the equilibrium path.

Having ruled out explicit monetary payments for consent, three scenarios are possible in decreasing order of the cost of denying consent:

1. Required Consent

Firm 1 can refuse to trade with a consumer who denies consent to linkage formation.

2. Voluntary Consent

Firm 1 cannot refuse to trade with a consumer who denies consent to linkage formation,

⁸Unlike in the case of transparency, where observing $q_1 = 0$ would be interpreted as a nil purchase level and result in the corresponding (linear) terms of trade in the second period, here denying consent allows the consumer to retain her private information in the second-period game.

but it can modify the terms of its offer, within some limits.

3. Right to Equal Service and Price

Firm 1 cannot refuse to trade with a consumer who denies consent, nor can it condition the terms of its offer on the consent decision.

In the first scenario, the consumer who denies consent pays the highest cost by completely losing the surplus from completing a transaction in the first market. In the second scenario, she can complete the transaction but possibly faces worse terms of trade. In the last scenario, her consent choice leaves her trading opportunities unaffected.

The complexity of this issue is reflected in the nuanced rules imposed by legislators in different countries and even across US states.⁹ For example, in Brazil and Japan, as well as in Nevada, the law requires individual consent for disclosure of personal data to a third party, but it does not prevent a firm from refusing service to a customer who denies consent to the sale of her data, as in the Required Consent scenario above. This is instead explicitly forbidden by the law in Canada and in several US states, including California and Texas. Nonetheless, under these regulations, if a consumer denies consent, firms are allowed to *“offer a different price, rate, level, or quality of goods or services to the consumer if that price or difference is reasonably related to the value provided to the business by the consumer’s data,”* as illustrated by CPRA §1798.125. In other words, under these laws the Required Consent scenario is illegal while Voluntary Consent is legal.

Even stronger restrictions are imposed by the EU GDPR and by the majority of state-level privacy laws in the US. These legislations fully embrace the Right to Equal Service and Price, explicitly forbidding firms from offering different prices or qualities to consumers based on their consent decision.

To contribute to the debate on optimal privacy regulation, we examine the implications on consumer surplus of each of these three types of consent rules. Do they improve on a simple transparency requirement? Which of them provides the highest consumer surplus? At first glance, one would conjecture that the most drastic form of regulation, the one establishing the right to equal service and price, is the most favorable to the consumer. We will show that once we take into account equilibrium behavior, this is not the case.

5.3.1 Required Consent

In the case of Required Consent, firm 1 can either commit to privacy, or ask for the consumer’s consent and commit to refusing trade if the latter is denied. Clearly, the firm will only ask

⁹See Appendix A for an exhaustive list of existing privacy laws.

for the consumer's consent if it expects to obtain it, else profits are nil, and if $(\lambda_1, \lambda_2) \in \Lambda^{PS}$, so that trading at the equilibrium terms (p_1^*, y_1^*) and selling the data is more profitable than committing to privacy and offering the privacy terms of trade.

In turn, when firm 1 asks for consent, the consumer correctly anticipates that if she denies it, she will only trade in period 2, without her period-1 type being inferred. If she does consent to the formation of a linkage, firm 1 will offer the equilibrium terms of trade under a data linkage, and firm 2 will make her a personalized offer.

Recall the consumer's first-period surplus, which includes the distortion and terms of trade effects, $U_1^*(\theta, \lambda_1, \lambda_2)$ defined in (17) and her second-period surplus $\hat{U}(\theta, m, \lambda_2)$ defined in (10). The consumer's participation constraint under required consent is given by

$$\mathbb{E}[U_1^*(\theta_1, \lambda_1, \lambda_2) + \hat{U}(\theta_2, (1 - \varphi)\mu + \varphi\theta_1, \lambda_2)] \geq \mathbb{E}[\hat{U}(\theta_2, \mu, \lambda_2)], \quad (23)$$

where the expectation is taken jointly over θ_1 and θ_2 .

Given the severity of the no-trade threat, the consumer consents to a linkage unless λ_1 is sufficiently small and λ_2 sufficiently large. Intuitively, if such linkages were formed, in period 1 she would have to distort her demand upwards a lot without the mitigating effect of a substantial improvement in the terms of trade. This would yield a negative surplus in period 1 that even the gain in surplus in period 2 would fail to compensate.

Proposition 8 summarizes the above discussion: a linkage is proposed, and consent is given, whenever $(\lambda_1, \lambda_2) \in \Lambda^{PS}$ and the consumer's participation constraint 23 is satisfied. We illustrate this set in Figure 6.

Proposition 8 (Required Consent).

When $\lambda_1 < 1$, a data linkage is formed if and only if $(\lambda_1, \lambda_2) \in \Lambda^{PS}$ and in addition,

$$\lambda_2 < h_{RC}^*(\lambda_1, \tilde{\sigma}_1, \varphi).$$

The threshold $h_{RC}^*(\lambda_1, \tilde{\sigma}_1, \varphi)$ is strictly increasing in $\tilde{\sigma}_1$, grows without bound as $\lambda_1 \rightarrow 1$, and satisfies $h_{RC}^* \geq 1/3$ for all $\lambda_1, \tilde{\sigma}_1$, and φ .

When $\lambda_1 \geq 1$, a data linkage is formed if and only if $(\lambda_1, \lambda_2) \in \Lambda^{PS}$.

We now ask whether this type of regulation improves consumer welfare compared to a simple transparency rule. With transparency, a linkage is formed if and only if $(\lambda_1, \lambda_2) \in \Lambda^{PS}$, namely, if and only if it improves producer surplus compared to privacy. With Required Consent, whenever the two firms propose a linkage, the consumer is given veto rights to protect her anonymity in the transaction with firm 2. However, the cost of that veto is the full surplus from transacting with firm 1. In equilibrium, the consumer never exercises

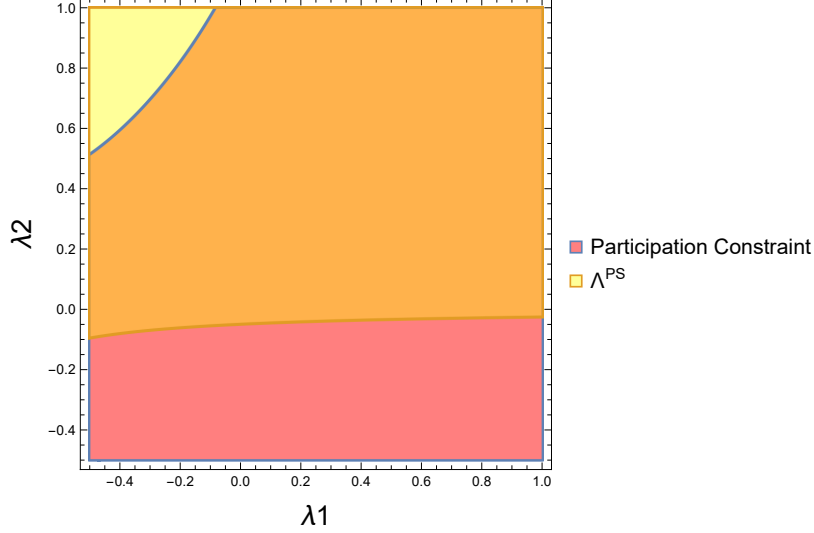


Figure 6: Linkages Formed under Required Consent

her veto rights. In the cases where she would, the firms refrain from proposing a linkage in the first place. Therefore, even this form of consent that is most costly for the consumers improves their surplus by preventing some of the linkages $(\lambda_1, \lambda_2) \in \Lambda^{PS}$ that are detrimental for consumer surplus and would be formed under a simple transparency rule (the yellow area in the top-left corner of Figure 6).

5.3.2 Voluntary Consent

We now consider the case of Voluntary Consent, where the law limits the burden that firm 1 can impose on consumers who deny consent to linkage formation. We capture these limitations by requiring the firm to serve the consumer regardless of her consent decision. Moreover, because firm 1 cannot commit to the terms of trade, equilibrium prices and quality levels will be given by the linear equilibria of Propositions 1 and 2, depending on whether or not a linkage forms.

Under these circumstances, the consumer will consent to the sharing of data if and only if $(\lambda_1, \lambda_2) \in \Lambda^{CS}$, that is, if the total value of the linkage ΔU is positive. In turn, firm 1 will ask for the consumer's consent only if it expects her to grant it, and if the linkage increases producer surplus, i.e., $(\lambda_1, \lambda_2) \in \Lambda^{PS}$. Therefore, as a result of mutual veto rights, only the linkages that constitute a Pareto improvement over anonymous trading will, in fact, be formed.¹⁰ We formalize this intuition in Proposition 9 and illustrate it in Figure 7

¹⁰In Appendix G, we analyze the case of informed consent decisions, restricting attention to pooling equilibria. Specifically, we derive conditions under which the equilibrium outcome of the game with uninformed consumers can be obtained as a pooling equilibrium of the game with informed consumers.

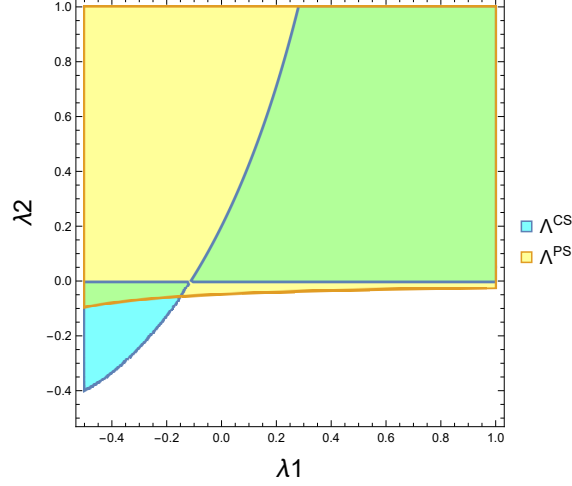


Figure 7: Linkages Formed under Voluntary Consent

Proposition 9 (Voluntary Consent).

If transparency and voluntary consent are required, a data linkage forms if and only if $(\lambda_1, \lambda_2) \in \Lambda^{PS} \cap \Lambda^{CS}$.

Protected by the service obligation, the consumer finds it optimal to deny consent for a wider set of (λ_1, λ_2) parameters. Therefore, the set of active linkages is a subset of the one obtained under required consent, and all the linkages that are removed are detrimental to consumer surplus. Therefore, voluntary consent improves consumer surplus, compared to required consent. Figure 8 compares the set Λ^{CS} of linkages that improve consumer surplus to the set of active linkages under required and voluntary consent, respectively. This clearly shows how voluntary consent offers a better approximation of Λ^{CS} .

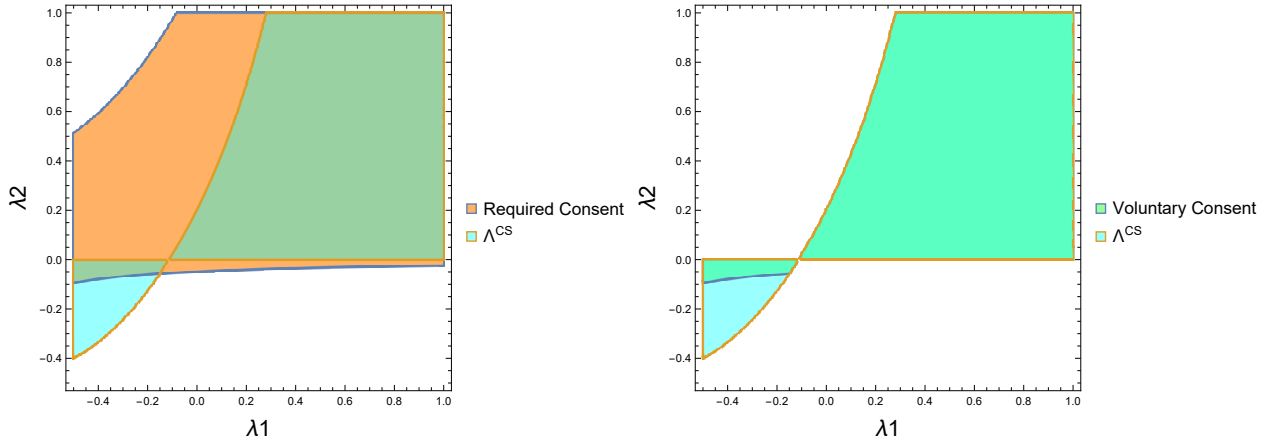


Figure 8: Required vs. Voluntary Consent

5.3.3 Right to Equal Service and Price

An even stronger form of privacy regulation allows firms to share transaction data if the consumer is made aware of the data sharing and explicitly consents to it, but forbids the firms from conditioning the terms of trade on the consent choice.

We model the right to equal service and price as a *simultaneous* decision of (p_1, y_1) by the firm and of consent by the consumer. (The firm can still offer a privacy guarantee to the consumer, so the game maintains mutual veto rights.) Making the consumer's consent decision unobservable to firm 1 rules out discriminatory behavior while maintaining the assumption that firm 1 does not commit to the terms of trade.¹¹

In this scenario, firm 1 must set the terms of trade in anticipation of whether the consumer will grant or deny consent. Therefore, if firm 1 anticipates that the consumer will grant her consent, it offers the terms of trade (p_1, y_1) as in the separating equilibrium (Proposition 2); otherwise, it offers the optimal terms of trade under privacy (Proposition 1).

