Data Linkages and Privacy Regulation*

Rossella Argenziano† Alessandro Bonatti‡

March 6, 2021

Abstract

We study data linkages among heterogeneous firms and examine how they shape the outcome of privacy regulation. A single consumer interacts sequentially with two firms: one firm collects data on consumer behavior; 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 then endogenize linkage formation under various forms of privacy regulation. We show that voluntary consent requirements are beneficial to consumers in equilibrium but that bans on discriminatory price and quality offers are harmful.

Keywords: consumer privacy; consumer consent; personal information; data linkages; data rights; price discrimination; transparency; ratchet effect.

JEL Classification: D44, D82, D83.

*Argenziano acknowledges financial support from through British Academy - Leverhulme 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, and seminar participants at Bergamo, Mannheim, MIT, THEMA, Toulouse, and Warwick for helpful discussions. We also thank 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

Motivation The widespread collection and distribution of individual data create *linkages* across seemingly unrelated activities. From browsing and search histories to geolocation data to social media activity, large online platforms gather vast amounts of information about their users. The data collected from one transaction enables targeted online advertising, tailored product offers and news stories, and even personalized prices across a myriad subsequent transactions.[1] Until recently, most data sharing occurred unbeknownst to consumers or without their explicit consent. In response to such market practices, a number of regulatory interventions such as the European Union's General Data Protection Regulation (GDPR) and the California Privacy Rights Act (CPRA) introduced transparency and consent requirements for data sharing. These policies aim to enable efficient data sharing by granting consumers property rights over their data.

However, this approach to *data governance* overlooks the fact that the data firms share with one another is rarely acquired from the consumer directly. Instead, most "big" datasets consist of information that firms "deduce from consumer behavior such as data about transactions on a platform (observed data) or about predictions on consumer behavior (inferred data)" (Vestager, 2020). In other words, 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 acquire these data to infer information about individual *preferences* and use it in future transactions.

A growing body of evidence supports the claim that the potential for future data sharing impacts consumers' behavior at the time of data collection.[2] With strategic consumers, the welfare implications of privacy regulation do not depend only on the allocation of property rights. Instead, to properly assess the impact of current policy and to clarify the need for further regulation, we must understand the impact of exercising those rights on the terms of trade in the transactions where the data are first collected.

[1] Marketing data from different transactions are linked through Customer Lifetime Value (CLV) scores—aggregate measures of profitability that merchants use to determine the level of service, prices, and perks to offer individual consumers (Bonatti and Cisternas, 2020). Another prominent example is the Chinese Social Credit system that determines access to credit, housing, and travel (Tirole, 2021).

[2] For example, Tang (2019) shows that disclosure requirements for loan applicants in an online peer-to-peer lending platform in China significantly affect application rates. Aridor et al. (2020) document the ability of some consumers to avail themselves of privacy-preserving technology (e.g., clearing cookies, anonymous browsing), even before the introduction of the GDPR. The incentive implications of linked transactions are even stronger in business-to-business interactions. For example, Paes Leme et al. (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 prices (Golrezaei et al., 2020).

Framework In this paper, we tease out the equilibrium effects of data linkages across heterogeneous transactions. We cast our analysis in a dynamic model of behavior-based price discrimination where a single consumer interacts sequentially with two firms. The active firm in each period sets both a quality level and a price (the *terms of trade*), and the consumer chooses how much to consume. The key object of interest is a *data linkage* between the two firms. An active data linkage enables the second firm to observe the entire outcome of the first period interaction, and to use the information so gained 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.

The presence of an active data linkage impacts both value creation (through quality discrimination) and value appropriation (through price discrimination) by the second-period firm. To capture the relative salience of price and quality discrimination, we introduce a firm-level characteristic that represents the sensitivity of the consumer's willingness to pay to the quality of the firm's product. We show that a data linkage increases the expected consumer surplus in the second period when the firm acquiring the data produces a good whose quality has a sufficiently large weight in the consumer's utility. This condition also characterizes the data linkages that benefit a naive consumer.

The effects of data linkages on strategic consumers, however, vary dramatically with the characteristics of the two firms, because a strategic consumer distorts her first period consumption level away from the static optimum to manipulate the second firm's beliefs. 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 (though lower quality) product in the second period—the canonical *ratchet* effect of Laffont and Tirole (1988). Conversely, if data is 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 way: if the consumer distorts her demand downward (upward), the first period firm lowers (raises) both its price and its quality level.[3]

Our results provide conditions on the two firms' characteristics and on the distribution of the consumer's willingness to pay under which a data linkage is beneficial to consumers.

[3] These strategic forces are robust to several extensions, including multiple or uncertain uses of period 1 data, and competition among period 2 firms. As these extensions do not affect our qualitative insights, we develop them in Appendices B and C.

The overall effect of a data linkage on consumer surplus is given by the combination of three forces: second-period discrimination, first-period demand distortion, and impact on (first period) terms of trade. In particular, the terms of trade effect suggests that consumers benefit from data linkages when the recipient firm is sufficiently similar to the collecting firm. Conversely, the first period distortion in behavior is costly for the consumer regardless of its direction. Therefore, the cost of distortions is minimized by interactions with second-period firms whose quality has a moderate weight, to avoid large (upward or downward) incentives to deviate from first-period utility maximization.

Privacy Regulation Having identified which data linkages increase consumer surplus, we turn to the question of which linkages form in equilibrium. To do so, 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 property rights to the consumer.[4]

Because the formation of a linkage is initiated by the firms, assigning decision rights to the consumers is equivalent to endowing them with formal veto power over linkage formation. However, since the data are traded contextually to the first-period transaction, the real strength of the consumer's rights depends on the (in)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 real-world privacy regulation, we examine the impact of various policies that differ in this respect.

We begin our analysis with the benchmark case of no regulation and no transparency. The second-period firm always has a positive value of information. Therefore, 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. Importantly, the two firms may be jointly worse off as a consequence.

Relative to this unregulated case, mandatory transparency forces the firms to announce the formation of a linkage to the consumer. However, the consumer has no veto rights and no ability to trade anonymously—the first period firm unilaterally decides whether a linkage is active or not. Transparency regulation thus grants valuable commitment power

[4] Our main 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 D. Likewise, we abstract from any signaling role of the consumer's consent decision by assuming a linkage is formed (or blocked) before she learns her type. Appendix E provides support for this assumption by analyzing pooling equilibria in a game where consent decisions are made by informed consumers.

to the firm, which internalizes both the impact of a linkage on the consumer's behavior and the value of selling the data to the second firm. Consequently, a linkage forms in equilibrium if and only if it improves the producers' total surplus.[5]

We then analyze the effects of mandatory consumer consent in three different forms, which roughly correspond to recently approved legislation in Nevada, California, and Maine, respectively. In its weakest form, consumer consent can be "required" by the firm: the consumer can choose to remain anonymous, but the first period firm can to 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 form, because the consumer has some real authority. Anticipating the consumer would forego trade in exchange for anonymity, the first period firm sometimes finds it more profitable to commit to anonymity.

A further strengthening of the consumer's property rights consists of voluntary consent. 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. Finally, we consider even more stringent laws that 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 the data linkages that do form, the consumer's privacy concerns lead to behavior distortions and modified terms of trade. Our results imply that requiring transparency by the firms allows them to commit to sharing data only when a linkage increases producer surplus. Because the sets of firm-optimal and consumer-optimal linkages are different, transparency policies have an ambiguous effect on consumer welfare. We find that adding a voluntary consent requirements improves consumer welfare, but banning discrimination unambiguously reduces it, relative to the case of pure voluntary consent. Under the right to equal service and price, the firm does not propose beneficial data linkages, and consumers consent to harmful linkages they would have vetoed under the equilibrium discriminatory offers.

[5] All data linkages create value for the second period firm by enabling price and quality discrimination. Conversely, in period 1, only linkages to period 2 firms whose quality is sufficiently valuable to consumers increase profits, stimulating demand, while linkages to firms with low-value quality depress period 1 demand. Therefore, the two firms are collectively better off creating a linkage only when the period 2 firm has quality above a critical value.

Related Literature Our paper joins a growing body of work on the economics of privacy and markets for information surveyed, e.g., by Acquisti et al. (2016), and Bergemann and Bonatti (2019). Our model is most directly related to the behavior-based price discrimination literature (Fudenberg and Villas-Boas, 2006), with seminal contributions by Taylor (2004), Acquisti and Varian (2005), Calzolari and Pavan (2006), and most recently by 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.

In the above papers, consumers do not control the quality of the information available to the firms. Several recent contributions, including Cummings et al. (2016), Frankel and Kartik (2019), Ball (2020), Bonatti and Cisternas (2020), and Jann and Schottmüller (2020), study how the consumer's manipulation incentives reduce the amount of information transmitted *in equilibrium*, and suggest mechanisms to mitigate this loss. 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, as well as the amount of information conveyed by her purchase history.

A distinct set of contributions, such as Conitzer et al. (2012), Belleflamme and Vergote (2016), Montes et al. (2019), and Ali et al. (2019), focus on the amount of information available to the firms when consumers can *actively* protect their privacy by remaining (fully or partially) anonymous, but doing so may involve a direct cost or carry signaling value. Ichihashi (2020b) studies *ex ante* information disclosure by a consumer to a seller who controls both a price and a (horizontal) quality dimension. Our model is simpler in this respect—all information is revealed in equilibrium. At the same time, information disclosure revelation occurs through a specific transaction, which allows us to focus on behavior distortions and their effects on the terms of trade and on welfare.

Motivated by recent privacy protection laws, Fainmesser et al. (2020) examine the distinction between data collection and data protection from adversaries. Jullien et al. (2020) study the related phenomenon data *leakages.* These leakages are modeled as reduced-form negative consequences of information diffusion. Likewise, Dosis and Sand-Zantman (2020) study a dynamic model of price discrimination in which a monopolist faces a tradeoff 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 we examine how it shapes the outcome of regulation.

A growing literature–including Choi et al. (2019), Bergemann et al. (2020), Ichihashi

(2020a), and Acemoglu et al. (2021)–studies the data externalities associated with collecting information from multiple consumers with correlated preferences. Relative to our paper, this literature largely abstracts from distortions in behavior arising due to the collection of personal data. While this is a reasonable approximation in the presence of a large number of consumers with correlated types, the extent of data externalities remains an open empirical question. One notable exception is Liang and Madsen (2020), who analyze the incentive effects of correlated types in a model of career concerns with data linkages, distinguishing between correlation in attributes and circumstances.

2 Model

A single consumer lives for two periods and interacts with a different firm in each period. The firm active at time $t = 1, 2$ sets a unit price p_t and a quality level y_t. The consumer, in turn, chooses a quantity level q_t. The consumer's per-period utility is given by

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

The variable θ represents the consumer's baseline consumption level (i.e., her ideal purchase size before adjusting for price and quality). We henceforth refer to θ as the consumer's type. 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 quality $b_t y_t - p_t$ as the *terms of trade* that firm t offers to the consumer.

Each firm t has a constant marginal cost of producing quantity q_t that we normalize to zero and a fixed per-consumer cost of producing quality y_t. Firm t's profits are then

$$\Pi(p_t, y_t, q_t) = p_t q_t - \frac{c y_t^2}{2}. \tag{2}$$

We assume that the sensitivity of the consumer's utility to quality satisfies $b_t \in [0, c\sqrt{2})$ in each period $t = 1, 2$, and we further normalize $c = 1$, so that the firm-level parameters $(b_1, b_2) \in [0, \sqrt{2})^2$ define the payoff environment.

While b_1 and b_2 are commonly known at the onset of the game, the type $\theta \in \Theta \subset \mathbb{R}_+$ is privately observed by the consumer. We denote the mean and the variance of the consumer's type by $\mu \triangleq \mathbb{E}[\theta]$ and $\sigma^2 \triangleq \text{var}[\theta]$, respectively. Firm 1 sets (p_1, y_1) on the basis of the prior distribution only. Firm 2's information sets, however, depend on the presence of a *data linkage*.

We define a data linkage as the ability by firm 2 to observe the outcome of the first period (p_1, y_1, q_1) before interacting with the consumer in the second period. In the presence of a data linkage, the timing of the game is the following:

1. Firm 1 offers a price p_1 and a quality level y_1 to the consumer.

2. The consumer learns her type θ and selects a quantity q_1.

3. Firm 2 observes the first-period outcome (p_1, y_1, q_1) before setting p_2 and y_2.

4. The consumer selects a quantity q_2.

In the absence of a data linkage, firm 2 does not observe anything prior to setting its terms of trade (p_2, y_2). Figure 1 provides an illustration of our model.

