Caviar and Yachts: How Your Purchase Data May Come Back to Haunt You

Benjamin Johnson
Carnegie Mellon University

Paul Laskowski
UC Berkeley

Thomas Maillart
UC Berkeley

John Chuang
UC Berkeley

Nicolas Christin
Carnegie Mellon University

ABSTRACT

A new generation of technologies allows firms to track online consumer behavior with increasing granularity, and to share this information with other firms. This promise of information sharing has driven considerable interest from firms; and its potential for monetization has allowed a large number of online and web services to be available free of charge to consumers (in exchange for behavioral tracking). At the same time, information sharing is not without risks, according to a variety of consumer advocates. We examine these risks and benefits using a game-theoretic model of information sharing. Online consumers decide whether to buy a product from two firms in sequence. Depending on the regulatory regime, the first firm may be allowed to sell its purchase information to the second firm. Unlike previous studies, our model employs a continuum of consumer types, revealing effects that have not been highlighted previously. Market outcomes depend critically on whether consumers are myopic, considering each purchase in isolation, or fully rational, considering the effect a purchase will have on future price discrimination. Myopic consumers always lose utility when firms are allowed to share information, though total welfare increases. On the other hand, fully rational consumers behave in a way that drastically limits firms' abilities to price discriminate, reducing total firm profits. As a result, both consumer surplus and overall welfare are higher when firms are allowed to share information. These results indicate that a more nuanced view of the benefits and drawbacks of information sharing might be in order: perhaps rational users do tolerate behavioral tracking not because they are ignorant of privacy implications, but because they perceive an economic benefit in doing so.

1. INTRODUCTION

Meglena Kuneva, Europe's Consumer Commissioner, made the now-famous declaration that "Personal data is the new oil of the Internet and the new currency of the digital world [14]." Merchants, service providers, and data brokers continue to amass huge amounts of personal data, and are using them to improve user experiences and generate profits at the same time. Traditionally, personal data has included basic demographic information, personal net worth estimations from public tax and property ownership records, and high-level lifestyle patterns from magazine subscriptions. But with the advent of electronic commerce on the web, data brokers are now also collecting data on consumer purchasing patterns, usually in collaboration with merchants. Notwithstanding the increased importance of data brokers, detailed purchase transaction data largely remain in the hands of merchants and payment service providers. There is thus significant interest in exploring and exploiting the potential of information sharing between merchants, either directly or indirectly via an intermediary, for better targeting of consumers.

Fueled by this opportunity, and fueling this opportunity at the same time, many online technologies for user tracking and personalized pricing are being developed and deployed. First, online tracking and user profiling has gained rapid technical sophistication in the past few years, going beyond standard web cookies to Flash cookies, evercookies, and canvas fingerprinting techniques, which all are techniques that attempt to provide an increased granularity of user fingerprinting [5]. A recent study by Acar et al. [5] uncovers a new technique called cookie syncing, a process by which two different trackers can link the identifiers they have assigned to the same user. Cookie syncing enables a new business model in conjunction with real-time bidding in an ad auction. Specifically, it allows the bidder and the ad network to refer to the user by the same identifier so that the bidder can place bids on a particular user in current and future auctions. By extension, this method can potentially facilitate information sharing about a user between two merchants, either directly or indirectly via a third party.

The design and delivery of personalized discounts is also being realized in a highly targeted manner. Predictive marketing platforms (e.g., Freshplum [1]) allow an online merchant to make personalized and exclusive discount offers to users who are predicted to leave the site without making a purchase. The predictions are performed in real-time, while the users are still at the site, based on a wide range of information the merchant or its partners have available on each user. By extension, the purchase decisions of the users at this merchant's site can potentially be re-used by the same predictive marketing platform for making predictions for other merchants in the future. This same concept of real-time personalized discounts is being extended to physical stores of major retailers by a number of companies, including Aislelabs [2], Nomi [3], and RetailNext [4].