In turn, the consumer chooses whether to consent to the formation of a data linkage taking firm 1's choice of terms of trade (p_1, y_1) as given. We show in Proposition 10 that the consumer prefers to grant consent if and only if λ_2 is positive but not too large.

Proposition 10 (Right to Equal Service and Price).

If transparency and consumer consent are required for data sharing and discrimination is forbidden, for any λ_1 a linkage is formed if and only if

$$\lambda_2 \in [0, t(\tilde{\sigma}_1^2)] , \quad (24)$$

where $t(\cdot)$ is a strictly increasing function.

Note that, because condition (24) requires $\lambda_2 \geq 0$, all linkages that satisfy this condition improve producer surplus, and hence firm 1 asks for the consumer's consent.

In the presence of this type of regulation, granting consent is always detrimental to consumer surplus in the first period. The reason is simple: when a data linkage is formed, the consumer distorts her demand away from the myopic optimum, but firm 1 cannot react by adjusting the terms of trade. Therefore, no compensating terms of trade effect can occur in the first period.¹²

¹¹It also has the realistic feature that most “consent boxes” appear on a webpage before the consumer can see the price for any product. Thus, the characterization of the equilibrium set of linkages in Proposition 10 below does not rely on this no-commitment assumption.

¹²This observation also explains why the terms of trade (p_1, y_1) do not impact the consumer's consent decision: their effect on the quantity purchased is independent of θ ; hence, they do not affect the distortion in quantity relative to the static optimum (which is given by $\lambda_2\theta$). An implication of this property is that commitment to the terms of trade would have no value for firm 1.

Why would the consumer then want to consent to the transmission of her data? If $\lambda_2 < 0$, information revelation would worsen her terms of trade in period 2, thereby reducing her period 2 surplus as well. Therefore, the consumer refuses consent for any $\lambda_2 < 0$, which explains the bottom half of Figure 9b. If instead $\lambda_2 > 0$, the consumer obtains a higher surplus in period 2 by granting consent, which can potentially offset the first-period loss. However, the consumer surplus “triangle” lost by distorting behavior is proportional to $(\lambda_2)^2$, while the value of a linkage in period 2 is proportional to λ_2 . Therefore, the consumer grants consent if λ_2 is positive but small. Finally, because the value of a linkage in period 2 is increasing in the prior uncertainty σ_1 , the threshold $t(\tilde{\sigma}_1^2)$ for granting consent is also increasing in $\tilde{\sigma}_1$.

How does consumer welfare under this policy compare to the outcome of the previous, less restrictive policies? Contrary to the common wisdom that discrimination allows predatory behavior, the comparison of the equilibrium set of linkages in Propositions 9 and 10 (see Figure 9) suggests the opposite.

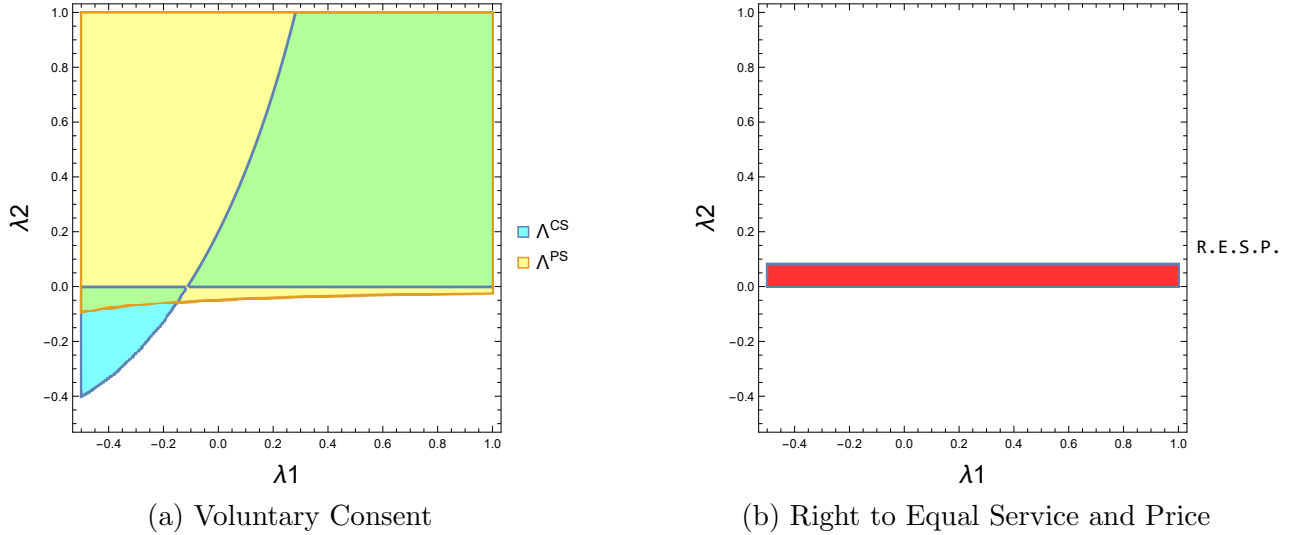


Figure 9: Voluntary Consent vs. Right to Equal Service and Price

In particular, for $\lambda_2 < 0$, no linkages form if discrimination is not allowed. However, for sufficiently negative λ_1 , the consumer would allow data sharing in exchange for better terms of trade, which is the equilibrium outcome under a voluntary consent policy. The same is true for $\lambda_2 > t(\tilde{\sigma}_1^2)$ and positive and sufficiently large λ_1 . Finally, for $0 < \lambda_2 < t(\tilde{\sigma}_1^2)$, firm 1 successfully proposes forming a linkage under right to equal service and price. When λ_1 is sufficiently negative, however, the consumer pays a higher price in the first period than she would under anonymity, if discrimination were allowed. In other words, she would be better off denying consent if that induced the equilibrium terms of trade under privacy.

We then draw a stark conclusion about adding the no-discrimination requirement to a policy that already requires the consumer’s explicit consent for data sharing.

Corollary 1 (Welfare Effects of Right to Equal Service and Price).

If transparency and consumer consent are required for data sharing, imposing equal terms of trade weakly damages consumers for all (λ_1, λ_2) , and strictly for some (λ_1, λ_2) .

Our results in this section inform the choice of the optimal price of anonymity, expressed in terms of the different terms of trade that emerge under alternative privacy choices. In this light, the optimal price of anonymity is limited, but it is not zero—maintaining some incentives for granting consent preserves the beneficial effects of the previous section.

5.4 Discussion

Our analysis has identified mandatory transparency and voluntary consent rules as the most favorable to consumer surplus. We have interpreted these rules as a form of mutual veto rights over linkages: for firms through transparency and for consumers through consent. These rules have two substantive implications for consumer surplus and social welfare. First, only Pareto-improving linkages will form, but some linkages that would increase social welfare will not. In Appendix F, we investigate a further scenario in which firms are allowed to offer a direct (positive or negative) payment to consumers so that all and only the linkages that improve social welfare are formed. At the moment, only the California CPRA allows firms to offer direct (positive) payments for consent.

Second, mandatory transparency and voluntary consent rules grant consumers the right to veto harmful data linkages, but they do not provide the power to impose the formation of new linkages. This prevents the creation of some linkages that would improve consumer surplus but harm producer surplus. The recent policy debate over “data portability” highlights the request from consumer advocates to give consumers the means to create beneficial linkages, not just to veto harmful ones. At present, major privacy laws like the GDPR and the CPRA only partially address this request by offering consumers the right to request that a firm share specific information with another firm, provided that this information was directly provided by the consumer in the first place.

6 Conclusions

We have provided a framework to assess the efficacy of different existing forms of privacy regulation. Our approach consists of developing a simple model of the equilibrium effects of

data linkages in the presence of privacy-conscious consumers. The model microfound a consumer’s preferences over the collection and transmission of behavior data by heterogeneous firms. The impact of data linkages on consumer surplus critically depends on the degree of similarity of the collecting and receiving firms in terms of how they respond to a perceived increase in demand. By endogenizing the formation of linkages under different types of privacy rules, our analysis shows how carefully designed transparency and consent requirements for linkage formation can benefit consumers. Voluntary Consent requirements provide firms with sufficient flexibility to *reward* consumers for granting consent, while curbing their power to *extort* the consumer’s consent by threatening to refuse service.

Our framework can easily be adapted to analyze a different set of policy issues. Mergers and acquisitions create data linkages that typically fall under different legislation from the linkages analyzed in this paper and raise concerns related to both competition and privacy. For example, by extending our analysis to allow for data collection and use by more than one firm, our model can be used to analyze the impact of a data-driven merger, such as the one between Google and Fitbit (Bourreau et al., 2020). Our framework can also be modified in various ways, allowing for different levels of consumer sophistication, idiosyncratic privacy preferences, and correlation between firms’ efficiency and the consumer’s taste parameter.

We have chosen to model linkages between commercial firms to analyze how existing privacy regulations affect consumer and producer surplus. Yet, the economic forces in our model apply to a broader class of linkages. The key mechanism is that a linkage gives a privacy-conscious consumer an incentive to distort her behavior away from the myopic optimum in order to influence how she is perceived in the future. This incentive arises whenever the consumer is concerned about a linkage between the first firm and a second entity she expects to interact with. Thus, the parameter λ_2 , which summarizes the nature of the second-period interaction and the direction of the distortion incentive, has a broader interpretation beyond market transactions with commercial firms.

For example, credit scoring algorithms create incentives for consumers to take out small, short-term installment loans to improve their future scores. Similarly, employee wellness programs can lead consumers to increase their engagement with fitness or health-tracking devices. Conversely, concerns about racial profiling or government surveillance can reduce demand for certain products. For instance, *The New York Times* reported that Chinese police used smartphone trackers to identify the presence of a Uyghur-to-Chinese dictionary app—an indicator that the phone likely belonged to a member of the heavily persecuted Uyghur minority. In this case, the use of trackers generates a data linkage, and individuals concerned about profiling have an incentive to avoid purchasing such apps.¹³

¹³<https://www.nytimes.com/2022/06/21/world/asia/china-surveillance-investigation.html>.

Appendix

A Consent Rules

We now summarize twenty-four existing privacy laws and the constraints they impose on how firms can respond when a consumer denies consent to a linkage. Laws that do not impose any restriction on the response are classified as Required Consent (R.C.) because firms are allowed to deny service to consumers who deny consent. Laws that forbid firms from denying service but allow them to charge a different price and/or offer a different quality are classified as Voluntary Consent (V.C.). Laws that forbid both denying service and conditioning price and quality on the consent choice are classified as Right to Equal Service and Price (R.E.S.P.).

Country / State	Law	Link	Consent Regime
Brazil	General Data Protection Law	https://iapp.org/resources/article/brazilian-data-protection-law-lgpd-english-translation/	R.C.
Japan	Act on the Protection of Personal Information	https://www.japaneselawtranslation.go.jp/en/laws/view/4241/en	R.C.
Nevada	Security and Privacy of Personal Information	https://www.leg.state.nv.us/NRS/NRS-603A.html	R.C.
Canada	Personal Information Protection and Electronic Documents Act	https://laws-lois.justice.gc.ca/eng/acts/p-8.6/FullText.html	V.C
California	California Privacy Rights Act	https://www.caprivacy.org/cpra-text/	V.C
Indiana	Indiana Consumer Data Protection Act	https://legiscan.com/IN/text/SB0005/id/2628665	V.C
Iowa	Iowa Consumer Data Protection Act	https://www.legis.iowa.gov/legislation/BillBook?ba=SF%20262&ga=90	V.C
Montana	Montana Consumer Data Privacy Act	https://leg.mt.gov/bills/2023/billpdf/SB0384.pdf	V.C
New Jersey	New Jersey Privacy Act	https://www.njleg.state.nj.us/bill-search/2022/S332	V.C
Tennessee	Tennessee Information Protection Act	https://legiscan.com/TN/text/HB1181/id/2672877	V.C
Texas	Texas Data Privacy and Security Act	https://capitol.texas.gov/tlodocs/88R/billtext/html/HB00004F.htm	V.C
Virginia	Virginia Consumer Data Protection Act	https://law.lis.virginia.gov/vacodefull/title59.1/chapter53/	V.C
EU	General Data Protection Regulation	https://gdpr-info.eu/	R.E.S.P.
Colorado	Colorado Privacy Act	https://leg.colorado.gov/sites/default/files/2021a_190_signed.pdf	R.E.S.P.
Connecticut	Connecticut Data Privacy Act	https://www.cga.ct.gov/2022/act/Pa/pdf/2022PA-00015-R00SB-00006-PA.PDF	R.E.S.P.
Delaware	Delaware Personal Data Privacy Act	https://legis.delaware.gov/BillDetail?LegislationId=140388	R.E.S.P.
Kentucky	Kentucky Consumer Data Protection Act	https://apps.legislature.ky.gov/record/24RS/hb15.html	R.E.S.P.
Maryland	Maryland Online Data Privacy Act	https://mgaleg.maryland.gov/mgawebsite/Legislation/Details/sb0541	R.E.S.P.
Minnesota	Minnesota Consumer Data Privacy Act	https://www.revisor.mn.gov/bills/bill.php?f=HF4757&b=house&y=2024&ssn=0	R.E.S.P.
Nebraska	Nebraska Data Protection Act	https://legiscan.com/NE/text/LB1074/2023	R.E.S.P.
New Hampshire	New Hampshire Privacy Act	https://legiscan.com/NH/text/SB255/id/2750381	R.E.S.P.
Oregon	Oregon Consumer Privacy Act	https://legiscan.com/OR/text/SB619/id/2830293	R.E.S.P.
Rhode Island	Rhode Island Data Transparency and Privacy Protection Act	https://legiscan.com/RI/text/H7787/2024	R.E.S.P.
Utah	Utah Consumer Privacy Act	https://le.utah.gov/~2022/bills/static/SB0227.html	R.E.S.P.