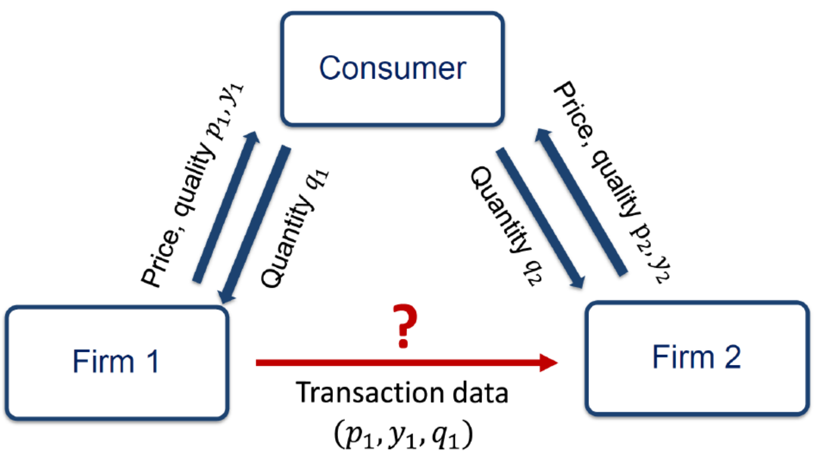

Figure 1: Model Sketch

We focus on *linear equilibria*—Bayesian Nash equilibria in which the consumer's strategy is linear in her type and the second-period firm's strategy is linear in any variable it observes.[6] In Section 3, we first analyze the linear equilibria of a static game with firm-level parameter b under different (exogenous) information structures. In Section 4,

[6]Linear equilibria are fully separating. However, because we define the consumer's type θ on a compact support, the consumer can choose actions that are off the equilibrium path. The linearity requirement disciplines firm 2's prices and qualities if this occurs, which has the effect of discouraging jumps in the consumer choice of q_1. See also the discussion in Ball (2020). Alternatively, we could have assumed that θ is distributed on $\mathbb{R}$ with full support, in which case all separating equilibria are linear.

we evaluate which data linkages benefit consumers and firms, respectively. Finally, 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 regulation.

3 Equilibrium Analysis

We begin our analysis with a static (one-period) benchmark, where we illustrate the welfare effects of information about the consumer's preferences when this is exogenously given to the firm. Therefore, Section 3.1 also serves as the naive-consumer benchmark, as well as the analysis of period $t = 2$ in the full game. Section 3.2 builds upon these results, proceeding by backward induction to uncover the value of information in a dynamic model where firm 2 can infer the consumer's type from the first period data.

3.1 Static Game

Consider a game between a consumer and a single firm that sells a product with quality of value b. The firm is endowed with an arbitrary information structure $\mathcal{I}$ consistent with the prior. The consumer observes the firm's offer (p, y) and simply maximizes her current-period utility (1). This yields consumer demand

$$q(\theta, p, y) = \theta + by - p. \tag{3}$$

The firm maximizes its expected profits (2) given the available information $\mathcal{I}$ and the demand function (3). This yields the following quality and price level

$$y^*(m, b) = \frac{bm}{2 - b^2}, \tag{4}$$
$$p^*(m, b) = \frac{m}{2 - b^2}, \tag{5}$$

where m denotes the firm's posterior mean

$$m \triangleq \mathbb{E}[\theta \mid \mathcal{I}].$$

The firm's choice of y is a fixed-cost investment in quality that shifts the consumer's demand function. Intuitively, the firm invests more when it anticipates the consumer will buy more units. Indeed, the optimal y^* in (4) maximizes total surplus given the monopoly price p^* in (5) and the consumer's demand curve.

The *terms of trade* (i.e., the price-adjusted quality level) summarize the role of the firm's beliefs for the consumer's problem:

$$\begin{aligned} by^*(m,b) - p^*(m,b) &= \lambda(b) \cdot m, \\ \lambda(b) &\triangleq \frac{b^2-1}{2-b^2}. \end{aligned} \tag{6}$$

Because the sensitivity parameter satisfies $b \in [0, \sqrt{2})$, the function λ has range $[-1/2, \infty)$. The value of λ captures the effect of the firm's beliefs on the (static) equilibrium terms of trade. We henceforth refer to λ_t as firm t's *type*. Intuitively, when the value of a firm's quality b is high, consumers with a higher type θ buy considerably more units at a higher quality level, which in turn justifies a large investment in y by the firm. Thus, firms with $\lambda > 0$ (i.e., $b > 1$) offer better terms of trade to higher-θ consumers.

Substituting (6) into the demand function (3), we obtain the realized consumer utility

$$U(\theta, m, b) = \frac{1}{2} q^*(\theta, y^*(m;b), p^*(m;b))^2 = \frac{1}{2}(\theta + \lambda(b) m)^2. \tag{7}$$

We may then ask how the availability of information affects the consumer and the firm *ex ante*. For this purpose, we consider the case of prior information only, where $m \equiv \mu$, and case of complete information, where $m = \theta$.

Proposition 1 (Value of Exogenous Information)

1. *Firm profits are higher under complete information for all* λ.

2. *Consumer surplus is higher under complete information for all* $\lambda > 0$.

3. *There exists a unique* $\lambda^* \in (-1/2, 0)$ *such that total surplus is higher under complete information if and only if* $\lambda > \lambda^*$.

The firm always benefits from information so as to tailor its price and quality to the consumer's type. Such discriminatory offers help the consumer in expectation if and only if $\lambda > 0$. Intuitively, information creates positive correlation between the firm's beliefs m and the consumer's type θ, and for $\lambda > 0$ this implies better terms of trade for the consumer when her true willingness to pay is in fact high. Finally, total surplus increases with information for $\lambda > 0$ as well as for some moderately negative λ.

All three effects in Proposition 1 are proportional to the prior variance σ^2, which measures the heterogeneity in the consumer's type. For illustration purposes, consider consumer surplus. Substituting $m = \mu$ (for prior information) and $m = \theta$ (for complete information) into (7), we obtain the ex ante value of information for the consumer,

$$\mathbb{E}_\theta \left[U(\theta, \theta, b) - U(\theta, \mu, b)\right] = \frac{1}{2}\sigma^2 (2 + \lambda(b)) \lambda(b). \tag{8}$$

This difference has the same sign of $\lambda(b)$ as shown in Proposition 1.

3.2 Dynamic Game

We now turn to the dynamic game played by the consumer with two firms of types λ_1 and λ_2 when the data linkage is active. Thus, firm 2 observes the terms of trade offered and the quantity purchased in the first period. Based on the previous analysis, we know the consumer benefits from the data linkage at t_2 if and only if $\lambda_2 > 0$. We now seek to identify the linkages that benefit a strategic consumer at t_1 and across both periods.

We begin our analysis of linear equilibria by illustrating the consumer's manipulation incentives. In any linear equilibrium, the first period quantity q_1 signals the consumer's type θ to firm 2. In particular, firm 2's (degenerate) posterior beliefs over θ are captured by an increasing, linear function $m(q_1)$. Since the consumer knows λ_2, she can compute her continuation payoff (7). She then solves the following problem

$$\max_{q_1} \left[U(\theta, q_1, p_1, y_1) + \frac{1}{2}(\theta + \lambda_2 m(q_1))^2\right]. \tag{9}$$

The consumer thus faces a trade-off between maximizing her t_1 utility and manipulating the t_2 terms of trade through a different choice of q_1. Critically, the direction of the manipulation incentives depends on the *sign* of λ_2, while the strength of such incentives depends on both the magnitude of λ_2 and on the sensitivity of firm 2's posterior $m(q_1)$.

Proposition 2 (Linear Equilibrium)
There exists a unique linear equilibrium of the game. In the linear equilibrium:

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

$$q_1^*(\theta, p_1, y_1) = \theta(1 + \lambda_2) + b_1 y_1 - p_1. \tag{10}$$

2. *Firm* 1 *offers terms of trade* (p_1^*, y_1^*) *that satisfy*

$$b_1 y_1^* - p_1^* = (1 + \lambda_2)\lambda_1 \mu. \tag{11}$$

3. *All players follow the second-period strategies (3)-(5), with* $m = \theta$.

Proposition 2 establishes the existence and uniqueness of linear equilibria in our dynamic game and characterizes the equilibrium strategies at t_1. Because this equilibrium is fully separating, the second-period game is played under complete information (i.e., with $m = \theta$) on the equilibrium path.

In the first period, the consumer distorts her equilibrium behavior away from the naive benchmark in (3). This distortion affects only the demand intercept through the weight placed on θ. In particular, the consumer places more weight on her type if $\lambda_2 > 0$ and less weight if $\lambda_2 < 0$.

To compute the magnitude of the equilibrium distortion in behavior, consider the second-period equilibrium utility of a type θ consumer as a function of the firm's beliefs, which is given in (7). Specifically, compute its derivative with respect to the firm's belief m, evaluated at the equilibrium (correct) beliefs $m = \theta$,

$$\frac{\partial U(\theta, \theta, \lambda_2)}{\partial m} = \lambda_2 (1 + \lambda_2)\, \theta. \tag{12}$$

From the first-order condition for the consumer's problem (9), we obtain the following equilibrium condition:

$$q_1^*(\theta, p_1, y_1) = \theta + b_1 y_1 - p_1 + \lambda_2 (1 + \lambda_2) \frac{\theta}{\alpha^*}, \tag{13}$$

where

$$\alpha^* \triangleq \frac{\partial q_1^*(\theta, p_1, y_1)}{\partial \theta} = \frac{1}{m'(q_1)}$$

is firm 2's equilibrium conjecture of the weight placed on θ by the consumer's strategy.

The first-order condition (13) rules out myopic behavior in equilibrium at t_1 if $\lambda_2 \neq 0$. Indeed, if firm 2 expected the consumer to maximize her t_1 utility (i.e., $\alpha^* = 1$), the right-hand side of (13) indicates that the consumer would instead place a weight of $1 + \lambda_2 (1 + \lambda_2)$ on θ. Recalling that $\lambda_2 \geq -1/2$, we immediately obtain that the consumer would buy more (less) quantity than optimal at t_1 depending on whether $\lambda_2 > 0$ (to raise firm 2's beliefs) or $\lambda_2 < 0$ (to depress them). Intuitively, a small deviation from static optimization has no first-order impact on t_1 utility but strictly improves the terms of trade at t_2.

Finally, matching the coefficients on (θ, p_1, y_1) in (13) yields the equilibrium demand function (10) and, in particular, $\alpha^* = 1 + \lambda_2$. From the perspective of firm 1 (because the mean type μ is positive), the consumer's behavior translates into an upward shift of the demand curve if $\lambda_2 > 0$ and a downward distortion if λ_2. This shift causes the equilibrium terms of trade (11) to shift by a factor of $1 + \lambda_2$, relative to the static benchmark in (6) with $\lambda = \lambda_1$. Combining the two parts of Proposition 2, it is immediate to verify that the quantity q_1^* traded in equilibrium also scales by the same factor $1 + \lambda_2$.

The driving forces of equilibrium behavior are robust to two important features of real-world data markets. In particular, in Appendix B, we allow for competition in the second period among firms with the same type but differentiated products. In Appendix C, we allow the first-period firm to form linkages with multiple, heterogeneous second-period firms. In both cases, the unique linear equilibrium shares the same qualitative properties as the one in Proposition 2.

4 Welfare Effects

In this section, we analyze the welfare implications of a data linkage between firms λ_1 and λ_2 in both periods of our game. The second-period welfare implications are described in Proposition 1. In the first period, a data linkage introduces behavior distortions and impacts the terms of trade, as described in 2. In particular, we have seen that the distortion in consumer behavior is a function of firm 2's type λ_2 only. However, the effect of this demand shift on the prevailing terms of trade at t_1 depends critically on both firms' types. We begin by illustrating two examples in Figure 2. In both examples, firm 1 is a pure price-setting firm (i.e., $b_1 = 0$; hence, $\lambda_1 = -1/2$).