Business practices like these have attracted considerable interest and investment, but also criticism from a variety of consumer advocates. Policies that govern the use of personal data are often unclear. Consumers may not understand whether a given merchant will share their information, or with what other parties. Fi-

.

nally, consumers may not recognize how their personal data will impact future shopping opportunities. As Odlyzko notes, the primary reason that businesses collect personal data is price discrimination [15]. In its simplest form, our willingness to pay for one product may be correlated with our willingness to pay for a second product. The title of this paper suggests the (untested) hypothesis that buyers of an expensive caviar are also likely to spend more on a yacht. If the correlation were strong enough, a savvy yacht company could observe our purchases of caviar to predict how much we were willing to pay for a yacht, then set the optimal price through a combination of discounts and special offers.

Motivated by these observations, this study will investigate the sharing of purchase data using a stylized game-theoretic model. Our model involves two firms that sell different goods to a set of consumers. Consumers first face one firm and make a decision about whether to buy or not. At that point, the first firm may sell to the second firm the information of which consumers bought its product. Finally, the second firm may use this information to condition its price on whether each consumer purchased the first good. Within this model, we compare two scenarios: a confidential, or "privacy" regime in which firms are forbidden from sharing information with each other, and a disclosure regime, in which one firm may sell information about its customers to another.

Like some previous studies [9,10], we will assume that consumer valuations for each good are perfectly correlated with each other. Although perfect correlation is idealistic, it will allow us to simplify our analysis and bring welfare effects into sharp relief. A more realistic partial correlation would diminish the strength of the effects we find, but likely preserve their direction.

A major advance provided by our work lies in our model of consumers. While previous studies assume that consumers' valuations for each good take on just two values ("high" and "low," respectively) [9, 10, 17], we base our framework around a *continuum* of consumer types with individual valuations. This adds to the realism of our model, avoids artifacts that can emerge from demand discontinuities, and highlights new economic effects.

Our results point to the central role played by consumer sophistication in determining economic outcomes. Guided by behavioral economic studies that demonstrate how consumers behave with regard to their privacy (e.g., [11]), we consider a myopic consumer that makes each purchase decision without considering the effect on future purchases. In this case, we find that information sharing between firms always leaves consumers worse off. The first firm chooses a higher price than it would in the privacy regime, making the data it collects more valuable to the second firm. The second firm segments the market, charging a higher price to those consumers that purchased the first good. While more consumers end up purchasing the second good, the firm keeps more of the average value generated by each sale, so the overall effect on consumers remains negative. In spite of this fact, overall welfare increases as the firms manage to extract more revenue in the disclosure regime.

We then consider the case of fully-rational consumers. When consumers understand that their purchase data can negatively impact their future purchases, they behave in a way that dramatically limits the value of their purchase data. In fact, our model predicts that firms earn less total profit than they would under the privacy regime. Similarly, both consumer surplus and overall welfare are higher under information sharing as long as consumers are fully-rational. This rather counter-intuitive (at least to us) result evidences that a nuanced view of the benefits and drawbacks of information sharing might be in order: perhaps rational users do tolerate behavioral tracking not because they are ignorant of privacy implications, but because they perceive an economic benefit in doing so.

2. RELATED WORK

Our work builds on a growing body of game-theoretic research that examines the sharing of personal information between two firms (Acquisti provides a survey [7]). In a typical setup, consumers contract with one firm followed by the second, and a comparison is made between a privacy regime, in which the sharing of purchase data is forbidden, and a disclosure regime, in which it is allowed. An emergent theme in this lineage is the role played by consumer sophistication: myopic consumers often lose out when firms share information, but the effect is mitigated or even reversed when consumers are fully strategic. Another feature of the studies described here is their focus on high-type-low-type models. That is, consumer valuations for each good are assumed to take on just two possible types. To the best of our knowledge, our study is the first to study the sale of purchase data under a continuous distribution of consumer preferences.

In a canonical paper, Taylor examines the case of two firms when consumer valuations for each good can take on two values, but these valuations are not perfectly correlated with each other [17]. Mirroring our results, he finds that a leading firm may elevate prices in order to gain information about its customers that is useful in the information market. Taylor finds that information sharing may increase or decrease consumer surplus and welfare, depending on the demand specification.