B Proofs of Propositions

Proof of Proposition 1. In a game without data linkages, the consumer's demand function is given by

$$q(\theta_t, p_t, y_t) = \theta_t + b_t y_t - p_t$$

in both periods $t = 1, 2$. Because $\mathbb{E}[\theta_1] = \mathbb{E}[\theta_2] = \mu$, each firm t 's profit can be written as

$$\mathbb{E}[\Pi(p_t, y_t, q_t)] = (p_t - k_t y_t)(\mu + b_t y_t - p_t) - c_t y_t^2 / 2.$$

This is a quadratic, concave function of p, y . Differentiating with respect to p_t and y_t and solving yields the expressions for $\hat{p}$ and $\hat{y}$ in (6) and (7), respectively. $\square$

Proof of Proposition 2. We begin by proving that in any linear equilibrium, the consumer's first-period demand function is linear in θ_1, p_1 and y_1 . Then, we show that such an equilibrium exists and is unique.

Fix a linear strategy by firm 2, i.e., two linear functions $p_2(q_1, p_1, y_1)$ and $y_2(q_1, p_1, y_1)$, and consider the consumer's first-period best-response problem,

$$\max_{q_1} \left[U(\theta_1, p_1, y_1, q_1) + \mathbb{E} \left[\max_{q_2} [U(\theta_2, p_2(q_1, p_1, y_1), y_2(q_1, p_1, y_1), q_2)] \mid \theta_1 \right] \right]. \quad (25)$$

Recall that the function $U(\theta, p, y, q)$ is quadratic in q and linear in p and y . Therefore, by the Envelope Theorem and the linearity of the expectations operator, we obtain that the consumer's first-period best reply q_1 to a linear strategy (p_2, y_2) is linear in θ_1, p_1 , and y_1 . The consumer's strategy in any linear equilibrium can then be written as

$$q_1 = \alpha \theta_1 + \beta y_1 + \gamma p_1 + \delta. \quad (26)$$

Upon observing the realized q_1 , firm 2 forms a degenerate posterior belief over the consumer's first- and second-period type,

$$m_1(q_1, p_1, y_1) = \frac{q_1 - \beta y_1 - \gamma p_1 - \delta}{\alpha} \quad (27)$$

$$m_2(q_1, p_1, y_1) = (1 - \varphi) \mu + \varphi \frac{q_1 - \beta y_1 - \gamma p_1 - \delta}{\alpha}. \quad (28)$$

If firm 1 conjectures the demand as in (26), it maximizes its expected profits,

$$\mathbb{E}[\Pi_1] = (p_1 - k_1 y_1) (\alpha \mu + \beta y_1 + \gamma p_1 + \delta) - \frac{c_1 y_1^2}{2}.$$

The first-order conditions for the firm's problem with respect to (p_1, y_1) are given by

$$\begin{aligned}(p_1 - k_1 y_1) \beta - c_1 y_1 - k_1 (\alpha \mu + \beta y_1 + \gamma p_1 + \delta) &= 0, \\ 2\gamma p_1 + y_1 (\beta - k_1 \gamma) + \delta + \alpha \mu &= 0.\end{aligned}$$

The first-period firm's optimal choices of price and quality are given by

$$p_1^* = -\frac{c_1 \alpha \mu + k_1 \mu \alpha \beta + k_1^2 \mu \alpha \gamma + c_1 \delta + k_1 \beta \delta + k_1^2 \gamma \delta}{2c_1 \gamma + (\beta + k_1 \gamma)^2} \quad (29)$$

$$y_1^* = -\frac{\mu \alpha \beta + k_1 \mu \alpha \gamma + \beta \delta + k_1 \gamma \delta}{2c_1 \gamma + (\beta + k_1 \gamma)^2}. \quad (30)$$

Next, we solve the consumer's problem and derive the equilibrium values of the coefficients of her demand. The consumer maximizes (12), i.e., the sum of her current flow surplus U_1 and her expected second-period surplus. Because $U_2 = \hat{U}(\theta_2, m_2(q_1), \lambda_2)$ in equilibrium, the consumer solves the following problem:

$$\max_{q_1} [U_1(\theta_1, p_1, y_1, q_1) + \mathbb{E}[\hat{U}(\theta_2, m_2(q_1, p_1, y_1), \lambda_2) \mid \theta_1]],$$

which can be rewritten as

$$\max_{q_1} \left[(\theta_1 + b_1 y_1 - p_1) q_1 - \frac{q_1^2}{2} + \mathbb{E}[(\varphi \theta_1 + \mu(1 - \varphi) + \varepsilon_2 + \lambda_2 m_2(q_1, p_1, y_1))^2 / 2] \right].$$

Using (28) above, the first-order condition with respect to q_1 can be written as

$$\theta_1 + b_1 y_1 - p_1 - q_1 + \frac{\lambda_2 \varphi}{\alpha} [\varphi \theta_1 + (1 - \varphi) \mu + \lambda_2 ((1 - \varphi) \mu + \frac{\varphi(q_1 - \beta y_1 - \gamma p_1 - \delta)}{\alpha})] = 0.$$

Solving for q_1 yields the following linear function of (θ_1, y_1, p_1) :

$$q_1^* = \frac{(\theta_1 + b_1 y_1 - p_1) + \frac{\lambda_2 \varphi}{\alpha} (\varphi \theta_1 + (1 - \varphi)(1 + \lambda_2) \mu) - \frac{\lambda_2^2 \varphi^2}{\alpha^2} (\beta y_1 + \gamma p_1 + \delta)}{1 - \frac{\lambda_2^2 \varphi^2}{\alpha^2}}.$$

Matching the coefficients to those in (26), we obtain a unique solution to the resulting system

of linear equations, which pins down the equilibrium strategies:

$$\begin{aligned}\alpha^* &= \frac{1}{2}(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1); \quad \beta^* = b_1; \\ \gamma^* &= -1; \quad \delta^* = \frac{2\lambda_2(\lambda_2 + 1)(1 - \varphi)\varphi\mu}{1 + \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1}}.\end{aligned}$$

Substituting into conditions (26), (29) and (30) yields the equilibrium strategies in parts (1) and (2) of the statement. Part (3) follows from the fact that $\mathbb{E}[\theta_2 \mid \theta_1] = (1 - \varphi)\mu + \varphi\theta_1$. $\square$

Proof of Proposition 3. By Proposition 1, replacing μ with m , i.e., firm 2's expectation of θ_2 , the realized surplus, profits, and welfare levels in period 2 can be written as

$$\begin{aligned}\hat{U}(\theta_2, m, \lambda_2) &= \frac{(\theta_2 + m\lambda_2)^2}{2} \\ \hat{\Pi}(\theta_2, m, \lambda_2) &= \frac{m(2\theta_2 - m)(1 + \lambda_2)}{2} \\ \hat{W}(\theta_2, m, \lambda_2) &= \frac{(\theta_2 + m\lambda_2)^2}{2} + \frac{m(2\theta_2 - m)(1 + \lambda_2)}{2}.\end{aligned}$$

Taking expectations over (θ_2, m) , we obtain the consumer's ex-ante welfare, the firm's profits, and total welfare. If the second-period firm learns θ_1 perfectly, the expected consumer surplus is given by

$$\begin{aligned}\mathbb{E}[\hat{U}(\theta_2, (1 - \varphi)\mu + \varphi\theta_1, \lambda_2)] &= \frac{1}{2}\mathbb{E}[(1 - \varphi)\mu + \varphi(\mu + \varepsilon_1) + \varepsilon_2 + \lambda_2((1 - \varphi)\mu + \varphi(\mu + \varepsilon_1))]^2 \\ &= \frac{1}{2}[\mu^2(1 + \lambda_2)^2 + \varphi^2(1 + \lambda_2)^2\sigma_1^2 + \sigma_2^2]\end{aligned}$$

If the firm has access only to the prior information, i.e., $m = \mu$, the expected consumer surplus is given by

$$\begin{aligned}\mathbb{E}[\hat{U}(\theta_2, \mu, \lambda_2)] &= \frac{1}{2}\mathbb{E}[(1 - \varphi)\mu + \varphi\mu + \varphi\varepsilon_1 + \varepsilon_2 + \lambda_2\mu]^2 \\ &= \frac{1}{2}[\mu^2(1 + \lambda_2)^2 + \varphi^2\sigma_1^2 + \sigma_2^2],\end{aligned}$$

so the difference in expected consumer surplus is

$$\begin{aligned}\Delta U_2 &= \mathbb{E}[\hat{U}(\theta_2, (1 - \varphi)\mu + \varphi\theta_1, \lambda_2)] - \mathbb{E}[\hat{U}(\theta_2, \mu, \lambda_2)] \\ &= \frac{1}{2}[\varphi^2(1 + \lambda_2)^2\sigma_1^2 - \varphi^2\sigma_1^2] \\ &= \frac{1}{2}\varphi^2\sigma_1^2\lambda_2(\lambda_2 + 2)\end{aligned}\tag{31}$$

which is positive if and only if $\lambda_2 > 0$.

If the second-period firm learns θ_1 , the realized and expected profits are given by:

$$\begin{aligned}\hat{\Pi}(\theta_2, (1 - \varphi)\mu + \varphi\theta_1, \lambda_2) &= \frac{(1 + \lambda_2)}{2} [(1 - \varphi)\mu + \varphi\mu + \varphi\varepsilon_1] [(1 - \varphi)\mu + \varphi\mu + \varphi\varepsilon_1 + 2\varepsilon_2] \\ &= \frac{(1 + \lambda_2)}{2} \left\{ [\mu + \varphi\varepsilon_1]^2 + 2\varepsilon_2 [\mu + \varphi\varepsilon_1] \right\} \\ \mathbb{E}[\hat{\Pi}(\theta_2, (1 - \varphi)\mu + \varphi\theta_1, \lambda_2)] &= \frac{(1 + \lambda_2)}{2} (\mu^2 + \varphi^2\sigma_1^2)\end{aligned}$$

Conversely, if the firm has access to the prior information only, then $m = \mu$, and

$$\begin{aligned}\hat{\Pi}(\theta_2, \mu, \lambda_2) &= \frac{(1 + \lambda_2)}{2} \mu [2(1 - \varphi)\mu + 2\varphi\mu + 2\varphi\varepsilon_1 + 2\varepsilon_2 - \mu], \\ &= \frac{(1 + \lambda_2)}{2} [\mu^2 + 2\mu\varphi\varepsilon_1 + 2\mu\varepsilon_2], \\ \mathbb{E}[\hat{\Pi}(\theta_2, \mu, \lambda_2)] &= \frac{(1 + \lambda_2)}{2} \mu^2.\end{aligned}$$

The change in firm 2's profits is given by

$$\Delta\Pi_2 = \mathbb{E}[\hat{\Pi}(\theta_2, (1 - \varphi)\mu + \varphi\theta_1, \lambda_2)] - \mathbb{E}[\hat{\Pi}(\theta_2, \mu, \lambda_2)] = \frac{(1 + \lambda_2)}{2} \varphi^2\sigma_1^2 > 0.$$

Finally, the change in social welfare is given by

$$\begin{aligned}\Delta W_2 &\triangleq \Delta U_2 + \Delta\Pi_2 = \frac{\varphi^2\sigma_1^2\lambda_2(\lambda_2 + 2)}{2} + \frac{(1 + \lambda_2)\varphi^2\sigma_1^2}{2} \\ &= \frac{\varphi^2\sigma_1^2}{2} (\lambda_2^2 + 3\lambda_2 + 1).\end{aligned}$$