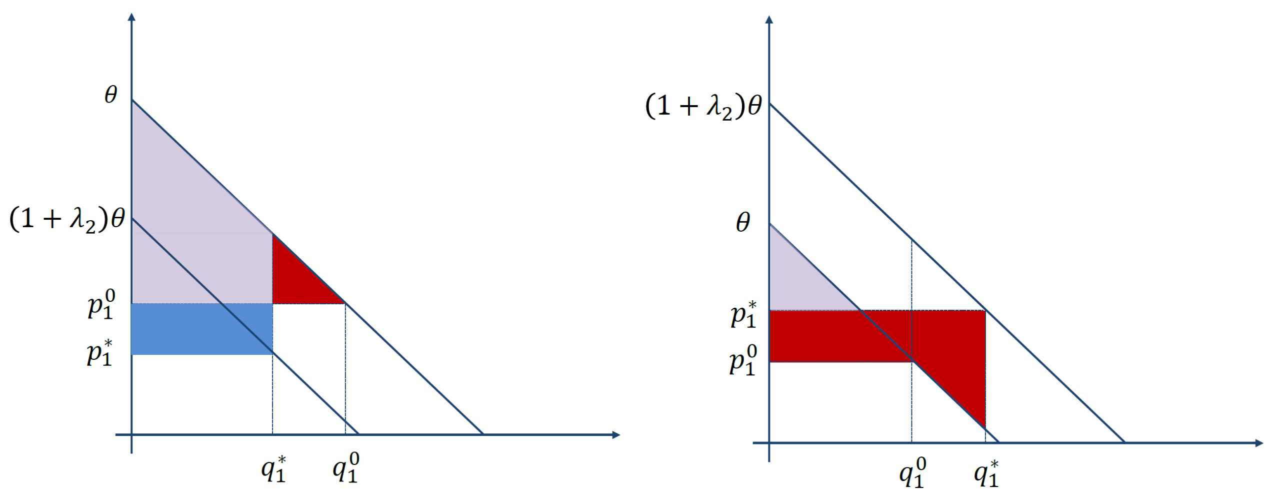

Figure 2: Demand and Price Shifts for $\lambda_1 = -1/2$, with $\lambda_2 < 0$ (left) and $\lambda_2 > 0$ (right)

The left panel captures the *ratchet* effect (Laffont and Tirole, 1988): because firm 2's type is $\lambda_2 < 0$, the consumer knows that a higher posterior belief by this firm leads to worse terms of trade. Firm 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 buys a smaller quantity q_1^* at a lower price p_1^* relative to the outcome (q_1^0, p_1^0) of a static game. Importantly, the consumer may benefit from this outcome, as the inframarginal discount $p_1^0 - p_1^*$ on q_1^* units (i.e., the blue rectangle) can compensate for the loss from foregoing consumption of the marginal units (i.e., the red triangle). In addition to these effects, the consumer faces a certain loss at t_2, when firm 2 learns θ perfectly.

The right panel captures the *niche envy* effect described by Turow (2008): firm 2's type is $\lambda_2 > 0$, which means the consumer wishes to manipulate the firm's beliefs upward to obtain better terms of trade. Thus, firm 1 expects a higher demand curve than that in a static game and charges a higher price, $p_1^* > p_1^0$. This price nonetheless leads the consumer to buy more units than statically optimal, $q_1^* > q_1^0$. Therefore, in the first period, the consumer buys "too many" units (the red trapezoid) at a higher price (the red rectangle). However, the consumer also enjoys better terms of trade at t_2.

More cases than those depicted in Figure 2 are possible. For example, if $\lambda_1 > 0$ and $\lambda_2 > 0$, then the consumer's upward demand shift in the first period would lead to more generous terms of trade at t_1. For any (λ_1, λ_2), however, the welfare effects of a data linkage at t_1 operate through the following two channels.

First, the terms of trade offered by firm 1 change to reflect the shifts in consumer demand. Comparing Propositions 1 and 2, the difference in terms of trade $b_1 y_1^* - p_1^*$ between the static and dynamic cases is related to the sign of λ_1 and λ_2, i.e.,

$$(b_1 y_1^* - p_1^*) - \left(b_1 y_1^0 - p_1^0\right) = \lambda_1 \cdot \lambda_2 \cdot \mu.$$

Thus, the consumer obtains better t_1 terms of trade if the two firms are similar in a very specific sense: the terms of trade improve when both firms produce quality that is of high *or* low value relative to money (i.e., λ_1 and λ_2 have the same sign). Specifically, if both firms are low quality, then the distortion causes a helpful reduction in price; and if they are both high quality, it causes a helpful increase in price-adjusted quality.

Second, the consumer's manipulation concerns introduce losses at t_1 due to the ensuing costly signaling that, despite being fully anticipated by firm 1, distorts the consumer's t_1 quantity away from the best reply to (p_1, y_1). Since demand is distorted up or down by an amount $\lambda_2\theta$, the magnitude of this loss is proportional to $(\lambda_2)^2$, as can also be seen from the red triangles in Figure 2. Thus, the consumer's cost of signaling is related to the strength of her manipulation incentives, regardless of their direction.

Proposition 3 summarizes the combination of these two effects—it compares the expected consumer and producer surplus at t_1 when a data linkage is active between firms λ_1

and λ_2 to the expected surplus levels in the static benchmark with firm type λ_1. Throughout, let $\hat{\sigma} \triangleq \sigma/\mu$ denote the coefficient of variation of the consumer's type distribution.

Proposition 3 (First-Period Welfare Effects)

1. *A data linkage increases consumer surplus at t_1 if and only if the following hold:*

$$\lambda_1 \cdot \lambda_2 > 0; \text{ and}$$
$$|\lambda_2| < |\lambda_1| \frac{2(1+\lambda_1)}{\hat{\sigma}^2 + 1 - \lambda_1^2} \text{ for all } \lambda_1 < \sqrt{1+\hat{\sigma}^2}.$$

2. *A data linkage increases firm 1's profits 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. Figure 3 illustrates the set of pairs (λ_1, λ_2) whose linkage is beneficial to consumers *in the first period.*

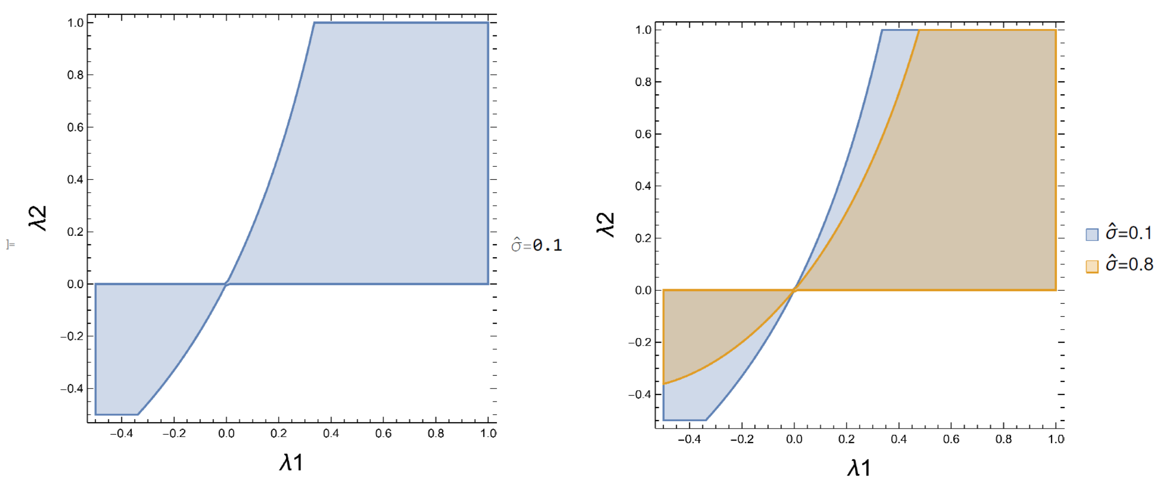

Figure 3: First-Period Consumer Surplus Improving Linkages ($\hat{\sigma} \in \{1/10, 8/10\}$)

Consistent with the trade-offs highlighted above, consumers can benefit in the first period only if two conditions are met: first, the firms' types must have the same sign; and second, λ_2 needs to be sufficiently small in magnitude such that the distortion in consumption does not trump the value of improved terms of trade. However, the latter condition applies only if λ_1 is smaller than the threshold—if the value of firm 1's quality is sufficiently large, then any $\lambda_2 > 0$ improves consumer surplus because the terms of trade effect dominates.

Finally, larger prior uncertainty $\hat{\sigma}$ does not affect the amount of distortion in behavior, but it unambiguously worsens its impact on expected consumer surplus. In particular, the region $\lambda_1 > \sqrt{1+\hat{\sigma}^2}$, where all $\lambda_2 > 0$ benefit the consumer, shrinks as $\hat{\sigma}$ increases. Because the distortion in behavior relative to the best reply at t_1 impacts the weight the consumer places on her type, the consumer suffers a convex loss equal to $(\lambda_2\theta)^2/2$. Thus, the variance of θ increases the expected loss to the consumer.[7]

Proposition 4 characterizes the welfare impact of a data linkage across both periods.

Proposition 4 (Intertemporal Welfare Effects)

1. *A data linkage increases consumer welfare if and only if the following hold:*

$$\begin{aligned} &\left(\hat{\sigma}^2 + \lambda_1(1+\lambda_1)\right)\cdot\lambda_2 > 0\text{; and} \\ &|\lambda_2| < \left|\hat{\sigma}^2 + \lambda_1(1+\lambda_1)\right|\frac{2}{1-\lambda_1^2} \text{ for all } \lambda_1 < 1. \end{aligned} \tag{14}$$

2. *A data linkage increases total firm profits if and only if*

$$\lambda_2 > \sqrt{\left(\frac{\hat{\sigma}^2}{2(\lambda_1+1)}\right)^2 + 1} - \frac{\hat{\sigma}^2}{2(\lambda_1+1)} - 1.$$

The total welfare impact of a data linkage combines the effect of exogenous information at t_2 with the first-period equilibrium forces. From the earlier Proposition 1 we know that the impact of a linkage on a naive consumer is independent of the point of collection of the data—naive consumers benefit from linkages to firms $\lambda_2 > 0$. From Proposition 4 we learn that for sophisticated consumers, instead, the effect of a linkage varies dramatically with the nature of the firm collecting the data.[8]

Figure 4 illustrates the sets of consumer- and firm-beneficial linkages (λ_1, λ_2) for different values of prior uncertainty $\hat{\sigma}$. We refer to these sets as Λ^{CS} and Λ^{PS}, respectively.

The overall effects of data linkages are thus best understood by comparing Propositions 3 and 4. In particular, two properties of Propositions 3 carry over to the intertemporal welfare effects: first, for any given λ_1, all λ_2 that benefit consumers have the same sign (left panels of Figure 4); second, all linkages with $\lambda_2 > 0$ benefit the firms. There are, however, important differences that we discuss below, beginning with the consumer's perspective.

[7] The social cost of data linkages also increases with $\hat{\sigma}$ because the impact on profits is constant in $\hat{\sigma}$.

[8] This distinction bears some resemblance to the experimental analysis by Lin (2019), who separates intrinsic and instrumental preferences for privacy. Our model does not have a separate intrinsic preference parameter—all our effects are through terms of trade at different times—but the ex ante value of information in the continuation game plays a very similar role in the analysis.

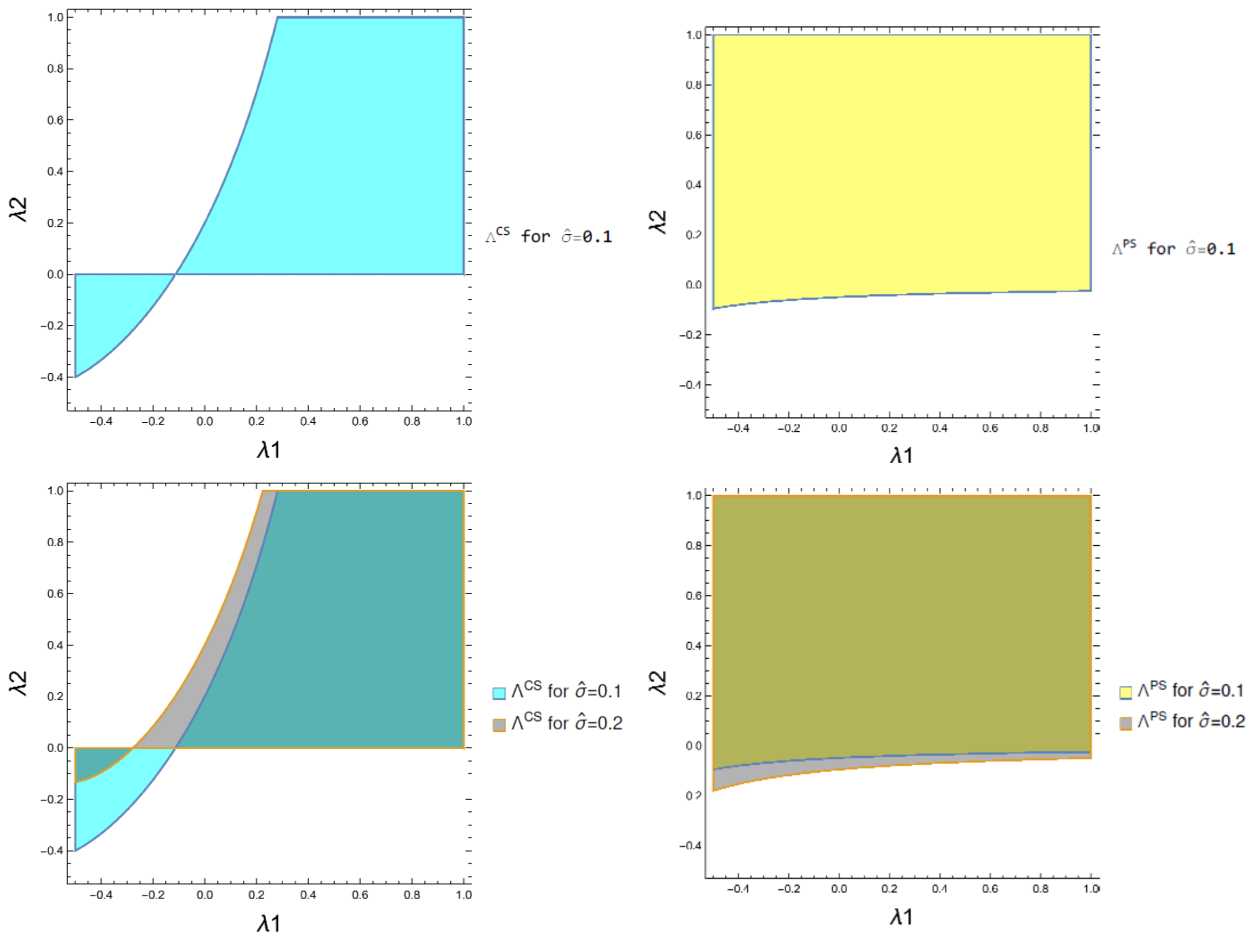

Figure 4: Consumer- and Producer-Optimal Linkages, ($\hat{\sigma} \in \{1/10, 1/5\}$)