Calzolari and Pavan describe a mechanism design framework, in which firms may offer arbitrary contracts to users [10]. This flexible setup can be used to describe a wide array of scenarios, beyond the case of sequential purchases. In the case of purchasing decisions, the authors find that a leading firm may find it rational to offer privacy guarantees to its customers, which pay a premium in return. This result depends on the full-rationality of consumers and the ability of firms to commit to the relevant contracts. In the general case, the authors note that the welfare effects of privacy regulation are ambiguous.

Acquisti and Varian look at a single monopolist that sells two goods in series [9]. While this is different than the present case of two firms, the authors similarly make the assumption that consumer tastes are perfectly correlated. They find that a firm may extract greater value from consumers when they behave myopically in the market for the first good, but this is not possible when consumers are fully strategic.

Hermalin and Katz provide examples of scenarios in which privacy regulation can enhance welfare, including insurance markets and investments in information gathering [13]. They further argue that intermediary increases in information can either increase or decrease welfare.

Taylor considers a scenario in which collecting information about customers is costly [16]. Without privacy regulation, firms overinvest in collecting personal data, decreasing welfare. Moreover, the effect is exacerbated when firms can sell the information they gather.

Hann et al. argue that unsolicited marketing imposes negative costs on consumers in the absence of privacy regulation [12].

Behavioral Economics of Privacy.

In addition to game theoretic work, a large body of research examines how consumers treat privacy from a behavioral economics perspective. Acquisti and Grossklags argue that consumers deviate from rational privacy decisions for at least three reasons [8]. Consumers have limited information about how firms use personal data. Due to bounded rationality, consumers often rely on simple heuristics when analyzing privacy scenarios. Finally, a variety of cognitive biases may influence privacy decisions.

Acquisti draws on previous work on immediate gratification to argue that consumers may be sophisticated, but time-inconsistent when it comes to privacy decisions [6]. The resulting behavior is myopic in the sense that future costs of sharing information are undervalued. In line with this research, we compute market outcomes for myopic consumers as well as fully-rational ones.

3. MODEL

The actors in our model are two firms and a continuum of consumers. Firm 1 sells good 1, while Firm 2 sells good 2. Each consumer has a valuation of these goods described by a type $\theta \in [0,1]$; and the valuations of good 1 and good 2 for each consumer are linearly related by a constant k. Specifically, a consumer of type θ values good 1 at θ and values good 2 at $k\theta$, where $k \in \mathbb{R}_{\geq 0}$ is a fixed parameter. For now, we assume that consumer types are distributed uniformly in $[0,1]$.

The firms and consumers make choices in a sequential game with imperfect information. Each firm makes choices to maximize its utility. We model consumers in two ways. First, we assume that consumers are myopic, in the sense that they make each purchase decision independently in order to maximize their immediate utility in the current round. Subsequently we perform the analysis under the assumption that the consumers are strategic while believing that their initial purchase information will later be sold.

The game proceeds sequentially in five rounds.

1. Firm 1 chooses a price $p_1 \in [0,1]$ to charge for good 1.

2. Each consumer chooses independently whether to buy good 1 at price p_1. The aggregate result of all consumer choices is a set of good 1 buying types $B_1 \subseteq [0,1]$. The complement $D_1 = [0,1] \smallsetminus B_1$ is the set of good 1 declining types.

3. Firm 1 chooses a price S to offer Firm 2 for revealing the set B_1. Firm 1 may also choose not to share this information.

4. Firm 2 chooses whether to buy B_1 at price S. If buying, it chooses two prices: p_{2B_1} for consumers whose types are in B_1, and p_{2D_1} for consumers whose types are in D_1. Firm 2 may also decline to buy the information, in which case she chooses a single price p_2 to offer all consumers.