Because $\lambda_2 \geq -1/2$, a data linkage improves social welfare for all

$$\lambda_2 \geq \underline{\lambda}_2 \triangleq -(3 - \sqrt{5})/2,$$

which is strictly negative. This completes the proof. $\square$

Proof of Proposition 4. (1.) Using Proposition 2, to calculate the consumer's first-period equilibrium quantity with a data linkage

$$\begin{aligned}q_1(\theta_1, p_1^*, y_1^*) &= \frac{\theta_1 \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1 \right)}{2} + \frac{\lambda_1\mu \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 2\varphi - 1 \right)}{2\varphi} \\ &\quad + \frac{2\lambda_2(\lambda_2 + 1)\mu(1 - \varphi)\varphi}{1 + \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1}},\end{aligned}$$

we derive her expected first-period equilibrium payoff:

$$\begin{aligned} \mathbb{E}[U_1^*(\theta_1, \lambda_1, \lambda_2)] = & \frac{\sigma_1^2 \left(-2\lambda_2(\lambda_2 + 1)\varphi^2 + \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1 \right)}{4} \\ & + \frac{(\lambda_1 + 1)\mu^2 \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 2\lambda_2(\lambda_2 + 1)\varphi + 1 \right)}{2 \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1 \right)^2} \\ & \cdot \left(\lambda_1 \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 2\lambda_2(\lambda_2 + 1)\varphi + 1 \right) \right. \\ & \left. + \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} - 2\lambda_2(\lambda_2 + 1)\varphi + 1 \right). \end{aligned} \quad (32)$$

Let $\tilde{\sigma}_1 \triangleq \sigma_1/\mu$. The first-period effect of a linkage is therefore equal to

$$\begin{aligned} \Delta U_1 &= \mathbb{E}[U_1^*(\theta_1, \lambda_1, \lambda_2)] - \mathbb{E}[\hat{U}(\theta_1, \mu, \lambda_1)] \\ &= \frac{2\lambda_2(\lambda_2 + 1)\mu^2\varphi}{\left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1 \right)^2} f(\lambda_1, \lambda_2, \varphi, \tilde{\sigma}_1), \end{aligned} \quad (33)$$

where

$$\begin{aligned} f(\lambda_1, \lambda_2, \varphi, \tilde{\sigma}_1) &\triangleq \lambda_1^2 \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + \lambda_2(\lambda_2 + 1)\varphi + 1 \right) \\ &+ \lambda_1 \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1 \right) - \lambda_2(\lambda_2 + 1)\varphi (\varphi^2 \tilde{\sigma}_1^2 + 1) \end{aligned} \quad (34)$$

is a convex quadratic expression in λ_1 .

Since $2\lambda_2(\lambda_2 + 1)\mu^2\varphi / \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1 \right)^2$ has the same sign as λ_2 , ΔU_1 is positive if λ_2 and $f(\lambda_1, \lambda_2, \varphi, \tilde{\sigma}_1)$ have the same sign. It can be easily verified that the smaller root of (34) is smaller than $-1/2$, hence outside the acceptable range for λ_1 . Therefore, the region where $\Delta U_1 \geq 0$ is characterized by

$$\lambda_2 (\lambda_1 - h_{CS}(\lambda_2, \tilde{\sigma}_1)) \geq 0,$$

where we denote the largest root λ_1 of $f(\lambda_1, \lambda_2, \varphi, \tilde{\sigma}_1)$ by

$$\begin{aligned} h_{CS}(\lambda_2, \tilde{\sigma}_1) &:= -\frac{1 + \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1}}{2 \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + \lambda_2(\lambda_2 + 1)\varphi + 1 \right)} \\ &+ \frac{\sqrt{\left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1 \right)^2 + 4\lambda_2(\lambda_2 + 1)\varphi \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + \lambda_2(\lambda_2 + 1)\varphi + 1 \right) (\varphi^2 \tilde{\sigma}_1^2 + 1)}}{2 \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + \lambda_2(\lambda_2 + 1)\varphi + 1 \right)}. \end{aligned} \quad (35)$$

Notice that $h_{CS}(0, \tilde{\sigma}_1) = 0$. Moreover, the denominator of (35) is positive for all $\lambda_2 > -1/2$, and the numerator is positive if and only if $\lambda_2 > 0$. Therefore, $h_{CS}(\lambda_2, \tilde{\sigma}_1)$ has the same sign as λ_2 . Finally, ΔU_1 is decreasing in $\tilde{\sigma}_1^2$ by inspection of (33) and (34).

(2.) For producer surplus, Proposition 2 establishes that expected profits with and without a linkage are given by

$$\mathbb{E}[\Pi^*(\theta_1, \lambda_1, \lambda_2)] = \frac{(\lambda_1 + 1)\mu^2 \left(2\lambda_2^2\varphi + \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 2\lambda_2\varphi + 1 \right)^2}{2 \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1 \right)^2}$$

and

$$\mathbb{E} \left[\hat{\Pi}(\theta_1, \mu, \lambda_1) \right] = \frac{\mu^2}{2} (1 + \lambda_1),$$

respectively. The difference between these two is equal to

$$\frac{2(\lambda_1 + 1)\lambda_2(\lambda_2 + 1)\mu^2\varphi \left(\lambda_2^2\varphi + \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + \lambda_2\varphi + 1 \right)}{\left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1 \right)^2},$$

and has the same sign as λ_2 . □

Proof of Proposition 5. (1.) Summing the changes in consumer surplus across the two periods in equations (31) and (33), the total effect of a data linkage is proportional to

$$\lambda_2 \left[(\lambda_2 + 2)\varphi\tilde{\sigma}_1^2 + \frac{\lambda_2 + 1}{\left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1 \right)^2} f(\lambda_1, \lambda_2, \varphi, \tilde{\sigma}_1) \right] \quad (36)$$

First, consider the case $\lambda_2 \geq 0$. If $\lambda_1 \geq h_{CS}(\lambda_2, \tilde{\sigma}_1)$, so that $f(\lambda_1, \lambda_2, \varphi, \tilde{\sigma}_1) \geq 0$, expression (36) is positive for any value of $\tilde{\sigma}_1$. If instead $\lambda_1 \leq h_{CS}(\lambda_2, \tilde{\sigma}_1)$, so that $f(\lambda_1, \lambda_2, \varphi, \tilde{\sigma}_1) \leq 0$ we rewrite the term in square brackets in (36) as

$$\begin{aligned} & \tilde{\sigma}_1^2 \frac{2\lambda_2\varphi^2 + \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} - 1}{2\lambda_2\varphi} \\ & + \frac{4(\lambda_1 + 1)(\lambda_2 + 1) \left(\lambda_1 \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + \lambda_2(\lambda_2 + 1)\varphi + 1 \right) - \lambda_2(\lambda_2 + 1)\varphi \right)}{\left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1 \right)^2}. \end{aligned} \quad (37)$$

Observe that, since the coefficient of $\tilde{\sigma}_1^2$ is positive, the whole expression (37), and hence

the total effect of a data linkage, is positive iff $\tilde{\sigma}_1 \geq \bar{\sigma}_1(\lambda_1, \lambda_2, \varphi)$, where

$$\bar{\sigma}_1^2(\lambda_1, \lambda_2, \varphi) \triangleq - \frac{2(\lambda_1 + 1)(\lambda_2 + 1) \left(\lambda_1 \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + \lambda_2(\lambda_2 + 1)\varphi + 1 \right) - \lambda_2(\lambda_2 + 1)\varphi \right)}{\varphi \left(2\lambda_2(\lambda_2 + 1)\varphi^2 + \lambda_2 \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 2\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + \lambda_2 + 2 \right)}. \quad (38)$$

Next, consider the case $\lambda_2 \leq 0$. If $\lambda_1 \geq h_{CS}(\lambda_2, \tilde{\sigma}_1)$, so that $f(\lambda_1, \lambda_2, \varphi, \tilde{\sigma}_1) \geq 0$, expression (36) is negative for any value of $\tilde{\sigma}_1$. Therefore, a necessary condition for a positive total change in consumer surplus is that $\lambda_1 \leq h_{CS}(\lambda_2, \tilde{\sigma}_1)$, so that $f(\lambda_1, \lambda_2, \varphi, \tilde{\sigma}_1) \leq 0$. Consider expression (37) The coefficient of $\tilde{\sigma}_1^2$ is positive even for $\lambda_2 \leq 0$, hence the necessary and sufficient condition for a positive effect on total consumer surplus is that $\tilde{\sigma}_1^2 \leq \bar{\sigma}_1^2(\lambda_1, \lambda_2, \varphi)$.

(2.) Similarly, summing the changes in profits in the two periods, the overall effect of the introduction of a linkage on the firms' profits is

$$\Delta\Pi = \frac{2(\lambda_2 + 1)\varphi(\lambda_1 + 1)\lambda_2\mu^2 \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + \lambda_2(\lambda_2 + 1)\varphi + 1 \right)}{\left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1 \right)^2} + \frac{1}{2}(\lambda_2 + 1)\sigma_1^2\varphi^2, \quad (39)$$

which has the same sign as

$$\frac{4(\lambda_1 + 1)\lambda_2 \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + \lambda_2(\lambda_2 + 1)\varphi + 1 \right)}{\left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1 \right)^2} + \tilde{\sigma}_1^2\varphi.$$

This expression is positive for $\lambda_2 \geq 0$. Moreover, if λ_2 is negative, the expression is positive for $\lambda_1 < h_{PS}(\lambda_2)$ where $h_{PS}(\lambda_2)$ is the following increasing function of λ_2 ,

$$h_{PS}(\lambda_2) := - \frac{\varphi\tilde{\sigma}_1^2 \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1 \right)^2}{4\lambda_2 \left(\sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + \lambda_2(\lambda_2 + 1)\varphi + 1 \right)} - 1, \quad (40)$$

which establishes the result. $\square$

Proof of Proposition 6. Linkage formation in the absence of privacy regulation is the outcome of a game wherein the first-period firm offers terms of trade to the consumer and simultaneously negotiates with firm 2 over the transfer of the first-period outcome data.

We now claim that there exists a unique equilibrium of this game, in which a data linkage forms for every pair (λ_1, λ_2) . Suppose, towards a contradiction, that there exists a pair (λ_1, λ_2) for which a linkage does not form with probability one. If so, not offering a linkage to firm 2 is a best response for firm 1. Now consider firm 1's alternative strategy

to offer the formation of a linkage to firm 2, while keeping the terms of trade (p_1, y_1) with the consumer unchanged. The transfer of data generates strictly positive value for firm 2 (see Proposition 3) and imposes no costs on firm 1, because the consumer cannot detect this deviation and therefore modify her purchasing behavior. Under our assumption of efficient trade over data, the two firms will then agree to a price and form the linkage. $\square$

Proof of Proposition 7. Under transparency regulation, we claim that all linkages that jointly benefit the two firms, and only those linkages, will form in the unique equilibrium. Indeed, consider a pair of firms with $(\lambda_1, \lambda_2) \in \Lambda^{PS}$. The firms' joint profits are larger under a *commonly known* data linkage than under privacy. Because the firms bargain efficiently, firm 1 can make an acceptable offer to firm 2 over the formation of a linkage that nonetheless increases firm 1's expected profits over the privacy benchmark. Conversely, if $(\lambda_1, \lambda_2) \notin \Lambda^{PS}$, no such profitable and acceptable offer exists, and the firms will not form a data linkage. $\square$

Proof of Proposition 8. Using expressions (10) and (32) for the consumer's equilibrium surplus with and without a linkage, the consumer's value of consenting to a linkage (23) is

$$\begin{aligned} & \frac{\mu^2}{4} \tilde{\sigma}_1^2 \left(1 + 2\lambda_2 \varphi^2 + \sqrt{1 + 4\lambda_2(1 + \lambda_2)\varphi^2} \right) + \frac{\mu^2}{\left(1 + \sqrt{1 + 4\lambda_2(1 + \lambda_2)\varphi^2} \right)^2} \\ & (1 + \lambda_1) \left(1 - 2\lambda_2(1 + \lambda_2)(-1 + \lambda_2 + \lambda_2^2)\varphi^2 + \sqrt{1 + 4\lambda_2(1 + \lambda_2)\varphi^2} \right. \\ & \left. + \lambda_1 \left(1 + \sqrt{1 + 4\lambda_2(1 + \lambda_2)\varphi^2} + 2\lambda_2(1 + \lambda_2)\varphi \left(1 + (1 + \lambda_2 + \lambda_2^2)\varphi + \sqrt{1 + 4\lambda_2(1 + \lambda_2)\varphi^2} \right) \right) \right). \end{aligned}$$