Consider the case of $\lambda_2 < 0$. Any linkage to such a firm 2 reduces consumer surplus at t_2 and introduces costly distortions in behavior at t_1. Thus, a linkage with $\lambda_2 < 0$ can be beneficial only if $\lambda_1 < 0$ (so that t_1 terms of trade improve) and $|\lambda_2|$ is sufficiently small such that distortions are not excessively costly. Conversely, all linkages with $\lambda_2 > 0$ have a positive effect on consumer surplus at t_2. Therefore, if a (λ_1, λ_2) linkage leads to a sufficiently small worsening of terms of trade and a sufficiently small behavior distortion, then such a linkage can be beneficial to consumers even if $\lambda_1 < 0 < \lambda_2$. This is exactly what occurs in Figure 4 (but not in Figure 3): above a threshold $\lambda_1 < 0$, the left-hand side of (14) is positive, and hence all beneficial linkages have $\lambda_2 > 0$.

From the firms' perspective (right panels of Figure 4), the problem is easier: all firms benefit from linkages with $\lambda_2 > 0$. These linkages increase profits at t_2 due to price discrimination and raise demand at t_1 by means of the consumer's manipulation incentives. By continuity, for any λ_1, there exists a threshold $\lambda_2 < 0$ above which linkages also 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 are less willing to link to negative λ_2 firms (i.e., their threshold λ_2 is higher) because the resulting downward distortion in consumer behavior is more costly for them.

Finally, the effects of data linkages depend quantitatively on the distribution of consumer types. In particular, the welfare effects at t_2 (Proposition 1) are increasing in σ^2, while the terms-of-trade effect is proportional to μ^2. Therefore, as $\hat{\sigma}$ increases, the cost of distortions increases and the relative importance of the t_2 effects grows relative to the terms of trade effect. This shifts the consumer-beneficial set of linkages to the left in Figure 4. Likewise, as $\hat{\sigma}$ increases, the value of information for the firms at t_2 grows, as does the set of firm-beneficial linkages, i.e., more $\lambda_2 < 0$ linkages become profitable.

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 ones being the General Data Protection Regulation (GDPR) adopted by the EU in 2016 and the only three state-level privacy laws in the U.S.A.: the California Consumer Privacy Act (CPRA) of 2020, the Maine Act to Protect the Privacy of Online Consumer Information of 2019 and the Nevada Internet Privacy Act of 2019.

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 the firms from penalizing the consumer 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, we assume that the firms cannot commit to 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, we assume that the consumer makes any decisions relative 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" independently of her current-day inclination to shop. Third, we assume that the two firms bargain efficiently over the transfer of the consumer's purchase data. For ease of exposition, we assume that firm 1 holds all the bargaining power. However, our results do not rely on a specific bargaining protocol. This is the case, for example, when firm 1 is a large online platform. Finally, we adopt ex ante consumer surplus as our welfare criterion.

Our main results are the following. 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 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.

We examine three specific forms of consent rules: *required consent,* whereby a firm can refuse to trade with a consumer who denies her consent to data transfer; *voluntary consent,* whereby a firm cannot refuse to trade but can condition the terms of trade on the consent decision, within some limits; and the *right to equal service and price*, whereby the firm cannot condition the transaction nor the terms of trade on the consent decision. While at first glance one would imagine that the strongest limits to discrimination best protect the 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 5, we show that, for every pair (λ_1, λ_2), a linkage is formed, and the market equilibrium of Proposition 2 is played.

Proposition 5 (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 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 we have shown in Proposition 1.

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 4. 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 4. 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 an attractive strategic advice, it is also a pillar of all recent privacy legislation. For example, the GDPR 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, the CPRA establishes that *"A consumer shall have the right to request that a business that collects personal information about the consumer disclose (...) 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 at time 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 4 and Figure 4.

Proposition 6 (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 profit loss from the ratchet effect in the first period exceeds the gains from the sale of information. The effect on consumer surplus depends on the two firms' types. On the one hand, 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. $(\lambda_1, \lambda_2) \notin \Lambda^{PS} \cup \Lambda^{CS}$. On the other hand, 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: it forces them to "ask" the consumer for permission. We examine the role of consent requirements next.

5.3 Consent

A second principle that is currently well established in privacy protection legislation is that consumer data processing and sharing requires the explicit consent of the consumer. For example, the CPRA 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.[9] 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 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.

[9] Unlike 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 at t_2, here denying consent allows the consumer to retain her private information in the second-period game.

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

3. **No Discrimination**
   Firm 1 cannot refuse to trade with a consumer who denies consent, nor it can condition the terms of its offer on the consent decision.

In the first scenario, the consumer who exercises her right to privacy 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 facing 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 different legislators. For example, the GDPR establishes that *"Consent should be given by a clear affirmative act establishing a freely given, specific, informed and unambiguous indication ..."* and additionally specifies that *"Consent should not be regarded as freely given if the data subject has no genuine or free choice."* This seems to rule out the case of Required Consent, where the consumer who denies consent is deprived of the opportunity to trade with the firm.

The only three US states to have passed a privacy law take three very different stances on this issue. In Nevada, the consumer has a right to deny consent to the sale and transfer of her data, but has no protection from the consequences of denying such consent. In other words, all three scenarios above are legal in Nevada. In California, the CPRA establishes the right to equal service and price but it also allows for an important exception, by allowing firms 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."* In other words, in California Voluntary Consent is legal. Finally, in Maine, the privacy law fully embraces the right to equal service and price by establishing that *"A provider may not (1) Refuse to serve a customer who does not provide consent...; or (2) Charge a customer a penalty or offer a customer a discount* *based on the customer's decision to provide or not provide consent..."* Therefore, in Maine only the third scenario listed above is legal.

The introduction of consent requirements in the EU has already proven to impact consumers and firms decisions.[10] To contribute to the debate on the 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 behaviour, this is not the case.

5.3.1 Required Consent

In the case of Required Consent, the consumer correctly anticipates that if she denies consent, she will only trade at time 2, without her type being inferred, while 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 infer her type and make a personalized offer.

We denote the consumer's first period equilibrium payoff (which includes the distortion and terms of trade effects) by $U^*(\theta, \mu, \lambda_1)$ and her second period payoff by $U(\theta, m, \lambda_2)$. With this notation, the consumer's participation constraint is given by

$$\mathbb{E}_\theta \left[U^*(\theta, \mu, \lambda_1) + U(\theta, \theta, \lambda_2)\right] \geq \mathbb{E}_\theta \left[U(\theta, \mu, \lambda_2)\right].$$

Given the severity of the no-trade threat, the consumer consents to almost any linkage, except those where λ_1 is sufficiently small and λ_2 sufficiently large, in which case a linkage would force her to distort her demand at time 1 too much.

In turn, Firm 1 will only ask the consumer's consent if it expects to obtain it, else profits are nil, and if 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. Hence, firm 1 offers a linkage whenever $(\lambda_1, \lambda_2) \in \Lambda^{PS}$ and the consumer's participation constraint is satisfied. We characterize this set in Proposition 7 and illustrate it in Figure 5.

[10] Recent empirical work on the effects of the European Union's General Data Protection Regulation (e.g., Aridor et al. (2020) and Johnson et al. (2020)) shows a drop in traffic and website interconnectivity following the introduction of the regulation.

Proposition 7 (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,

$$\lambda_2 < \frac{\lambda_1}{1-\lambda_1} + \frac{\hat{\sigma}^2}{1-\lambda_1^2} + \sqrt{\frac{1+\hat{\sigma}^2}{(1-\lambda_1)^2} + \left(\frac{\hat{\sigma}^2}{1-\lambda_1^2}\right)^2}.$$

When $\lambda_1 \geq 1$ a data linkage is formed if and only if $(\lambda_1, \lambda_2) \in \Lambda^{PS}$. Whenever a linkage is not formed, the firm commits to privacy.

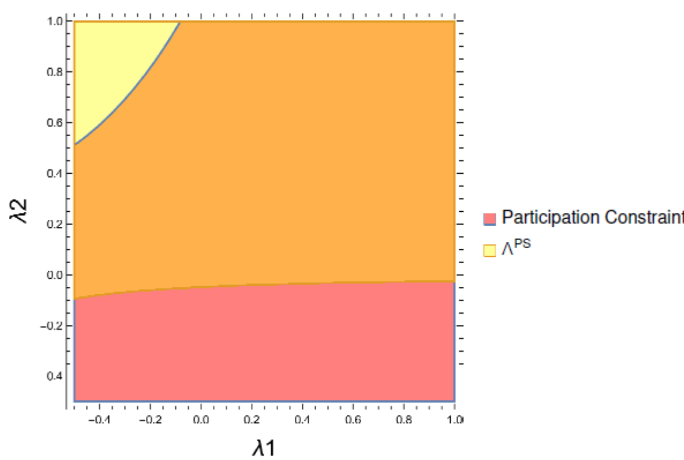

Figure 5: Required Consent

We now ask whether this type of regulation improves consumer welfare compared to a simple transparency rule. With transparency, a linkage is formed is and only if $(\lambda_1, \lambda_2) \in \Lambda^{PS}$, i.e. if and only if it improves producer surplus compared to privacy. With Required Consent, the consumer is given the option to veto a linkage to protect her anonymity in the transaction with firm 2, but she exercises it very rarely because the cost of the veto is the full surplus from a transaction with firm 1. We can therefore conclude that even this form of consent requirement most costly for the consumers does marginally improve their surplus by preventing some, thought very few, 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 5).

5.3.2 Voluntary Consent

We now consider the case of Voluntary Consent, where the law limits the negative consequences 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, the equilibrium prices and quality levels will be given by the linear equilibria of Propositions 1 and 2, depending on whether a linkage forms or not.

Under these circumstances, the consumer she will grant consent to data sharing if and only if this increases her expected surplus, i.e., if $(\lambda_1, \lambda_2) \in \Lambda^{CS}$:

$$\mathbb{E}_\theta \left[U^* \left(\theta, \mu, \lambda_1\right) + U \left(\theta, \theta, \lambda_2\right)\right] \geq \mathbb{E}_\theta \left[U \left(\theta, \mu, \lambda_1\right) + U \left(\theta, \mu, \lambda_2\right)\right],$$

In turn, firm 1 will only ask for the consumer's consent only if (1) it expects her to grant consent and (2) the linkage improves producer surplus i.e. $(\lambda_1, \lambda_2) \in \Lambda^{PS}$. Therefore, as a result of mutual veto rights, the linkages that will form are only those that constitute a Pareto improvement over anonymous trading.[11] We formalize this intuition in Proposition 8 and illustrate in Figure 6

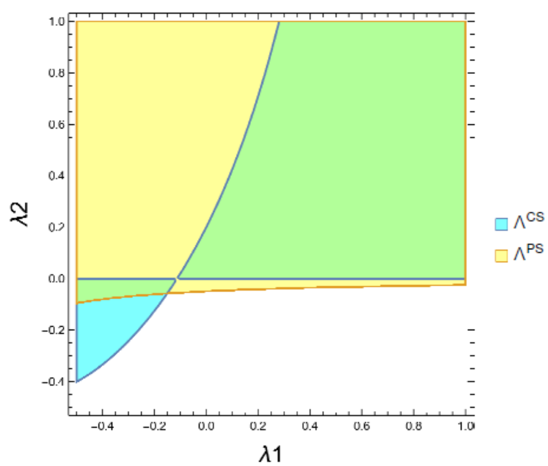

Figure 6: Voluntary Consent

Proposition 8 (Voluntary Consent)
If transparency and voluntary consumer 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 is able 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

[11] In Appendix E, we analyze the case of informed consent decisions, restricting attention to pooling equilibria. In particular, 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.

obtained under required consent, and all the linkages that are removed are detrimental to consumer welfare. Therefore, voluntary consent improves consumer surplus, compared to mandatory consent. Figure 7 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}.

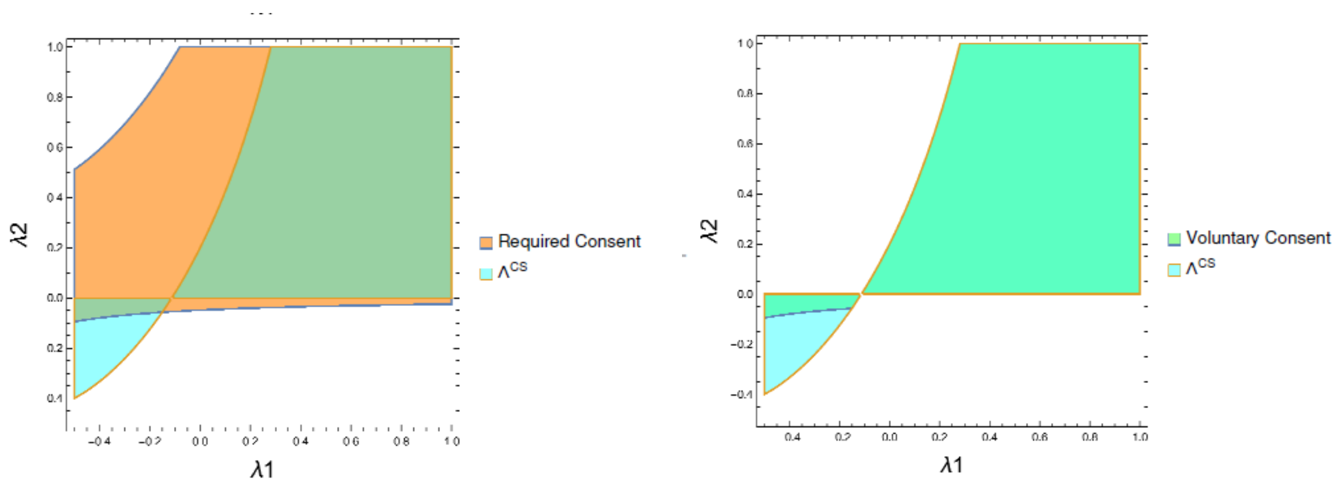

Figure 7: Required vs. Voluntary Consent

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 to condition the terms of trade on the consent choice.

In our model, we model a no-discrimination clause 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 the double 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.[12]

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

[12]It also has the realistic feature that most "consent boxes" appear on a webpage before the consumer can see the price for any product. With that said, the characterization of the equilibrium set of linkages in Proposition 9 below does not rely on this no-commitment assumption.

In turn, the consumer chooses whether to consent to the formation of a data linkage $\lambda_1 \to \lambda_2$ taking firm 1's choice of terms of trade (p_1, y_1) as given. We show in Proposition 9 that the consumer prefers to grant consent if and only if

$$\lambda_2(2\widehat{\sigma}^2 - \lambda_2) \geq 0, \tag{15}$$

for any first-period terms of trade. Furthermore, because condition (15) can hold only if $\lambda_2 \geq 0$, firm 1 will profitably propose all linkages that satisfy this condition because they all improve producer surplus. We can then characterize and illustrate the equilibrium set of data linkages as follows.

Proposition 9 (No Discrimination)
If transparency and consumer consent are required for data sharing and discrimination is forbidden, the linkage $\lambda_1 \to \lambda_2$ is formed if and only if $\lambda_2 \in [0, 2\widehat{\sigma}^2]$, for all λ_1.

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.[13]

Why, then, would the consumer want to consent to the transmission of her data? If $\lambda_2 < 0$, information revelation would worsen her terms of trade at time 2, thus reducing her time 2 surplus as well. Therefore, the consumer refuses consent for any $\lambda_2 < 0$, which explains the bottom half of Figure 8. If instead $\lambda_2 > 0$, the consumer obtains a higher surplus at time 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 information at time 2 is proportional to λ_2. Therefore, the consumer grants consent if λ_2 is positive but small. Finally, because the value of information is increasing in the prior uncertainty σ, the threshold λ_2 for granting consent is also increasing in $\hat{\sigma}$.

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 8 and 9 (see Figure 8) suggests the opposite.

[13]This 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.

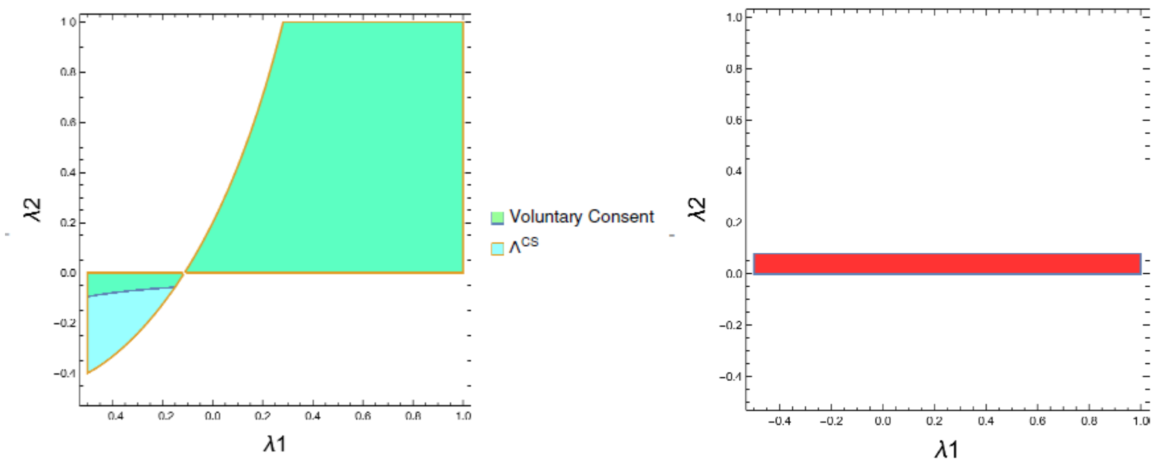

Figure 8: Voluntary Consent: with and without Discrimination

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 $\lambda_2 > 2\widehat{\sigma}^2$ and positive and sufficiently large λ_1. Finally, for $0 < \lambda_2 < 2\widehat{\sigma}^2$, firm 1 successfully proposes forming a linkage under a no-discrimination policy. 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, by doing so, she induced the equilibrium terms of trade under privacy, but that cannot happen under this policy.

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 (Consent and Discrimination)
If transparency and consumer consent are mandatory requirements for data sharing, banning discrimination over the terms of trade weakly damages consumers for any (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.

6 Conclusions

We have developed a simple model that microfounds a consumer's preferences over the collection and transmission of behavior data by heterogeneous firms. We have shown that 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. Our welfare results inform the evaluation of current privacy regulation both in the EU and in the US. In particular, we have shown how carefully designed mandatory and consent requirements for the formation of a data linkage can benefit consumers. In particular, voluntary consent provides firms sufficient flexibility to *reward* consumers for granting consent, while curbing their power to *extort* the consumer's consent by threatening to otherwise refuse service.

Even under mandatory consent and transparency, the instruments available to firms to compensate consumers for any losses in privacy remain imperfect. In particular, assigning consumers veto rights over harmful linkages reduces the number of such linkages that form in equilibrium. Because firms still hold proposal power over linkage formation, however, no welfare beneficial linkages can form if producers do not jointly benefit from them. Thus, mutual veto rights may lead to a socially suboptimal level of data sharing. The recent policy debate over "data portability" is an important first step in this sense—suggesting a framework to give consumers the means to create beneficial linkages, not just to veto harmful ones.

Appendix

A Proofs of Propositions

Proof of Proposition 1. The realized utility, profits, and welfare are, respectively:

$$\begin{aligned}
U(\theta, m, \lambda) &= \frac{(\theta + m\lambda)^2}{2} \\
\Pi(\theta, m, \lambda) &= \frac{m(2\theta - m)(1+\lambda)}{2} \\
W(\theta, m, \lambda) &= \frac{(\theta + m\lambda)^2}{2} + \frac{m(2\theta - m)(1+\lambda)}{2}
\end{aligned}$$

Taking expectations over (θ, m), we obtain the consumer's ex ante welfare, the firm's profits and total welfare. Denote the complete information structure $\mathcal{I}^*$ and the prior information structure by $\varnothing$. Under complete information, the expected utility, profits and welfare are given by

$$\begin{aligned}
\mathbb{E}[U \mid \mathcal{I}^*] &= \frac{1}{2}\left(\mu^2 + \sigma^2\right)(1+\lambda)^2, \\
\mathbb{E}[\Pi \mid \mathcal{I}^*] &= \frac{1}{2}\left(\mu^2 + \sigma^2\right)(1+\lambda), \\
\mathbb{E}[W \mid \mathcal{I}^*] &= \frac{1}{2}\left(\mu^2 + \sigma^2\right)(1+\lambda)(2+\lambda).
\end{aligned}$$

If instead the firm has only access to the prior information, we have $m = \mu$, and the expected utility and profits are given by:

$$\begin{aligned}
\mathbb{E}[U \mid \varnothing] &= \mathbb{E}_\theta\left[\frac{(\theta + \mu\lambda)^2}{2}\right] = \frac{\mu^2(1+\lambda)^2 + \sigma^2}{2} \\
\mathbb{E}[\Pi \mid \varnothing] &= \mathbb{E}_\theta\left[\frac{\mu(2\theta - \mu)(1+\lambda)}{2}\right] = \frac{1}{2}\mu^2(1+\lambda)
\end{aligned}$$

(1.) The change in consumer surplus is

$$\Delta U \triangleq \mathbb{E}[U \mid \mathcal{I}^*] - \mathbb{E}[U \mid \varnothing] = \frac{\sigma^2}{2}\lambda(\lambda + 2)$$

which is positive iff $\lambda > 0$.

(2.) When the firm has exogenous information and $m = \theta$, the change in profits is

$$\Delta\Pi \triangleq \mathbb{E}[\Pi \mid \mathcal{I}^*] - \mathbb{E}[\Pi \mid \varnothing] = \frac{\sigma^2(1+\lambda)}{2} > 0.$$

(3.) Finally, the change in social welfare is

$$\Delta W \triangleq \Delta U + \Delta\Pi = \frac{\sigma^2}{2}\left(\lambda^2 + 3\lambda + 1\right).$$

Given the domain of $\lambda \in [-1/2, \infty)$, a data linkage improves social welfare for

$$\lambda \geq \lambda^* = -\left(3 - \sqrt{5}\right)/2,$$

which is strictly negative. ■

Proof of Proposition 2. We seek to construct an equilibrium where the consumer's first-period strategy takes the form

$$q_1 = \alpha\theta + \beta y_1 + \gamma p_1 + \delta. \tag{16}$$

With this linear demand function, the firm maximizes its expected profits,

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

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

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

Therefore, if firm 1 conjectures the demand as in (16), its optimal choices of price and quality are given by

$$p_1^* = -\frac{\delta + \alpha\mu}{\beta^2 + 2\gamma} \tag{17}$$

$$y_1^* = -\beta\frac{\delta + \alpha\mu}{\beta^2 + 2\gamma}. \tag{18}$$

Next, we solve the consumer's problem and derive the equilibrium values of the coefficients of her linear demand.