Because the coefficient on $\tilde{\sigma}_1^2$ is strictly positive, we can rewrite the consumer's participation constraint as

$$4(1 + \lambda_1)g(\lambda_1, \lambda_2) + \tilde{\sigma}_1^2 \geq 0, \quad (41)$$

where

$$\begin{aligned} g(\lambda_1, \lambda_2) := & \frac{1 - 2\lambda_2(1 + \lambda_2)(-1 + \lambda_2 + \lambda_2^2)\varphi^2 + \sqrt{1 + 4\lambda_2(1 + \lambda_2)\varphi^2} + \lambda_1 \left(1 + \sqrt{1 + 4\lambda_2(1 + \lambda_2)\varphi^2} \right)}{\left(1 + \sqrt{1 + 4\lambda_2(1 + \lambda_2)\varphi^2} \right)^2 \left(1 + 2\lambda_2\varphi^2 + \sqrt{1 + 4\lambda_2(1 + \lambda_2)\varphi^2} \right)} \\ & + \frac{2\lambda_2(1 + \lambda_2)\varphi \left(1 + (1 + \lambda_2 + \lambda_2^2)\varphi + \sqrt{1 + 4\lambda_2(1 + \lambda_2)\varphi^2} \right)}{\left(1 + \sqrt{1 + 4\lambda_2(1 + \lambda_2)\varphi^2} \right)^2 \left(1 + 2\lambda_2\varphi^2 + \sqrt{1 + 4\lambda_2(1 + \lambda_2)\varphi^2} \right)}. \end{aligned}$$

It is then immediate to establish the following properties of the function g . First if $\lambda_1 > 1$, then $g(\lambda_1, \lambda_2) > 0$ and (41) holds. Conversely, if $\lambda_1 < 1$, then

$$g(\lambda_1, \lambda_2) \leq 0 \iff \lambda_2 \geq -\frac{1}{2} + \frac{\sqrt{5 + 6\lambda_1 + 5\lambda_1^2 + \frac{4-4\lambda_1^2}{\varphi}}}{2(1 - \lambda_1)}, \quad (42)$$

which is strictly increasing in λ_1 on $[-1/2, 1]$ and strictly decreasing in φ . In particular, note that $g(\lambda_1, \lambda_2) < 0$ implies $\lambda_2 > 1/3$. Second, one can show that the cross-partial derivative $g_{\lambda_1 \lambda_2}(\lambda_1, \lambda_2)$ is positive whenever $\lambda_2 > 0$. Thus, the partial derivative $g_{\lambda_2}(\lambda_1, \lambda_2)$ is increasing in λ_1 over the region where $g(\lambda_1, \lambda_2) \leq 0$. Moreover, solving the inequality in (42) for λ_1 and substituting in the derivative $g_{\lambda_2}(\lambda_1, \lambda_2)$ yields

$$g_{\lambda_2}(\lambda_1, \lambda_2) \leq -\frac{2(1 + 2\lambda_2)\varphi \left(1 + 2\lambda_2(1 + \lambda_2)\varphi^2 + \sqrt{1 + 4\lambda_2(1 + \lambda_2)\varphi^2}\right)}{\sqrt{1 + 4\lambda_2(1 + \lambda_2)\varphi^2}} < 0,$$

wherever $g(\lambda_1, \lambda_2) \leq 0$. Finally, for any $\lambda_1 < 1$, we have

$$\lim_{\lambda_2 \rightarrow \infty} g(\lambda_1, \lambda_2) = -\infty.$$

Combining these results, the consumer's participation constraint (41) is always satisfied when $\lambda_1 \geq 1$. When $\lambda_1 < 1$, the left-hand side of (41) is strictly positive for all $\tilde{\sigma}_1 > 0$ when λ_2 is equal to the threshold in (42), strictly decreasing in λ_2 , and strictly negative as λ_2 grows without bound. This establishes the desired threshold property. $\square$

Proof of Proposition 9. Under voluntary consent, both transparency and consumer agreement are required for the formation of a linkage. Further, because neither firm can commit to the terms of trade in advance, the game in periods 1 and 2 will be played under privacy or a data linkage, as in the baseline model.

Therefore, both the consumer and the coalition of the two firms can secure the equilibrium expected payoff under privacy by refusing or not proposing a linkage, respectively. It then follows that a data linkage forms if and only if it constitutes a Pareto improvement, i.e., $(\lambda_1, \lambda_2) \in \Lambda^{PS} \cap \Lambda^{CS}$. $\square$

Proof of Proposition 10. Fix the terms of trade offered by the first-period firm. If the consumer gives consent, her demand in period 1 is given by equation (13) and her expected surplus in the first period is given by

$$\begin{aligned} \mathbb{E}[U_1(\theta_1, p_1, y_1, q_1^*)] &= \frac{-4\mu(p_1 - b_1 y_1) + 2(p_1 - b_1 y_1)^2}{4} \\ &\quad + \frac{\mu^2(-2(\lambda_2^2 + \lambda_2 - 1)\varphi^2 + \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} - 1)}{4\varphi^2} \\ &\quad + \frac{\sigma_1^2(-2\lambda_2(\lambda_2 + 1)\varphi^2 + \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1)}{4}. \end{aligned}$$

In the second period, the consumer then obtains expected surplus $\mathbb{E}[\hat{U}(\theta_2, \theta_1, \lambda_2)]$. If

instead she denies consent, in the first period she consumes

$$q(\theta_1, p_1, y_1, b_1) = \theta_1 + b_1 y_1 - p_1$$

and obtains expected surplus

$$\mathbb{E}[U_1] = \mathbb{E}[(\theta_1 + b_1 y_1 - p_1)^2 / 2] = ((\mu + b_1 y_1 - p_1)^2 + \sigma_1^2) / 2,$$

and her second-period surplus is given by $\mathbb{E}[\hat{U}(\theta_2, \mu, \lambda_2)]$.

Simplifying the above expressions, the difference in payoff in period 1 is given by

$$-\frac{\mu^2(2\lambda_2(\lambda_2 + 1)\varphi^2 - \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1)(\varphi^2 \tilde{\sigma}_1^2 + 1)}{4\varphi^2}$$

and the difference in period 2 was calculated in the proof of Proposition 3:

$$\frac{1}{2}\varphi^2 \sigma_1^2 \lambda_2 (\lambda_2 + 2)$$

Therefore, granting consent is optimal if and only if the following expression is non-negative,

$$-2\lambda_2(\lambda_2 + 1)\varphi^2 + \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + \varphi^2 \tilde{\sigma}_1^2 \left(2\lambda_2\varphi^2 + \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} - 1\right) - 1 \quad (43)$$

which is a linear expression in $\tilde{\sigma}_1^2$. It can be easily verified that the constant is negative, as it is equal to zero when evaluated at $\varphi^2 = 0$ and decreasing in φ^2 , while the coefficient of $\tilde{\sigma}_1^2$ has the same sign as λ_2 , because it is equal to zero when evaluated at $\varphi^2 = 0$ and its derivative w.r.t. φ^2 has the same sign as λ_2 . We can therefore conclude that granting consent is never optimal for $\lambda_2 < 0$, while for $\lambda_2 > 0$ it is optimal if and only if

$$\tilde{\sigma}_1^2 \geq \tau(\lambda_2) \triangleq \frac{2\lambda_2(\lambda_2 + 1)\varphi^2 - \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} + 1}{\varphi^2 \left(2\lambda_2\varphi^2 + \sqrt{4\lambda_2(\lambda_2 + 1)\varphi^2 + 1} - 1\right)} \quad (44)$$

The r.h.s. of (44) is increasing in λ_2 , as its derivative w.r.t. λ_2 is null at $\varphi^2 = 0$ and increasing in φ^2 . Therefore, $\tau(\lambda_2)$ can be inverted and granting consent is optimal iff

$$0 \leq \lambda_2 \leq t(\tilde{\sigma}_1^2), \quad (45)$$

where $t(\tilde{\sigma}_1^2)$ is the inverse of $\tau(\lambda_2)$. □

Supplementary Appendix

“Data Linkages and Privacy Regulation”

Rossella Argenziano and Alessandro Bonatti

C Nonlinear Equilibria

Weng, Wu and Yin (2023) study nonlinear separating Perfect Bayesian Equilibria in linear-quadratic games such as ours. For the special case $\varphi = 1$, $k = 0$, $c = 1$, $\sigma_2 = 0$, i.e. perfectly persistent types, they show that our game admits a unique nonlinear PBE. In particular, let the consumer’s type $\theta \in [\underline{\theta}, \bar{\theta}]$ and consider the consumer’s first-period PBE strategy $q_1(\theta, p_1, y_1)$.

When $\lambda_2 < 0$, Weng, Wu and Yin (2023, Section 5.2) show that this strategy takes the form

$$\frac{(q_1 - (\theta + b_1 y_1 - p_1) - \theta \lambda_2)^{\lambda_2}}{[q_1 - (\theta + b_1 y_1 - p_1) + (1 + \lambda_2)\theta]^{-1-\lambda_2}} = \frac{(-\bar{\theta} \lambda_2)^{\lambda_2}}{[(1 + \lambda_2)\bar{\theta}]^{-1-\lambda_2}}. \quad (46)$$

When $\lambda_2 > 0$, this strategy takes the following form:

$$\frac{(-q_1 + (\theta + b_1 y_1 - p_1) + \theta \lambda_2)^{\lambda_2}}{[q_1 - (\theta + b_1 y_1 - p_1) + (1 + \lambda_2)\theta]^{-1-\lambda_2}} = \frac{(\underline{\theta} \lambda_2)^{\lambda_2}}{[(1 + \lambda_2)\underline{\theta}]^{-1-\lambda_2}}. \quad (47)$$

Under these strategies, the highest (resp. lowest) consumer type chooses the myopic strategy $q_1 = \theta + b_1 y_1 - p_1$ when $\lambda_2 < 0$ (resp. > 0). Figure 10 compares the Weng, Wu and Yin (2023) with the myopic consumer’s best reply to prices and quality p_1, y_1 and with the linear equilibrium strategy in our paper.

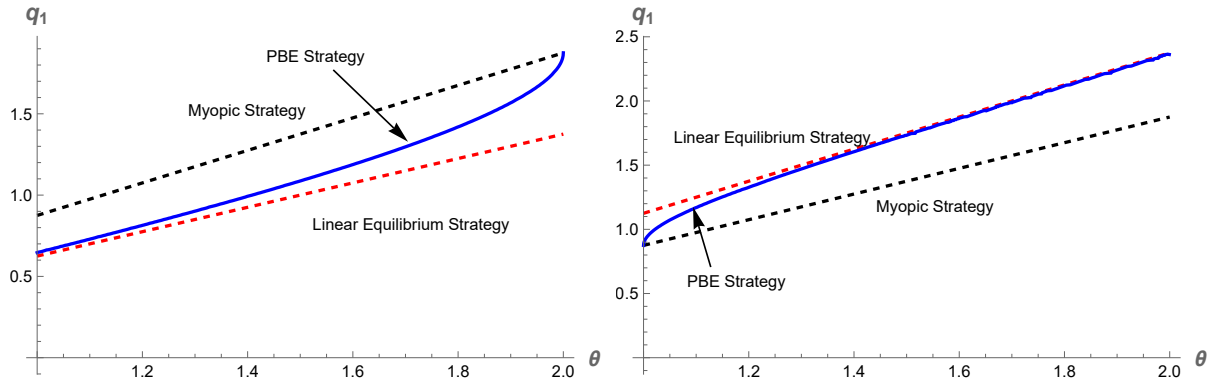


Figure 10: Linear and Perfect Bayesian Equilibria (fixed support)

The characterization in Weng, Wu and Yin (2023) implies (but their paper does not show) that, for all $\theta < \bar{\theta}$, the Perfect Bayesian Equilibrium strategy converges pointwise to

our Linear Bayesian equilibrium strategy as the support of the type distribution grows. In particular, let q^{LIN} and q^{PBE} denote the Linear and Perfect Bayesian equilibrium strategies, respectively. We then obtain the following new result.

Proposition 11 (Approximation). *If $\lambda_2 < 0$, we have $q^{PBE}(\theta) \rightarrow q^{LIN}(\theta)$ for every $\theta < \bar{\theta}$ as $\bar{\theta} \rightarrow \infty$. If $\lambda_2 > 0$, we have $q^{PBE}(\theta) \rightarrow q^{LIN}(\theta)$ for every θ as $\underline{\theta} \rightarrow 0$.*

Proof of Proposition 11. Consider the case $\lambda_2 < 0$ first. The closed-form solution for the PBE strategy is provided in equation (46). Fix θ and let $\bar{\theta} \rightarrow \infty$. Because $1 + 2\lambda_2 > 0$, the right-hand side of (46) diverges, but the denominator on the left-hand side remains bounded because the solution to the best-response problem of each type $\theta < \bar{\theta}$ is given by a finite q_1 . Therefore, the numerator must vanish (because the exponent is $\lambda_2 < 0$), which occurs at the linear equilibrium strategy

$$q_1 = (1 + \lambda_2)\theta + b_1 y_1 - p_1.$$

For the case of $\lambda_2 > 0$, Weng, Wu and Yin (2023) establish that the PBE and linear equilibrium coincide when $\underline{\theta} = 0$. The convergence result then follows from the continuity of (47) in q and $\underline{\theta}$ for every θ . $\square$

In Figure 11, we augment the picture in the Weng, Wu and Yin (2023) paper by comparing the PBE strategies for different supports of the consumer's type. (Note that the PBE first-period price and quality p_1 and y_1 depend on the distribution of types, and as such, they will vary with the support. In Figure 11, we hold them fixed to highlight the consumer's equilibrium strategy $q_1(\theta_1)$.)

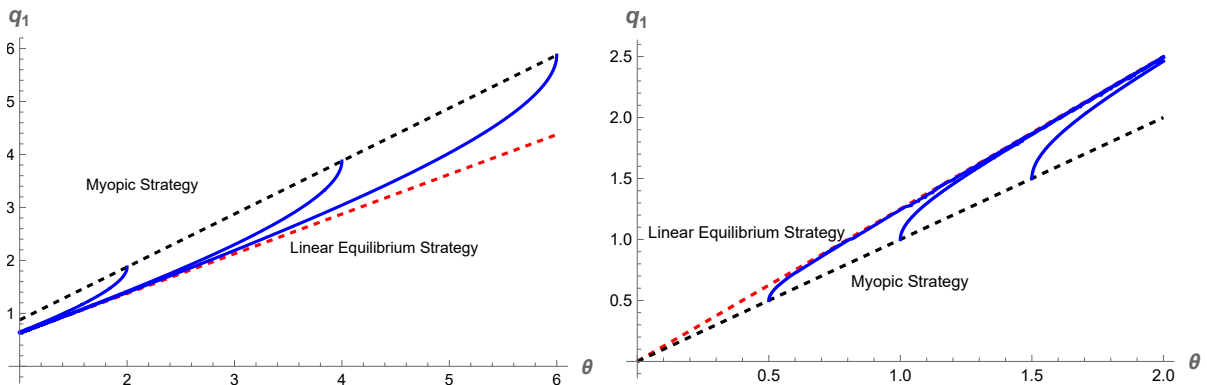


Figure 11: Linear and Perfect Bayesian Equilibria (varying support)

Weng, Xi, Fan Wu, and Xundong Yin. 2023. Linear Riley equilibria in quadratic signaling games. *Journal of Economic Theory*, 213.

D Two-Part Tariffs

We now relax the assumption that the second-period firm must adopt a linear pricing rule for each consumer. Specifically, we now assume that the second-period firm is able to offer a single two-part tariff $T + pq$ that conditions on the first-period information.

In this section, we derive conditions under which a linear equilibrium exists in both periods. In such an equilibrium, the second-period firm's quality provision, as well as the fixed and variable parts of the optimal tariff are linear in the first period outcome data. Furthermore, the consumer's quantity choice is a linear function of the current-period terms of trade. In particular, the consumer's strategy in the first period is given by

$$q_1 = \alpha\theta_1 + \beta y_1 + \gamma p_1 + \delta. \quad (48)$$

Throughout this section, we let $\varepsilon_2 \in [-\Delta, \Delta]$ and assume Δ is small enough that the second-period firm offers a two-part tariff that serves every type θ_2 . We then obtain the following result.

Proposition 12 (Linear Equilibrium). *If $\lambda_2 \in (-4\varphi^2/(1 + 4\varphi^2), 0)$, there exists a unique linear equilibrium in which the consumer plays strategy (48) with*

$$\alpha^* = \sqrt{\frac{(\lambda_2 + 1)\varphi^2}{\lambda_2} + \frac{1}{4} + \frac{1}{2}}, \quad \beta^* = b_1, \quad \gamma^* = -1, \quad \delta^* = \frac{(1 - \alpha^*)(2\Delta + \mu(\varphi - 1))}{\varphi}.$$

Proof. Consider the second period. In any separating (e.g., linear) equilibrium, firm 2 holds degenerate beliefs over the consumer's first-period type θ_1 . Under our full coverage assumption (i.e., Δ small), the fixed part of the optimal two-part tariff for firm 2 extracts the entire willingness to pay of the lowest type in the support of its beliefs. Because the consumer's second-period demand function is given by

$$q_2 = \theta_2 + b_2 y - p,$$

the fixed part of the optimal second-period tariff is given by

$$T^*(m) = \frac{(m - \Delta + b_2 y - p)^2}{2}.$$

Consequently, the second-period firm maximizes the following profit function

$$\Pi_2(p, y) := [(p - ky)(m + b_2 y - p) - c_2 y^2/2 + (m - \Delta + b_2 y - p)^2/2].$$

Note that the function Π_2 is globally concave if and only if

$$\lambda_2 := \frac{(b_2 - k)^2 - c_2}{2c_2 - (b_2 - k)^2} < 0.$$

When $\lambda_2 < 0$ the second-period firm's optimal (variable) price and quality satisfy

$$p_2^*(m) = \frac{\Delta (\lambda_2(b_2 + k) + k) - k(2\lambda_2 + 1)m}{\lambda_2(b_2 - k)}, \quad y_2^*(m) = \frac{(2\lambda_2 + 1)(\Delta - m)}{\lambda_2(b_2 - k)}.$$

These price and quality levels yield the following indirect utility function for the consumer:

$$U_2(\theta, m) = \frac{(\Delta - m + \theta_2)(2\Delta - 2m + (\Delta - 3m + \theta_2)\lambda_2)}{2\lambda_2}.$$

Under the linear conjecture for the first-period demand function by the consumer (26) and upon observing the outcome of the first period, firm 2 forms (degenerate) beliefs about θ_1 ,

$$m_2(q_1, p_1, y_1) = \varphi \frac{q_1 - (\beta y_1 + \gamma p_1 + \delta)}{\alpha} + (1 - \varphi)\mu.$$

Under this conjecture, the consumer then solves the following problem in the first period

$$q_1 = \arg \max_q [q_1(\theta_1 + b_1 y_1 - p_1) - q_1^2/2 + U_2(\varphi\theta_1 + (1 - \varphi)\mu, m(q_1, p_1, y_1))]. \quad (49)$$

Taking the first-order condition in problem (49) and matching coefficients with the conjectured strategy (48) yields the result. $\square$

In particular, Proposition 12 implies that $\alpha^* < 1$ for all $\lambda_2 < 0$ that admit a linear equilibrium, as in the baseline analysis (Proposition 2). The major difference with our baseline setting is that for $\lambda_2 > 0$ or φ large, the returns to quality are unbounded and hence no linear equilibrium exists in the static game.

However, whenever a linear equilibrium exists, it has qualitatively similar implications for markets for information as under linear pricing. Figure 12 shows the set of (λ_1, λ_2) pairs for which total producer surplus is higher with a data linkage than without.

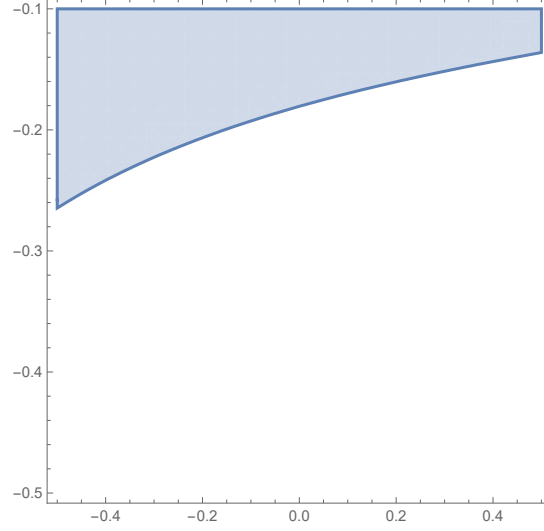


Figure 12: Linkages $(\lambda_1, \lambda_2) \in \Lambda^{PS}$ with two-part tariffs

Absent regulation, the pairs of firms in the shaded region will actively trade the consumer's information. The parameter values are $\mu = 1, \varphi = 1/6$, and the first and second period types ε_t are uniformly distributed on $[-4/9, 4/9]$. (The details of the calculations of producer surplus are available from the authors.)

E Noisy Data

We now extend our baseline model to the case where the second-period firm observes the first-period outcome with noise. For ease of exposition, we set the marginal cost of production $k = 0$, and we let types be fully persistent, i.e., $\varphi = 1$. We then introduce the following information structure.

The consumer's first-period type $\theta \in \mathbb{R}$ is drawn from a Gaussian distribution

$$\theta \sim N(\mu, 1/\tau_\theta).$$

Under a data linkage, the second-period firm observes the realized p_1, y_1 as well as a signal of the consumer's interaction with firm 1,

$$s \sim N(q_1, 1/\tau_q).$$

The analysis in the second period is unchanged relative to the baseline model—the consumer's second-period payoff can be written as a function of her type and the firm's posterior mean beliefs m ,

$$U_2(\theta, m) = \frac{1}{2}(\theta + \lambda_2 m)^2, \quad \text{where } \lambda_2 \triangleq \frac{b_2^2 - c_2}{2c_2 - b_2^2}.$$

In the first period, the consumer chooses q_1 to solve the following problem,

$$U_1(\theta, p_1, y_1) = \max_q \left[(\theta + b_1 y_1 - p_1)q - q^2/2 + \int U_2(\theta, m) dF(m | q) \right],$$

where $F(m | q)$ denotes the distribution of firm 2's posterior mean m given the consumer's quantity choice and the firm's conjecture about the consumer's strategy. Under the Gaussian information structure we have assumed, the consumer's choice of q leads to shifts in $F(m)$. It is then natural to look for a linear equilibrium.

Proposition 13 (Linear Equilibrium with Noisy Signal). *There exists a unique equilibrium in linear strategies. In this equilibrium, the consumer's first period demand is given by*

$$q_1^*(\theta, p_1, y_1) = \alpha^* \theta + b_1 y_1 - p_1 + \delta^*,$$

where α^* solves

$$\lambda_2 = (\alpha^* - 1) \left(\frac{\tau_\theta}{\alpha^{*2} \tau_q} + 1 \right), \tag{50}$$

and

$$\delta^* = \frac{\alpha^* \lambda_2^2 \mu \tau_q \tau_\theta}{(\alpha^{*2} \tau_q + \tau_\theta)^2}.$$

Proof of Proposition 13. To characterize a linear equilibrium, suppose firm 2 conjectures a linear strategy

$$\hat{q}_1(\theta, p_1, y_1) = \hat{\alpha}\theta + \hat{\beta}y_1 + \hat{\gamma}p_1 + \hat{\delta}. \quad (51)$$

The consumer's first-period best reply is then given by

$$q_1(\theta, p_1, y_1) = \theta + b_1 y_1 - p_1 + \lambda_2 \frac{\partial \mathbb{E}[m \mid q]}{\partial q} (\theta + \lambda_2 \mathbb{E}[m \mid q]).$$

Under a conjectured linear strategy (51) for the consumer, the second-period firm's posterior mean as a function of the realized signal s is given by

$$m := \mathbb{E}[\theta \mid s] = \frac{\mu \tau_\theta + \frac{s - \hat{\delta} - \hat{\gamma} p_1 - \hat{\beta} y_1}{\hat{\alpha}} \hat{\alpha}^2 \tau_q}{\tau_\theta + \hat{\alpha}^2 \tau_q}.$$

From the consumer's perspective, the expectation of the firm's belief m is a function of the first-period quantity choice, which is given by

$$\mathbb{E}[m \mid q] = \frac{\mu \tau_\theta + \frac{q - \hat{\delta} - \hat{\gamma} p_1 - \hat{\beta} y_1}{\hat{\alpha}} \hat{\alpha}^2 \tau_q}{\tau_\theta + \hat{\alpha}^2 \tau_q}.$$

Substituting in the linear best reply for the consumer and matching the coefficients yields the result. $\square$

The linear equilibrium of the model with noisy observations has the following properties, which are immediate from condition (50).

Corollary 2 (Equilibrium Properties).

1. α^* is continuous and strictly increasing in λ_2 , with $\alpha^* = 1$ if $\lambda_2 = 0$.
2. $\alpha^* > 0$ is decreasing in the signal precision τ_q if $\lambda_2 < 0$ and increasing if $\lambda_2 > 0$.
3. For all $\lambda_2 > -1/2$, $\alpha^* \rightarrow 1 + \lambda_2$ as $\tau_q \rightarrow \infty$.

Therefore, as expected, the linear equilibrium of a model with noisy observations converges to the linear equilibrium of the baseline model as the precision of the first-period quantity signal grows without bound.

F Direct Payments for Consent

We now describe the equilibrium outcome under a complete and efficient market for consumer information. In this section, we assume that transparency and consumer consent are required and that firm 1 is allowed to offer a direct (positive or negative) payment to the consumer in exchange for her consent to form a linkage with firm 2. For ease of exposition only, we let $\varphi = 1$ and $\varepsilon_2 = 0$, and we assume that firm 1 has all the bargaining power vis-à-vis firm 2 and the consumer. That is, firm 1 extracts all the surplus from the formation of a link. As the two firms bargain efficiently, firm 1 proposes forming the linkage (λ_1, λ_2) if and only if this linkage increases *social* surplus. Proposition 14 establishes our characterization result.

Proposition 14 (Social Welfare).

There exist two thresholds $\tilde{\lambda}_2(\lambda_1, \tilde{\sigma})$ and $\tilde{\tilde{\lambda}}_2(\lambda_1, \tilde{\sigma})$ satisfying $\tilde{\lambda}_2(\lambda_1, \tilde{\sigma}) < 0$ for all $\lambda_1, \tilde{\sigma} \geq 0$, and $\tilde{\tilde{\lambda}}_2(\lambda_1, \tilde{\sigma}) > 0$ for all $\lambda_1 < 0 < \tilde{\sigma}$, such that the following hold.