The consumer maximizes (9), i.e., the sum of her current flow utility U_1 and her expected second period utility, which is given by (7). Under first-period demand (16), firm 2 forms a degenerate posterior belief over the consumer's type,

$$m(q_1) = \frac{q_1 - \beta y_1 - \gamma p_1 - \delta}{\alpha}. \tag{19}$$

The consumer anticipates (19) and therefore, upon observing the choice of (p_1, y_1), she solves the following problem:

$$\max_{q_1} \left[U_1(\theta, q_1) + U_2^*(\theta, m(q_1)) \right].$$

The first-order condition with respect to q_1 is given by

$$\theta + b_1 y_1 - p_1 - q_1 + \lambda_2 m'(q_1)(\theta + \lambda_2 m(q_1)) = 0.$$

Under (19) above, this is a linear equation in q_1. Solving this condition for q_1 yields the following linear function of (θ, y_1, p_1):

$$q_1^* = \frac{\theta + b_1 y_1 - p_1 + \frac{\lambda_2}{\alpha}\left(\theta + \lambda_2 \frac{-\beta y_1 - \gamma p_1 - \delta}{\alpha}\right)}{1 - (\lambda_2)^2/\alpha^2}.$$

Matching the coefficients to those in (16) we obtain a unique solution to the resulting system of linear equations, which pins down the equilibrium strategies:

$$\begin{aligned} \alpha^* &= 1 + \lambda_2 \\ \beta^* &= b_1 \\ \gamma &= -1 \\ \delta &= 0. \end{aligned}$$

Substituting into conditions (16)-(18) yields the equilibrium strategies in the statement.

■

Proof of Proposition 3. (1.) By Proposition 2, the consumer's first-period realized payoff with a data linkage can be written as

$$U(p_1^*, y_1^*, q_1^*) = \frac{1}{2}(\lambda_2 + 1)(\theta + \mu\lambda_1)\left[\theta(1 - \lambda_2) + \mu\lambda_1 + \mu\lambda_1\lambda_2\right].$$

Therefore, the expected first-period consumer surplus is

$$\mathbb{E}U_1 = \frac{\mu^2}{2}(\lambda_2+1)(1+\lambda_1)(1-\lambda_2+\lambda_1+\lambda_1\lambda_2) + \frac{\sigma^2}{2}\left(1-\lambda_2^2\right)$$

Without a data linkage, the consumer's expected utility is given by

$$\mathbb{E}U_1^p = \frac{\mu^2}{2}(1+\lambda_1)^2 + \frac{\sigma^2}{2}.$$

The value of the linkage is therefore equal to

$$\Delta U_1 \triangleq \mathbb{E}U_1 - \mathbb{E}U_1^p = \frac{\mu^2}{2}\left(\lambda_2(\lambda_1+1)(2\lambda_1-\lambda_2+\lambda_1\lambda_2) - \hat{\sigma}^2\lambda_2^2\right)$$

Solving the right hand side for λ_2 yields two roots:

$$\lambda_2 \in \left\{0, \lambda_1\frac{2(1+\lambda_1)}{\hat{\sigma}^2+1-\lambda_1^2}\right\}.$$

For all $\lambda_1 < \sqrt{1+\hat{\sigma}^2}$, the coefficient on the quadratic term in λ_2 is negative and the surplus-improving linkages are in between the two roots. For $\lambda_1 > \sqrt{1+\hat{\sigma}^2}$ the coefficient is positive and the second root negative, therefore all $\lambda_2 > 0$ improve consumer surplus.

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

$$\mathbb{E}\Pi_1 = \mathbb{E}_\theta\left[\Pi_1^*\right] = \frac{\mu^2}{2}(\lambda_2+1)^2(\lambda_1+1)$$

and

$$\mathbb{E}\Pi_1^p = \frac{\mu^2}{2}(1+\lambda_1),$$

respectively. The difference between these two then has the same sign as λ_2. ■

Proof of Proposition 4. (1.) Summing the changes in consumer surplus in the two periods (Propositions 3 and proof of Proposition 1), the overall change due to the introduction of a linkage is proportional to

$$\Delta U = (\lambda_1+1)\lambda_2(2\lambda_1-\lambda_2+\lambda_1\lambda_2) + 2\hat{\sigma}^2\lambda_2. \tag{20}$$

Solving for λ_2 we obtain the two roots

$$\lambda_2 \in \left\{0, \frac{2\left(\hat{\sigma}^2 + \lambda_1\left(1+\lambda_1\right)\right)}{1-\lambda_1^2}\right\}.$$

If $\lambda_1 > 1$, the second root is negative and the quadratic term on λ_2 is positive in (20). Therefore values of λ_2 for which $\Delta U > 0$ are all $\lambda_2 > 0$. Conversely, if $\lambda_1 < 1$, all values for which $\Delta U > 0$ are in between the two roots.

(2.) Summing the changes in profits in the two periods, the overall effect of the introduction of a linkage is proportional to

$$\Delta\Pi = \left(1+\lambda_1\right)\lambda_2\left(\lambda_2+2\right) + \hat{\sigma}^2\left(1+\lambda_2\right). \tag{21}$$

Because $\lambda_2 > -1$, the right-hand side of expression (21) is increasing in λ_2. Solving for λ_2 and selecting the larger root yields

$$\lambda_2 \geq \sqrt{\left(\frac{\hat{\sigma}^2}{2\left(\lambda_1+1\right)}\right)^2 + 1} - \frac{\hat{\sigma}^2}{2\left(\lambda_1+1\right)} - 1,$$

which is the condition for $\Delta\Pi \geq 0$ in the statement. ■

The proofs for Propositions 5, 6, and 8 in Section 5 are given in the text.

Proof of Proposition 7. When asked for her consent to a data linkage, the consumer's participation constraint is given by

$$\underbrace{\mathbb{E}_\theta\left[\left(\theta + b_1y_1 - p_1\right)\left[\left(1+\lambda_2\right)\theta + b_1y_1 - p_1\right] - \frac{\left[\left(1+\lambda_2\right)\theta + b_1y_1 - p_1\right]^2}{2}\right]}_{\text{net surplus from trade at } t_1} + \underbrace{\frac{\sigma^2}{2}\lambda_2\left(\lambda_2+2\right)}_{\text{value of information at } t_2} \geq 0$$

Simplifying it yields

$$\begin{aligned} & \left(\mu + b_1y_1 - p_1\right)\left[\left(1+\lambda_2\right)\mu + b_1y_1 - p_1\right] - \frac{\left[\left(1+\lambda_2\right)\mu + b_1y_1 - p_1\right]^2}{2} \\ + \quad & \sigma^2\left(1+\lambda_2\right)\left(1 - \frac{\left(1+\lambda_2\right)}{2}\right) + \frac{\sigma^2}{2}\lambda_2\left(\lambda_2+2\right) \geq 0, \end{aligned}$$

and hence

$$\left(\mu\left(1+\lambda_2\right) + b_1y_1 - p_1\right)\left(\mu\left(1-\lambda_2\right) + b_1y_1 - p_1\right) + \sigma^2\left(2\lambda_2+1\right) \geq 0. \tag{22}$$

Because firm 1 cannot commit to the terms of trade, the consumer anticipates that it will offer the equilibrium terms of trade with an active linkage,

$$\begin{aligned} p_1^* &= (1+\lambda_1)(1+\lambda_2)\mu, \\ y_1^* &= \sqrt{(\lambda_1+1)(2\lambda_1+1)}(1+\lambda_2)\mu. \end{aligned}$$

Substituting (p_1^*, y_1^*) into (22), we obtain

$$(\lambda_2+1)(\lambda_1+1)(\lambda_1(1+\lambda_2)+1-\lambda_2)+\widehat{\sigma}^2(2\lambda_2+1)\geq 0.$$

When $\lambda_1 < 1$, we select the unique root $\lambda_2 \geq -1/2$ and write the consumer's participation constraint as

$$\lambda_2 \leq \frac{1}{1-\lambda_1^2}\left(\lambda_1(1+\lambda_1)+\sigma^2+\sqrt{(1+\lambda_1)^2+\sigma^2\left(2\lambda_1+\sigma^2+\lambda_1^2+1\right)}\right). \tag{23}$$

When $\lambda_1 \geq 1$, the consumer's participation constraint is satisfied for all λ_2.

In turn, firm 1 only offers a linkage if it anticipates the consumer will accept it, and if the data linkage improves the producers' joint profits. Therefore, the set of active linkages is given by all $(\lambda_1, \lambda_2) \in \Lambda^{PS}$ that also satisfy (23). ■

Proof of Proposition 9. If the consumer gives consent, her expected utility is

$$\begin{aligned} \mathbb{E}_\theta U_1 &= \mathbb{E}_\theta\left\{(\theta+by_1-p_1)\left[(1+\lambda_2)\theta+by_1-p_1\right]-\frac{\left[(1+\lambda_2)\theta+by_1-p_1\right]^2}{2}\right\} \\ &= (\mu+by_1-p_1)\left[(1+\lambda_2)\mu+by_1-p_1\right]-\frac{\left[(1+\lambda_2)\mu+by_1-p_1\right]^2}{2}+\frac{\sigma^2\left(1-\lambda_2^2\right)}{2} \end{aligned}$$

in the first period and

$$\mathbb{E}_\theta U_2 = \frac{\left(\mu^2+\sigma^2\right)(1+\lambda_2)^2}{2}$$

in the second. If instead she denies consent, her expected utility is

$$\mathbb{E}_\theta U_1^p = \mathbb{E}_\theta\left[\frac{(\theta+by_1-p_1)^2}{2}\right] = \frac{(\mu+by_1-p_1)^2+\sigma^2}{2}$$

in the first period and

$$\mathbb{E}_\theta U_2^p = \frac{\mu^2(1+\lambda_2)^2+\sigma^2}{2}$$

in the second period. Giving consent is optimal iff

$$\mathbb{E}_\theta U_1 + \mathbb{E}_\theta U_2 - \mathbb{E}_\theta U_1^p - \mathbb{E}_\theta U_2^p \geq 0$$

The above condition can be simplified to

$$\frac{1}{2}\lambda_2\left(2\widehat{\sigma}^2 - \lambda_2\right) \geq 0,$$

which establishes the result. ■

Online Appendices

B Competing Firms

We now return to our baseline model, but we introduce competition in the second period. In particular, the consumer interacts with a monopolist firm of type λ_1 in the first period. She then faces two period-2 firms that sell differentiated products and compete in prices and qualities. The second-period firms share a common value of quality b_2. We let (p_{2j}, y_{2j}, q_{2j}) denote the second-period actions, with $j = 1, 2$. To maintain the assumption of linear demand, let the consumer's utility function in the second period be given by

$$U_2(p, y, q) \triangleq \frac{1}{2}\sum_{j=1}^{2}\left[(\theta + b_2 y_{2j} - p_{2j})q_{2j} - \frac{1}{2}q_{2j}^2\right] - s q_{21} q_{22}, \tag{24}$$

where $s \in [0, 1)$ captures with the degree of substitutability of the two products, i.e., the intensity of second-period competition.

We now characterize the unique linear equilibrium of the game in which the first-period firm has formed a linkage with both second-period competitors.

Proposition 10 (Equilibrium with Second-Period Competition) *For any $s \in [0, 1)$, there exists a unique linear equilibrium of the game.*

1. *The consumer's t_1 demand function is given by*

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

where

$$\alpha^*(s) \triangleq \frac{1}{2} + \frac{1}{2}\sqrt{4\hat{\lambda}(s) + 1}, \tag{25}$$

$$\hat{\lambda}(s) \triangleq \frac{b_2^2 + s^2 - 1}{(2 - b_2^2 + s(1-s))^2}. \tag{26}$$

2. *Firm 1 offers terms of trade (p_1^*, y_1^*) that satisfy*

$$b_1 y_1^* - p_1^* = \alpha^*(s)\,\lambda_1 \mu.$$

For moderately fierce competition in the second stage, the consumer's behavior is qualitatively identical to the case of a monopoly. Distortions again affect only the coefficient on θ, and the terms of trade effect is unchanged from the baseline case as a function of the coefficient α^*. Competition does, however, have a quantitative effect on the consumer's equilibrium behavior. As we can see from expressions (25) and (26), the equilibrium coefficient α^* is increasing in s for all b_2. Furthermore, α^* is larger than one for all $b_2 \geq \sqrt{1-s^2}$, which is strictly lower than the threshold $b_2 = 1$ in the case of monopoly. Intuitively, fiercer competition in the second period alleviates the ratchet effect. Conversely, competition creates a greater incentive for the consumer to be perceived as high type to receive higher-quality products at lower prices than under monopoly.