1. *For $\lambda_1 \geq 0$ all linkages with $\lambda_2 \geq \tilde{\lambda}_2(\lambda_1, \tilde{\sigma})$ increase social welfare.*
2. *For $\lambda_1 < 0$, all linkages $\lambda_2 \in [\tilde{\lambda}_2(\lambda_1, \tilde{\sigma}), \tilde{\tilde{\lambda}}_2(\lambda_1, \tilde{\sigma})]$ increase social welfare.*

Proof of Proposition 14. The total change in social welfare ΔW due to the formation of a linkage can be obtained by adding ΔU and $\Delta \Pi$:

$$\Delta W = \frac{\mu^2}{2} (\lambda_1 + 1) \lambda_2 (2\lambda_1 + \lambda_1 \lambda_2 + 2) + \frac{\tilde{\sigma}^2}{2} (3\lambda_2 + 1)$$

Dividing by μ^2 , multiplying by 2, and rearranging, we obtain

$$\Delta W \propto \lambda_2 (\lambda_2 + 2) \lambda_1^2 + \lambda_2 (\lambda_2 + 4) \lambda_1 + (3\lambda_2 + 1) \tilde{\sigma}^2 + 2\lambda_2. \quad (52)$$

This is a quadratic expression in λ_2 with a coefficient $\lambda_1 (1 + \lambda_1)$ on the quadratic term. The two roots are given by

$$\tilde{\lambda}_2(\lambda_1, \tilde{\sigma}) \triangleq \frac{-3\tilde{\sigma}^2 - 2(1 + \lambda_1)^2 + \sqrt{-4\tilde{\sigma}^2 \lambda_1 (1 + \lambda_1) + (3\tilde{\sigma}^2 + 2(1 + \lambda_1)^2)^2}}{2\lambda_1 (1 + \lambda_1)},$$

and

$$\tilde{\tilde{\lambda}}_2(\lambda_1, \tilde{\sigma}) \triangleq \frac{-3\tilde{\sigma}^2 - 2(1 + \lambda_1)^2 - \sqrt{-4\tilde{\sigma}^2 \lambda_1 (1 + \lambda_1) + (3\tilde{\sigma}^2 + 2(1 + \lambda_1)^2)^2}}{2\lambda_1 (1 + \lambda_1)}.$$

The term in the root is always positive. Furthermore, the following properties hold.

Whenever $\lambda_1 \geq 0$, we have $0 > \tilde{\lambda}_2(\lambda_1, \tilde{\sigma}) > -1/2 > \tilde{\tilde{\lambda}}_2(\lambda_1, \tilde{\sigma})$ for all $\tilde{\sigma} \geq 0$, and the expression (52) has a positive coefficient on the quadratic term. Therefore, all $\lambda_2 \geq \tilde{\lambda}_2(\lambda_1, \tilde{\sigma})$ increase social welfare.

Whenever $\lambda_1 < 0$, we have $-1/2 < \tilde{\lambda}_2(\lambda_1, \tilde{\sigma}) < 0 < \tilde{\tilde{\lambda}}_2(\lambda_1, \tilde{\sigma})$ and (52) has a negative coefficient on the quadratic term. Therefore, all $\lambda_2 \in [\tilde{\lambda}_2, \tilde{\tilde{\lambda}}_2]$ increase social welfare. $\square$

In Figure 13, we illustrate the set of welfare-improving linkages (λ_1, λ_2) .

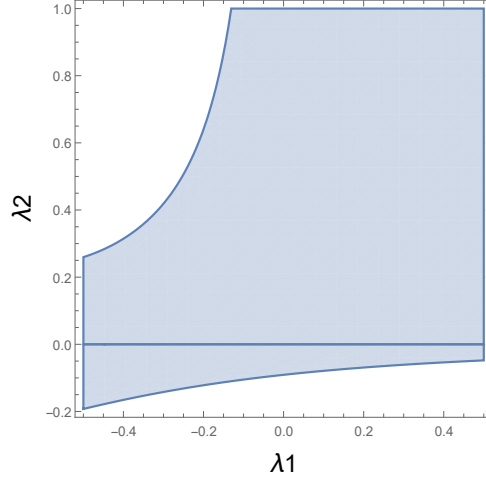


Figure 13: Socially Efficient Linkages ($\tilde{\sigma} = 1/2$)

In a static version of our model, the social value of information is positive for all λ larger than a threshold $\lambda^* < 0$. In a dynamic model with a data linkage, the consumer has an incentive to distort her demand, and the situation becomes more complex. Specifically, suppose the second-period firm has a large $\lambda_2 > 0$: if the first-period firm has $\lambda_1 < 0$, any linkage between these two firms causes a considerable loss in consumer surplus due to higher monopoly prices and upward quantity distortions in the first period. Likewise, for large $\lambda_1 > 0$, any linkage with $\lambda_2 < 0$ causes an inefficient reduction in consumer demand and underprovision of quality. The resulting loss is more severe for larger values of λ_1 , for which the consumer's average consumption is higher. Thus, relative to the λ_2 cutoff policy of Proposition 3, the social planner forms all linkages such that (heuristically) $\lambda_1 \cdot \lambda_2$ is sufficiently large, and would only form linkages with $\lambda_2 > 0$ as λ_1 becomes large.

In this scenario, the consequences for consumer welfare relative to the outcome of regulation depend heavily on the distribution of bargaining power. In our stylized setting, where firm 1 has all the bargaining power, the consumer is as well off as under privacy for any (λ_1, λ_2) . This outcome is weakly worse than that under required consent for any (λ_1, λ_2) , but the ranking relative to *laissez faire*, transparency, or no discrimination is sensitive to the specific values of λ_1, λ_2 , and $\tilde{\sigma}$.

G Consent by Informed Consumers

In this section, we revisit the most favorable privacy regulation in our baseline model, i.e., the case of Voluntary Consent (Section 5.3.2). We now analyze a game where consumers make their consent decisions after learning their first-period type. Our goal is to investigate the robustness of our conclusions regarding the active linkages in equilibrium under uninformed consent, which are given by $(\lambda_1, \lambda_2) \in \Lambda^{CS} \cap \Lambda^{PS}$ as in Proposition 9.

Our approach consists of characterizing pooling equilibria (when they exist) for any given pair of firms (λ_1, λ_2) . Because in these equilibria all types θ grant or deny consent, these equilibria are outcome-equivalent to uninformed decisions by consumers in our baseline model.

For expositional convenience, we let $\varphi = 1$ and $\varepsilon_2 = 0$, i.e., $\theta_1 = \theta_2$, and we denote the consumer's type by just θ . We further impose a specific assumption on the distribution of the consumer's type, namely that θ is uniformly distributed over the interval

$$\theta \sim U[\bar{\theta}/2, \bar{\theta}], \quad (53)$$

parameterized by $\bar{\theta} > 0$. (Under this assumption, we show below that the existence of pooling equilibria is independent of $\bar{\theta}$ and is thus much easier to illustrate.)

We begin our analysis with the relevant subgame, in which the firms have offered a linkage to the consumer. In this subgame, we show that there always exists a pooling equilibrium without consent. In contrast, a pooling equilibrium with consent exists only for linkages in the set Λ^* defined by the inequalities in (56) at the bottom of this section.

Proposition 15 (Pooling Equilibria).

1. *For any (λ_1, λ_2) , there exists an equilibrium where all types θ deny consent.*
2. *There exists an equilibrium where all types θ grant consent if and only if $(\lambda_1, \lambda_2) \in \Lambda^*$.*

The first part of Proposition 15 states that consumers can always be “trapped” into denying consent. On the equilibrium path, the two firms do not update their prior beliefs on the consumer's type. In contrast, the second part of Proposition 15 establishes that pooling equilibria where every consumer grants consent do not always exist. Indeed, if such an equilibrium exists, it can be supported by degenerate off-path beliefs. In particular, a consumer who denies consent will face the worst possible terms of trade in a static game in both periods. However, on the candidate equilibrium path, the consumer must play her dynamic best response to the firms' prices, which entails costly behavior distortions. When these equilibrium distortions are too high, the consumer is then willing to face worse terms

of trade in both periods in order to avoid them. This is the case if, for example $\lambda_1 < 0$ and λ_2 is sufficiently large.

Proof of Proposition 15. (1.) We build a pooling equilibrium with no consent. To do so, it is sufficient to specify the off-path beliefs (if a consumer grants consent) in such a way that both firms assign probability one to the worst type for that transaction (i.e., $\bar{\theta}/2$ if $\lambda_t > 0$ and $\bar{\theta}$ if $\lambda_t < 0$). Because the firms' beliefs are degenerate, the consumer cannot signal her type through her purchase level either. Thus, the terms of trade she faces in each period are the optimal ones (for the firms holding those beliefs) in a static game. Therefore, she would be better off denying consent and facing the equilibrium terms of trade for an anonymous consumer in a static game.

(2.) Consider the consumer's decision to grant consent. If she grants consent, she receives the equilibrium surplus level

$$U^*(\theta, \mu, \lambda_1) + U(\theta, \theta, \lambda_2). \quad (54)$$

If she denies consent, under our uniform distribution assumption (53), she receives the terms of trade for the worst type in each period. This type is given by

$$\hat{\theta}(\lambda) \triangleq \frac{\bar{\theta}}{2} + \mathbf{1}_{\{\lambda < 0\}} \frac{\bar{\theta}}{2}.$$

Consequently, the consumer's surplus off-path is given by

$$U(\theta, \hat{\theta}(\lambda_1), \lambda_1) + U(\theta, \hat{\theta}(\lambda_2), \lambda_2). \quad (55)$$

The resulting difference in surplus levels (54)-(55) can be written as a quadratic function of θ , with a coefficient of λ_2 on the term θ^2 . Evaluating the difference at the endpoints of the support of the type distribution (if $\lambda_2 < 0$) or at the unique critical point (if $\lambda_2 > 0$), and substituting the definition of the worst type $\hat{\theta}(\lambda)$, we obtain the set of linkages $(\lambda_1, \lambda_2) \in \Lambda^*$ for which

$$U^*(\theta, \mu, \lambda_1) + U(\theta, \theta, \lambda_2) \geq U(\theta, \hat{\theta}(\lambda_1), \lambda_1) + U(\theta, \hat{\theta}(\lambda_2), \lambda_2)$$

for all θ . This set is given by the union of the regions in (λ_1, λ_2) space described in (56) and illustrated in Figure 14. Recall that the set Λ^* is independent of $\bar{\theta}$ under our distribution assumption (53). $\square$

We now compare the set of linkages Λ^* with the set of linkages Λ^{CS} that benefit consumers ex ante. Although the two share similar qualitative properties, they are distinct. However, as $\tilde{\sigma} \rightarrow 0$, the set Λ^{CS} becomes a strict subset of Λ^* . This is shown in Figure 15.

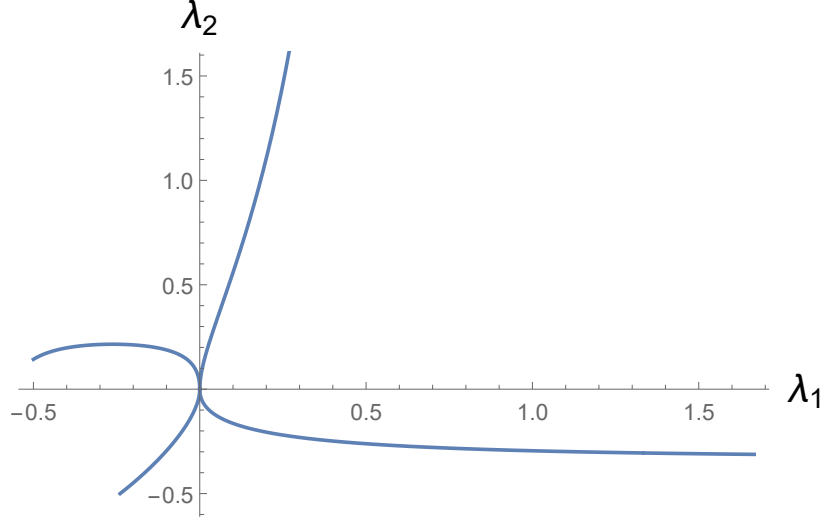


Figure 14: Linkages $(\lambda_1, \lambda_2) \in \Lambda^*$

A fortiori, as the degree of uncertainty over the consumer's type vanishes, the set of linkages that do form under ex ante consent $\Lambda^{CS} \cap \Lambda^{PS}$ is a strict subset of Λ^* . When this is the case, we can select the pooling equilibrium with consent for any $(\lambda_1, \lambda_2) \in \Lambda^{CS} \cap \Lambda^{PS}$ and the pooling equilibrium without consent for all other (λ_1, λ_2) .