Proof of Proposition 10. We begin with the second-period behavior and equilibrium values. Taking the first order conditions in (24) with respect to (q_{21}, q_{22}) and solving for q_{21} and q_{22}, we obtain

$$q_{2i} = \frac{1}{1-s^2}\left((1-s)\theta + b_2\left(y_{2i} - sy_{2j}\right) - p_{2i} + sp_{2j}\right), \text{ for } i,j = 1,2 \text{ and } i \neq j.$$

Given a posterior belief m, each firm i at t_2 then maximizes

$$\Pi_{2i} = p_{2i}q_{2i} - \frac{1}{2}y_{2i}^2,$$

which yields the following first order conditions:

$$\begin{aligned} 0 &= \frac{1}{1-s^2}\left((1-s)m + b_2y_{2i} - sb_2y_{2j} - 2p_{2i} + sp_{2j}\right), \\ 0 &= \frac{1}{1-s^2} \cdot p_{2i}b_2 - y_{2i}. \end{aligned}$$

Solving for a symmetric equilibrium yields the following expressions:

$$\begin{aligned} y_2^* &= \frac{bm}{2 - b^2 + s\left(1-s\right)} \\ p_2^* &= \frac{(1-s^2)m}{2 - b^2 + s\left(1-s\right)}. \end{aligned}$$

Note that $b^2 \leq 2$ and $s^2 \leq s$ so the denominator is non-zero. The resulting terms of trade in period 2 are given by

$$b_2y_2^* - p_2^* = \frac{b^2 - (1-s^2)}{2 - b^2 + s\left(1-s\right)}m \triangleq \lambda_2\left(s\right)m. \tag{27}$$

We then compute the second-period utility of a consumer of type θ when interacting with a pair of firms with beliefs m, which is given by

$$U_2^* (\theta, m) = \frac{1}{2} \frac{(\theta + \lambda_2 (s) \cdot m)^2}{1+s}.$$

Now suppose there was a linear equilibrium where

$$q_1 = \alpha\theta + \beta y_1 + \gamma p_1 + \delta.$$

Given the first period linear demand, the second period firms form a degenerate belief on the consumer's type

$$m(q_1) = \frac{q_1 - \beta y_1 - \gamma p_1 - \delta}{\alpha}.$$

The consumer anticipates this and therefore solves the following problem

$$\max_{q_1} \left[U_1 (\theta, q_1) + U_2^* (\theta, m(q_1)) \right],$$

the first order condition for which is given by

$$\theta + b_1 y_1 - p_1 - q_1 + m'(q_1) \frac{\lambda_2 (s)}{1+s} (\theta + \lambda m(q_1)) = 0.$$

Solving for q_1 yields

$$q_1^* = \frac{\theta + b_1 y_1 - p_1 + \frac{\lambda_2(s)}{\alpha(1+s)} \left(\theta + \lambda_2 (s) \frac{-\beta y_1 - \gamma p_1 - \delta}{\alpha}\right)}{1 - \frac{\lambda_2(s)^2}{(1+s)\alpha^2}}$$

Matching the coefficients above we obtain a unique system of linear equations which pins down the strategies described in (25).

$$\begin{aligned} \alpha^* &= \frac{1}{2} + \frac{1}{2}\sqrt{1 + \frac{4\lambda_2 (s) (\lambda_2 (s) + 1)}{s+1}} \\ \beta^* &= b_1 \\ \gamma &= -1 \\ \delta &= 0 \end{aligned}$$

To complete the proof, we need to show that the term in the square root is always positive. Recall the definition of $\lambda_2(s)$ in (27), which implies

$$\frac{\lambda_2(s)(\lambda_2(s)+1)}{s+1} = \frac{b_2^2 + s^2 - 1}{(2 - b_2^2 + s(1-s))^2}$$

This expression is minimized at $b_2 = s = 0$, yielding a value of $-1/4$, so the square root is in fact always positive. ■

C Multiple Data Uses

Consider a consumer who interacts with a single firm at $t = 1$ and with a continuum of heterogeneous firms at $t = 2$. We refer to λ_t as the type of firm t. While the type of the period-1 firm λ_1 is commonly known, the type of the each second-period firm λ_2 is drawn from a distribution F with support $\Lambda \subseteq [-1/2, \infty)$. Thus, once collected, the consumer's data can be used in a large number of ways. An alternative, equivalent interpretation is that the consumer faces uncertainty over the type of the period-2 firm.

Recall that the expected surplus of consumer θ when interacting with second-period firm λ_2 is given by (7), i.e.,

$$U_2^* (\theta, m, \lambda_2) = \frac{1}{2} (\theta + \lambda_2 m)^2 .$$

Clearly, the firm's posterior belief m will vary depending on whether the firm has access to the period-1 outcome data.

We now characterize the equilibrium strategies and payoffs when the first-period outcome is observed by a measurable subset of period-2 firms $\Lambda^o \subseteq \Lambda$. Thus, all firms $\lambda_2 \in \Lambda^o$ observe (p_1, y_1, q_1) prior to setting their price and quality levels, while the remaining firms $\lambda_2 \in \Lambda \setminus \Lambda^o$ operate under the prior distribution only.

Upon receiving a first-period offer (p_1, y_1) and facing the prospect of firms $\lambda_2 \in \Lambda^o$ observing the first-period outcome, the consumer solves the following problem

$$\max_{q_1} \left[U_1 (\theta, q_1, p_1, y_1, \lambda_1) + \int_{\Lambda^o} U_2^* (\theta, m(q_1), \lambda_2) \, dF(\lambda) + \int_{\Lambda \setminus \Lambda^o} U_2^* (\theta, \mu, \lambda_2) \, dF(\lambda) \right] .$$

Proposition 11 characterizes the equilibrium strategies for an arbitrary "linked set" Λ^o.

Proposition 11 (Equilibrium with Multiple Uses)
For any linked set Λ^o, there exists a unique linear equilibrium of the game.

1. *In the first period, the consumer's demand function is given by*

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

where

$$\alpha^* (\Lambda^o) \triangleq \frac{1}{2} \left(1 + \sqrt{4k (\Lambda^o) + 1} \right), \text{ and} \tag{28}$$

$$k (\Lambda^o) \triangleq \int_{\Lambda^o} (1 + \lambda) \lambda dF(\lambda) . \tag{29}$$

2. Firm 1 offers terms of trade $(p_1^*(\Lambda^o), y_1^*(\Lambda^o))$ *that satisfy*

$$b_1 y_1^*(\Lambda^o) - p_1^*(\Lambda^o) = \alpha^*(\Lambda^o)\lambda_1\mu.$$

3. In the second period, all players follow the strategies in Proposition 1, with each firm λ *forming its beliefs according to its information set.*

As in the case of a deterministic λ_2, the consumer's manipulation incentives introduce a distortion in her first-period behavior that affects only the weight of the consumer's type in the equilibrium quantity.[14] Furthermore, the first-period terms of trade effect is entirely unchanged: firms $\lambda_1 > 0$ raise prices and quality levels when the set of firm-2 linked firms Λ^o leads the consumer to manipulate upward, i.e., to set $\alpha^* > 1$.

However, the consumer's incentives to manipulate her behavior are more responsive to their true type when the future interaction is uncertain. To formalize this comparison, we rewrite the function $k(\Lambda^o)$ in (29) as

$$k(\Lambda^o) = F(\Lambda^o)\left[\mathbb{E}[\lambda|\Lambda^o] + \mathbb{E}[\lambda|\Lambda^o]^2 + \text{var}[\lambda|\Lambda^o]\right].$$

The consumer responds more aggressively to her type when the nature of the second-period interaction is stochastic, relative to the deterministic case in which $\text{var}[\lambda|\Lambda^o] = 0$. This occurs because the incentives to manipulate are related to the consumer's type through the product of two terms: first, the marginal value of a higher θ on the continuation value vis-à-vis firm λ is given by $1 + \lambda$; second, the marginal value of manipulating the firm's belief is itself λ. Thus, the marginal benefit of manipulation is a convex function of λ.

Proof of Proposition 11. We now characterize a linear equilibrium in which the consumer plays the first period strategy

$$q_1 = \alpha\theta + \beta y_1 + \gamma p_1 + \delta. \tag{30}$$

In the second period, firms set prices as in (4) and (5), with $m = \mu$ for $\lambda \notin \Lambda^o$ and $m = m(q)$ as in (19). The consumer accordingly uses her myopic demand function and obtains $U_2^*(\theta, \lambda_2, m)$ as in (7).

[14] The case of a single, deterministic λ_2 corresponds to the case where the distribution $F(\lambda_2)$ is degenerate. In that case, the right-hand side of (28) reduces to $\alpha(\lambda_2) = 1 + \lambda_2$, which is the expression in Proposition 2.

Under this period-2 conjecture, the consumer's period 1 can be written as

$$W(\theta) = \max_q \left[\begin{array}{c} (\theta + b_1 y_1 - p_1) q - \frac{q^2}{2} + \frac{1}{2} \int_{\Lambda^o} \left(\theta + \lambda \frac{q - (\beta y_1 + \gamma p_1 + \delta)}{\alpha} \right)^2 \mathrm{d}F(\lambda) \\ + \frac{1}{2} \int_{\Lambda \setminus \Lambda^o} (\theta + \lambda \mu)^2 \mathrm{d}F(\lambda) . \end{array} \right] \tag{31}$$

If the strategy (30) is an equilibrium, then it satisfies the first-order condition for the consumer's problem (31)

$$\theta + b_1 y_1 - p_1 - q_1 + \int_{\Lambda^o} \frac{\lambda}{\alpha} \left(\theta + \lambda \frac{q_1 - (\beta y_1 + \gamma p_1 + \delta)}{\alpha} \right) \mathrm{d}F(\lambda) = 0$$

as well as (30). Substituting the latter into the f.o.c., we obtain

$$0 = \theta + b_1 y_1 - p_1 - (\alpha \theta + \beta y_1 + \gamma p_1 + \delta) + \frac{\theta}{\alpha} \int_{\Lambda^o} (1 + \lambda) \lambda \mathrm{d}F(\lambda) .$$

Matching coefficients, we obtain the unique solution

$$\beta = b_1, \ \gamma = -1, \ \delta = 0,$$

and

$$1 - \alpha + \frac{k(\Lambda^o)}{\alpha} = 0,$$

where $k(\Lambda^o)$ is defined as in (29). We solve for α and select the unique positive root for α, so that the resulting prices and qualities in (32)-(33)

$$p_1^*(\Lambda^o) = \frac{\alpha^*(\Lambda^o)}{2 - b_1^2} \mu \tag{32}$$

$$y_1^*(\Lambda^o) = b_1 \frac{\alpha^*(\Lambda^o)}{2 - b_1^2} \mu \tag{33}$$

are non-negative. This yields equation (28), and completes the proof. ■

D Direct Payments for Consent

We describe the equilibrium outcome under a complete and efficient market for consumer information. We assume that transparency and consumer consent are mandatory and that firm 1 is allowed to offer a direct (positive or negative) payment to the consumer in exchange for her consent to forming a linkage with firm 2. For ease of exposition, assume further that firm 1 has all the bargaining power vis-à-vis firm 2 and the consumer, i.e., that it extracts all the surplus from the formation of a link. Because bargaining is assumed efficient, firm 1 proposes forming the linkage $\lambda_1 \to \lambda_2$ if and only if this linkage increases *social* surplus. Proposition 12 establishes our characterization result.

Proposition 12 (Social Welfare)

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

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

Proof of Proposition 12. The total change in social welfare ΔW due to the formation of a linkage can be obtained by adding lines (20) and (21):

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

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

$$\Delta W \propto \lambda_2(\lambda_2 + 2)\lambda_1^2 + \lambda_2(\lambda_2 + 4)\lambda_1 + (3\lambda_2 + 1)\hat{\sigma}^2 + 2\lambda_2. \tag{34}$$

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

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

and

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

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

Whenever $\lambda_1 \geq 0$, we have $0 > \tilde{\lambda}_2(\lambda_1, \hat{\sigma}) > -1/2 > \tilde{\tilde{\lambda}}_2(\lambda_1, \hat{\sigma})$ for all $\hat{\sigma} \geq 0$, and the expression (34) has a positive coefficient on the quadratic term. Therefore, all $\lambda_2 \geq \tilde{\lambda}_2(\lambda_1, \hat{\sigma})$ increase social welfare.

Whenever $\lambda_1 < 0$, we have $-1/2 < \tilde{\lambda}_2(\lambda_1, \hat{\sigma}) < 0 < \tilde{\tilde{\lambda}}_2(\lambda_1, \hat{\sigma})$ and (34) has a negative coefficient on the quadratic term. Therefore all $\lambda_2 \in [\tilde{\lambda}_2, \tilde{\tilde{\lambda}}_2]$ increase social welfare. ■

In Figure 9, we illustrate the set of welfare-improving linkages (λ_1, λ_2).

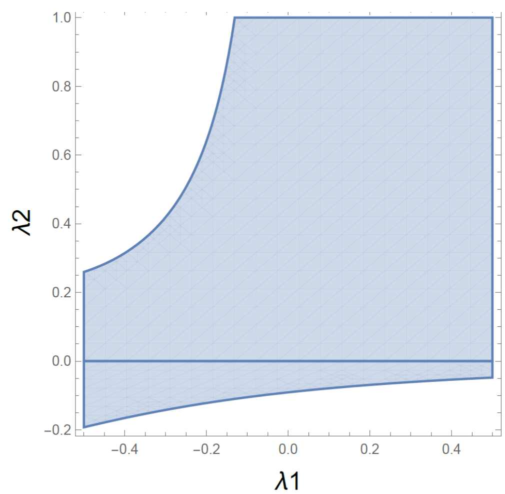

Figure 9: Socially Efficient Linkages $(\hat{\sigma} = 1/2)$

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 underinvestment in product 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 1, the social planner forms all linkages such that (heuristically) $\lambda_1 \cdot \lambda_2$ is large enough, and would only form linkages with $\lambda_2 > 0$ as λ_1 grows 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 mandatory 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 $\hat{\sigma}$.

E 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 type θ. Our goal is to investigate the robustness 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 8.

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 impose a specific assumption on the distribution of the consumer's type θ, namely that θ is uniformly distributed over the interval

$$\theta \sim U\left[\bar{\theta}/2, \bar{\theta}\right], \tag{35}$$

parametrized by $\bar{\theta} > 0$. (With this assumption, we show below that the existence of pooling equilibria is independent of $\bar{\theta}$, and 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 only exists for linkages in the set Λ^* defined by the inequalities in (38) at the bottom of this section.

Proposition 13 (Pooling Equilibria)

1. *For any $(\lambda_1, \lambda_2) \in [-1/2, \infty)^2$, there exists a pooling equilibrium where all types θ deny consent.*

2. *There exists a pooling equilibrium where all types θ grant consent if and only if $(\lambda_1, \lambda_2) \in \Lambda^*$.*

The first part of Proposition 13 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. Off the equilibrium path (if a consumer grants consent), both firms can hold degenerate beliefs that she is of the worst type for that transaction (i.e., θ_L if $\lambda_t > 0$ and θ_H 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.

In contrast, the second part of Proposition 13 establishes that pooling equilibria where every consumer grants consent do not always exist. Indeed, if a such an equilibrium exists, it can be supported by off-path punishments as above. In particular, a consumer who denies consent will face the worst possible terms of trade in a static game at both t_1 and t_2. 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.

More formally, consider the consumer's decision to grant consent. If she grants consent, she receives the equilibrium utility level

$$U^* (\theta, \mu, \lambda_1) + U (\theta, \theta, \lambda_2) . \tag{36}$$

If she denies consent, under our uniform distribution assumption (35), 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 utility off path is given by

$$U \left(\theta, \hat{\theta} (\lambda_1) , \lambda_1\right) + U \left(\theta, \hat{\theta} (\lambda_2) , \lambda_2\right) . \tag{37}$$

The resulting difference in utility levels (36)-(37) 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

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

for all θ. This set is given by the union of the regions in (λ_1, λ_2) space described in (38) and illustrated in Figure 10. Recall that the set Λ^* is independent of $\bar{\theta}$ under our distribution assumption (35).

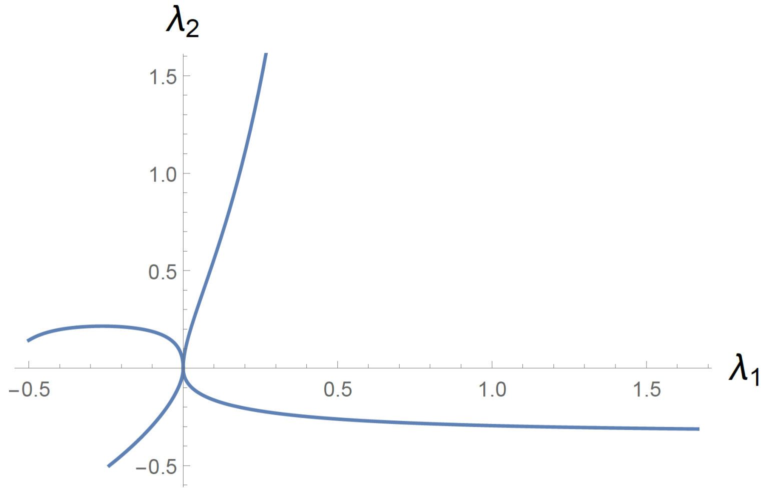

Figure 10: Linkages $(\lambda_1, \lambda_2) \in \Lambda^*$

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 of course distinct. However, as $\hat{\sigma} \to 0$, the set Λ^{CS} becomes a strict subset of Λ^*. This is shown in Figure 11.

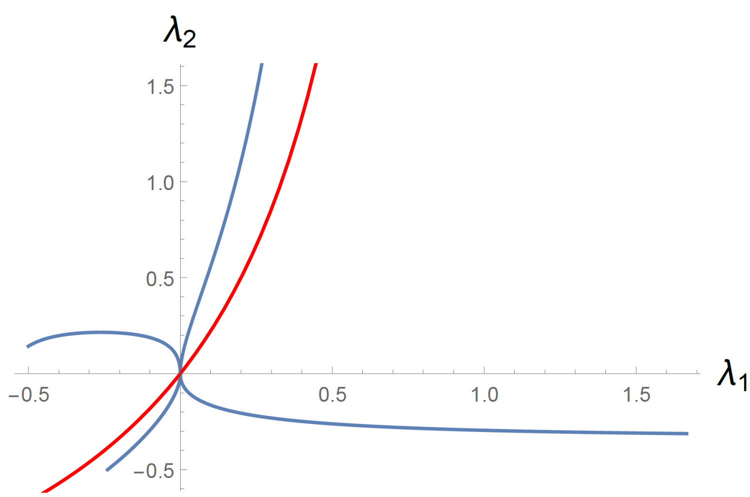

Figure 11: Comparison of Λ^* and Λ^{CS}

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 $(\lambda_1, \lambda_2) \in \Lambda^*$ and the pooling equilibrium without consent on $\Lambda \setminus \Lambda^*$. We then obtain the following corollary.

Corollary 2 *As $\hat{\sigma} \to 0$, 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.

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

$$\begin{gathered}
\left\{\lambda_1 > \frac{4}{3}, \lambda_2 \in \left[\frac{2\sqrt{\frac{8\lambda_1+8\lambda_1^2+3\lambda_1^3}{4+3\lambda_1}} - 3\lambda_1}{3\lambda_1 - 4}, 0\right]\right\} \\
\left\{\lambda_1 \in [0, 4/3], \lambda_2 \in \left[-\frac{2\sqrt{\frac{8\lambda_1+8\lambda_1^2+3\lambda_1^3}{4+3\lambda_1}} - 3\lambda_1}{4 - 3\lambda_1}, 0\right]\right\} \\
\left\{\lambda_1 \in \left[-\frac{1}{2}, \frac{4}{55}\left(3\sqrt{5} - 10\right), \lambda_2 \in \left[-\frac{1}{2}, 0\right]\right]\right\} \\
\left\{\lambda_1 \in \left[\frac{4}{55}\left(3\sqrt{5} - 10\right), 0, \lambda_2 \in \left[-\frac{4\sqrt{\frac{2\lambda_1+2\lambda_1^2+3\lambda_1^3}{4+3\lambda_1}} - 3\lambda_1}{4 - 3\lambda_1}, 0\right]\right]\right\} \\
\left\{\lambda_1 > \frac{2}{3}, \lambda_2 > 0\right\} \\
\left\{\lambda_1 \in \left[0, \frac{2}{3}\right], \lambda_2 \in \left[0, \frac{3\lambda_1}{2 - 3\lambda_1} + 2\sqrt{\frac{3\lambda_1^3 + 4\lambda_1^2 + 2\lambda_1}{(3\lambda_1 - 2)^2 (2 + 3\lambda_1)}}\right]\right\} \\
\left\{\lambda_1 \in \left[-\frac{1}{2}, 0\right], \lambda_2 \in \left[0, \frac{3\lambda_1}{2 - 3\lambda_1} + 2\sqrt{2}\sqrt{\frac{6\lambda_1^3 + 2\lambda_1^2 - \lambda_1}{(3\lambda_1 - 2)^2 (2 + 3\lambda_1)}}\right]\right\}.
\end{gathered} \tag{38}$$

References

Acemoglu, D., A. Makhdoumi, A. Malekian, and A. Ozdaglar (2021). Too much data: Prices and inefficiencies in data markets. *American Economic Journal: Microeconomics*, forthcoming.

Acquisti, A., C. Taylor, and L. Wagman (2016). The economics of privacy. *Journal of Economic Literature 54*, 442–492.

Acquisti, A. and H. R. Varian (2005). Conditioning prices on purchase history. *Marketing Science 24*(3), 367–381.

Ali, S. N., G. Lewis, and S. Vasserman (2019). Voluntary disclosure and personalized pricing. Technical Report 26592, National Bureau of Economic Research.

Aridor, G., Y.-K. Che, and T. Salz (2020). The economic consequences of data privacy regulation: Empirical evidence from gdpr. Technical Report 26900, National Bureau of Economic Research.

Ball, I. (2020). Scoring strategic agents. Technical report, Yale University.

Baye, M. R. and D. E. 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, P. and W. Vergote (2016). Monopoly price discrimination and privacy: The hidden cost of hiding. *Economics Letters 149*, 141–144.

Bergemann, D. and A. Bonatti (2019). Markets for information: An introduction. *Annual Review of Economics 11*, 85–107.

Bergemann, D., A. Bonatti, and T. Gan (2020). The economics of social data. Technical Report 2203R, Cowles Foundation for Research in Economics, Yale University.

Bonatti, A. and G. Cisternas (2020). Consumer scores and price discrimination. *Review of Economic Studies 87*, 750–791.

Calzolari, G. and A. Pavan (2006). On the optimality of privacy in sequential contracting. *Journal of Economic Theory 130*(1), 168–204.

Choi, J., D. Jeon, and B. Kim (2019). Privacy and personal data collection with information externalities. *Journal of Public Economics 173*, 113–124.

Conitzer, V., C. R. Taylor, and L. 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. In *ACM-EC (Economics and Computation) 2016*.

Dosis, A. and W. Sand-Zantman (2020). The ownership of data. Technical report, Toulouse School of Economics.

Fainmesser, I., A. Galeotti, and R. Momot (2020). Digital privacy. Technical report, Johns Hopkins University.

Frankel, A. and N. Kartik (2019). Improving information from manipulable data. *arXiv preprint arXiv:1908.10330*.

Fudenberg, D. and J. M. Villas-Boas (2006). Behavior-based price discrimination and customer recognition. *Handbook on Economics and Information Systems*.

Golrezaei, N., A. Javanmard, and V. Mirrokni (2020). Dynamic incentive-aware learning: Robust pricing in contextual auctions. *Operations Research 69*(1), 297–314.

Ichihashi, S. (2020a). The economics of data externalities. Technical report, Bank of Canada.

Ichihashi, S. (2020b). Online privacy and information disclosure by consumers. *American Economic Review 110*(2), 569–595.

Jann, O. and C. Schottmüller (2020). An informational theory of privacy. *Economic Journal 130*(625), 93–124.

Johnson, G., S. Shriver, and S. Goldberg (2020). Privacy & market concentration: Intended & unintended consequences of the gdpr. Technical report, Boston University Questrom School of Business.

Jullien, B., Y. Lefouili, and M. Riordan (2020). Privacy protection, security, and consumer retention. Technical report, Toulouse School of Economics.

Laffont, J. and J. Tirole (1988). The dynamics of incentive contracts. *Econometrica 59*, 1735–1754.

Liang, A. and E. Madsen (2020). Data and incentives. Technical report, Northwestern University.

Lin, T. (2019). Valuing intrinsic and instrumental preferences for privacy. Technical report, Boston University.

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

Paes Leme, R., M. Pal, and S. Vassilvitskii (2016). A field guide to personalized reserve prices. In *Proceedings of the 25th international conference on world wide web*, pp. 1093–1102.

Shen, Q. and J. M. Villas-Boas (2018). Behavior-based advertising. *Management Science 64*(5), 2047–2064.

Tang, H. (2019). The value of privacy: Evidence from online borrowers. Technical report, HEC Paris.

Taylor, C. R. (2004). Consumer privacy and the market for customer information. *Rand Journal of Economics 35*(4), 631–650.

Tirole, J. (2021). Digital dystopia. *American Economic Review*, forthcoming.

Turow, J. (2008). *Niche envy: Marketing discrimination in the digital age*. MIT Press.

Vestager, M. (2020). Statement to Committee on the Judiciary. *United States House of Representatives*.