Corollary 3. *As $\tilde{\sigma} \rightarrow 0$, the equilibrium outcome of the game with uninformed voluntary consent can be obtained as a pooling equilibrium of the game with informed consent for all (λ_1, λ_2) .*

Finally, note that there may exist other equilibria for some sets of linkages, such as threshold equilibria in which the consumer's types are partitioned into two intervals. In any one of these equilibria, high consumer types may grant or deny consent, depending on the firms' types. However, these equilibria exist for limited ranges of parameters, even in the uniform case.

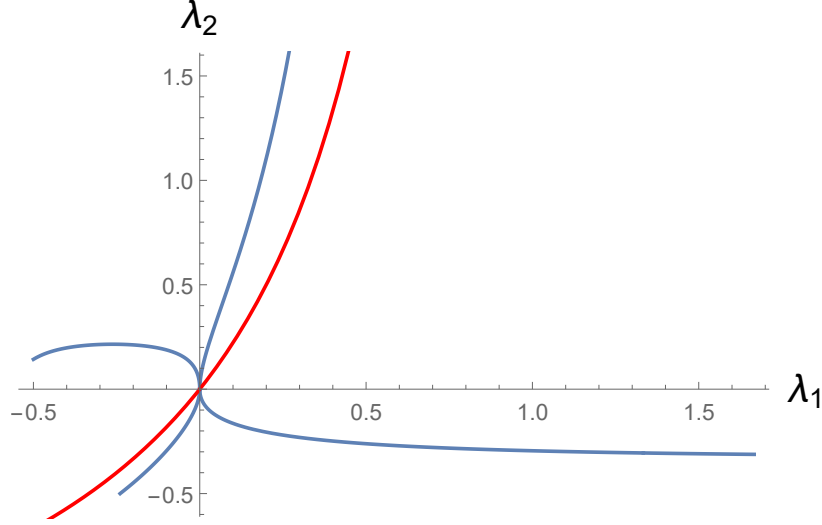


Figure 15: Comparison of Λ^* and Λ^{CS}

Note: the set Λ^* is given by the union of the following regions. (A Mathematica file with the calculations is available from the authors.)

$$\begin{aligned}
& \left\{ \lambda_1 > \frac{4}{3}, \lambda_2 \in \left[\frac{2\sqrt{\frac{8\lambda_1+8\lambda_1^2+3\lambda_1^3}{4+3\lambda_1}} - 3\lambda_1}{3\lambda_1 - 4}, 0 \right] \right\} \\
& \left\{ \lambda_1 \in [0, 4/3], \lambda_2 \in \left[-\frac{2\sqrt{\frac{8\lambda_1+8\lambda_1^2+3\lambda_1^3}{4+3\lambda_1}} - 3\lambda_1}{4 - 3\lambda_1}, 0 \right] \right\} \\
& \left\{ \lambda_1 \in \left[-\frac{1}{2}, \frac{4}{55} (3\sqrt{5} - 10) \right], \lambda_2 \in \left[-\frac{1}{2}, 0 \right] \right\} \\
& \left\{ \lambda_1 \in \left[\frac{4}{55} (3\sqrt{5} - 10), 0 \right], \lambda_2 \in \left[-\frac{4\sqrt{\frac{2\lambda_1+2\lambda_1^2+3\lambda_1^3}{4+3\lambda_1}} - 3\lambda_1}{4 - 3\lambda_1}, 0 \right] \right\} \\
& \left\{ \lambda_1 > \frac{2}{3}, \lambda_2 > 0 \right\} \\
& \left\{ \lambda_1 \in \left[0, \frac{2}{3} \right], \lambda_2 \in \left[0, \frac{3\lambda_1}{2 - 3\lambda_1} + 2\sqrt{\frac{3\lambda_1^3 + 4\lambda_1^2 + 2\lambda_1}{(3\lambda_1 - 2)^2 (2 + 3\lambda_1)}} \right] \right\} \\
& \left\{ \lambda_1 \in \left[-\frac{1}{2}, 0 \right], \lambda_2 \in \left[0, \frac{3\lambda_1}{2 - 3\lambda_1} + 2\sqrt{2}\sqrt{\frac{6\lambda_1^3 + 2\lambda_1^2 - \lambda_1}{(3\lambda_1 - 2)^2 (2 + 3\lambda_1)}} \right] \right\}. \tag{56}
\end{aligned}$$

References

- Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian, and Asu Ozdaglar.** 2022. “Too much data: Prices and inefficiencies in data markets.” *American Economic Journal: Microeconomics*, 14(4): 218–256.
- Acquisti, A., and H.R. Varian.** 2005. “Conditioning Prices on Purchase History.” *Marketing Science*, 24: 367–381.
- Acquisti, A., C. Taylor, and L. Wagman.** 2016. “The Economics of Privacy.” *Journal of Economic Literature*, 54: 442–492.
- Ali, S Nageeb, Greg Lewis, and Shoshana Vasserman.** 2023. “Voluntary Disclosure and Personalized Pricing.” *Review of Economic Studies*, 90(2).
- Argenziano, Rossella, and Alessandro Bonatti.** 2023. “Data markets with privacy-conscious consumers.” Vol. 113, 191–196, American Economic Association.
- Aridor, Guy, Yeon-Koo Che, and Tobias Salz.** 2023. “The effect of privacy regulation on the data industry: empirical evidence from GDPR.” *RAND Journal of Economics*, 54(4): 695–730.
- Bae, Donghwa, Raveesh Mayya, and Tingting Nian.** 2023. “Privacy Regulation and Its Unintended Consequence on Consumption Behaviors: Evidence From CCPA.” Vol. 18.
- Ball, Ian.** 2025. “Scoring strategic agents.” *American Economic Journal: Microeconomics*, 17(1): 97–129.
- Baye, Michael R, and David EM Sappington.** 2020. “Revealing Transactions Data to Third Parties: Implications of Privacy Regimes for Welfare in Online Markets.” *Journal of Economics & Management Strategy*, 29(2): 260–275.
- Belleflamme, Paul, and Wouter Vergote.** 2016. “Monopoly Price Discrimination and Privacy: The Hidden Cost of Hiding.” *Economics Letters*, 149: 141–144.
- Bergemann, Dirk, A. Bonatti, and T. Gan.** 2022. “The Economics of Social Data.” *RAND Journal of Economics*, 53(2): 263–296.
- Bergemann, Dirk, and A. Bonatti.** 2019. “Markets for Information: An Introduction.” *Annual Review of Economics*, 11: 85–107.

- Bonatti, Alessandro, and Gonzalo Cisternas.** 2020. "Consumer Scores and Price Discrimination." *Review of Economic Studies*, 87: 750–791.
- Bourreau, Marc, Cristina Caffarra, Zhijun Chen, Chongwoo Choe, Gregory S Crawford, Tomaso Duso, Christos Genakos, Paul Heidhues, Martin Peitz, Thomas Rønde, et al.** 2020. *Google/Fitbit will monetise health data and harm consumers*. Centre for Economic Policy Research.
- Burke, Jeremy, Julian Jamison, Dean Karlan, Kata Mihaly, and Jonathan Zinman.** 2023. "Credit building or credit crumbling? A credit builder loan's effects on consumer behavior and market efficiency in the United States." *Review of Financial Studies*, 36(4): 1585–1620.
- Calzolari, G., and A. Pavan.** 2006. "On the Optimality of Privacy in Sequential Contracting." *Journal of Economic Theory*, 130: 168–204.
- Canayaz, Mehmet, Ilja Kantorovitch, and Roxana Mihet.** 2022. "Consumer privacy and value of consumer data." *Swiss Finance Institute Research Paper*, 22-68.
- Chen, Zhijun.** 2022. "Privacy Costs and Consumer Data Acquisition: An Economic Analysis of Data Privacy Regulation." *Available at SSRN 4085923*.
- Choi, J., D. Jeon, and B. Kim.** 2019. "Privacy and Personal Data Collection with Information Externalities." *Journal of Public Economics*, 173: 113–124.
- Conitzer, Vincent, Curtis R Taylor, and Liad Wagman.** 2012. "Hide and seek: Costly consumer privacy in a market with repeat purchases." *Marketing Science*, 31(2): 277–292.
- Cummings, R., K. Ligett, M. Pai, and A. Roth.** 2016. "The Strange Case of Privacy in Equilibrium Models." 401–415. New York, NY:ACM Press.
- Dengler, Sebastian, and Jens Prüfer.** 2021. "Consumers' Privacy Choices in the Era of Big Data." *Games and Economic Behavior*, 130: 499–520.
- Dosis, Anastasios, and Wilfried Sand-Zantman.** 2023. "The ownership of data." *The Journal of Law, Economics, and Organization*, 39(3): 615–641.
- Doval, Laura, and Vasiliki Skreta.** 2024. "Purchase history and product personalization." *RAND Journal of Economics*.
- Fainmesser, Itay P, Andrea Galeotti, and Ruslan Momot.** 2023. "Digital privacy." *Management Science*, 69(6): 3157–3173.

- Frankel, Alex, and Navin Kartik.** 2022. “Improving information from manipulable data.” *Journal of the European Economic Association*, 20(1): 79–115.
- Fudenberg, Drew, and J Miguel Villas-Boas.** 2006. “Behavior-Based Price Discrimination and Customer Recognition.” *Handbook on Economics and Information Systems*.
- Fudenberg, Drew, and J Miguel Villas-Boas.** 2012. “Price Discrimination in the Digital Economy.” In *The Oxford Handbook of the Digital Economy*. Oxford University Press.
- Golrezaei, Negin, Adel Javanmard, and Vahab Mirrokni.** 2020. “Dynamic incentive-aware learning: Robust pricing in contextual auctions.” *Operations Research*, 69(1): 297–314.
- Ichihashi, Shota.** 2020. “Online Privacy and Information Disclosure by Consumers.” *American Economic Review*, 110(2): 569–595.
- Ichihashi, Shota.** 2021. “The Economics of Data Externalities.” *Journal of Economic Theory*, 196: 105316.
- Jann, Ole, and Christoph Schottmüller.** 2020. “An Informational Theory of Privacy.” *Economic Journal*, 130(625): 93–124.
- Johnson, Garrett A, Scott K Shriver, and Samuel G Goldberg.** 2023. “Privacy and market concentration: intended and unintended consequences of the GDPR.” *Management Science*, 69(10): 5695–5721.
- Jullien, B., Y. Lefouili, and M. Riordan.** 2020. “Privacy Protection, Security, and Consumer Retention.” Toulouse School of Economics.
- Ke, T Tony, and K Sudhir.** 2023. “Privacy rights and data security: GDPR and personal data markets.” *Management Science*, 69(8): 4389–4412.
- Kostka, Genia, and Lukas Antoine.** 2020. “Fostering model citizenship: Behavioral responses to China’s emerging social credit systems.” *Policy & Internet*, 12(3): 256–289.
- Krähmer, Daniel, and Roland Strausz.** 2023. “Optimal nonlinear pricing with data-sensitive consumers.” *American Economic Journal: Microeconomics*, 15(2): 80–108.
- Laffont, J.J., and J. Tirole.** 1988. “The Dynamics of Incentive Contracts.” *Econometrica*, 59: 1735–1754.
- Liang, Annie, and Erik Madsen.** 2024. “Data and incentives.” *Theoretical Economics*, 19(1): 407–448.

- Montes, Rodrigo, Wilfried Sand-Zantman, and Tommaso Valletti.** 2019. “The value of personal information in online markets with endogenous privacy.” *Management Science*, 65(3): 1342–1362.
- Paes Leme, Renato, Martin Pal, and Sergei Vassilvitskii.** 2016. “A Field Guide to Personalized Reserve Prices.” 1093–1102. Montreal, Canada:ACM.
- Peukert, Christian, Stefan Bechtold, Michail Batikas, and Tobias Kretschmer.** 2022. “Regulatory spillovers and data governance: Evidence from the GDPR.” *Marketing Science*, 41(4): 746–768.
- Shen, Qiaowei, and J Miguel Villas-Boas.** 2018. “Behavior-based advertising.” *Management Science*, 64(5): 2047–2064.
- Tang, Huan.** 2019. “The Value of Privacy: Evidence from Online Borrowers.” HEC Paris.
- Taylor, Curtis R.** 2004. “Consumer privacy and the market for customer information.” *RAND Journal of Economics*, 35(4): 631–650.
- Tirole, Jean.** 2021. “Digital Dystopia.” *American Economic Review*, 111(6): 2007–2048.
- Turow, Joseph.** 2008. *Niche envy: Marketing discrimination in the digital age*. MIT Press.
- Vestager, Margrethe.** 2020. “Statement to Committee on the Judiciary.” *United States House of Representatives*.
- Villas-Boas, J Miguel.** 1999. “Dynamic Competition with Customer Recognition.” *RAND Journal of Economics*, 604–631.
- Weng, Xi, Fan Wu, and Xundong Yin.** 2023. “Linear Riley equilibria in quadratic signaling games.” *Journal of Economic Theory*, 213: 105733.
- Zhang, Juanjuan.** 2011. “The Perils of Behavior-Based Personalization.” *Marketing Science*, 30(1): 170–186.