5. Each consumer chooses whether to buy good 2 at the price she is offered. The aggregate result is a set of good 2 buying types B_2 with compliment $D_2 = [0,1] \smallsetminus B_2$. In the context of prior choices, B_2 may be further decomposed as a disjoint union of two (possibly empty) intervals: $(B_1 \cap B_2) \cup (D_1 \cap B_2)$. Similarly D_2 may be decomposed as a disjoint union of two (possibly empty) intervals: $(B_1 \cap D_2) \cup (D_1 \cap D_2)$.

The utility of Firm 1 is its derived revenue from the sale of good 1 to consumers, together with the additional revenue S if Firm 1 sells its information.

$$R_1 = \begin{cases} p_1 \cdot |B_1| + S & \text{if Firm 1 sells its information} \\ p_1 \cdot |B_1| & \text{otherwise} \end{cases} . \qquad (1)$$

The utility of Firm 2 is its revenue from the sale of good 2, minus the price S if it chooses to buy the sales information from Firm 1.

$$R_2 = \begin{cases} p_{2B_1} \cdot |B_1 \cap B_2| & \\ +p_{2D_1} \cdot |D_1 \cap B_2| & \\ -S & \text{if Firm 1 sells its info} \\ p_2 \cdot |B_2| & \text{otherwise .} \end{cases} \qquad (2)$$

The myopic consumer evaluates her utility separately in rounds 2 and 5. In round 2, her utility is

$$u_{C2} = \begin{cases} \theta - p_1 & \text{if } \theta \in B_1 \\ 0 & \text{if } \theta \in D_1 \end{cases} ; \qquad (3)$$

while in round 5, her utility is

$$u_{C5} = \begin{cases} k\theta - p_2 & \text{if } \theta \in B_2 \\ 0 & \text{if } \theta \in D_2 \end{cases} . \qquad (4)$$

The strategic consumer evaluate her utility in round 2 with awareness of possible price discrimination in round 5. The utility function for a consumer of type θ is expressed as a function of her combined choices in the following table.

Table 1: Utility of a strategic consumer of type θ, assuming that Firm 1 sells its information.

Consumer Choice	Consumer Utility
$\theta \in B_1 \cap B_2$	$\theta - p_1 + k\theta - p_{2B_1}$
$\theta \in D_1 \cap B_2$	$k\theta - p_{2D_1}$
$\theta \in B_1 \cap D_2$	$\theta - p_1$
$\theta \in D_1 \cap D_2$	0

4. ANALYSIS

We begin our analysis by examining the case of the myopic consumer in the disclosure regime. The myopic consumer makes each purchase decision separately based only on her valuations of the goods. Subsequently, we consider the case of the strategic consumer in the disclosure regime. Finally, we compare these results with their analogues in the Privacy regime. In each case, our analysis proceeds by reverse induction on the rounds of the game.

4.1 Myopic Consumer Best Responses

Assume for this section that each consumer is myopic. That is, she makes a choice in round 2 to maximize her utility only in that round, without considering the future consequences on her utility in future rounds. The utilities for such a consumer are given in Equations (3) and (4) as a function of her choices in rounds 2 and 5, respectively. We now proceed by reverse induction to analyze each actor's preferred strategic choices in response to the choices made by other players in previous rounds.

4.1.1 Round 5

In round 5, each consumer must choose whether or not to buy good 2 at price p_2. Since a consumer of type θ values good 2 at $k\theta$, she will maximize her utility by making the following choice:

$$\begin{cases} \theta \in B_2 & \text{if } k\theta \geq p_2 \\ \theta \in D_2 & \text{if } k\theta < p_2, \end{cases} \qquad (5)$$

where p_2 is used generically to refer to the price charged by Firm 2 to this consumer.

4.1.2 Round 4

In round 4, Firm 2 decides the price p_2 to charge for good 2, based on the information it has available about consumers. To perform the analysis, we must consider separate cases depending on whether Firm 2 chooses to buy the consumer sales information from Firm 1.

- In the case where Firm 2 buys the sales information, Firm 2 will choose two prices: a price p_{2B_1} for consumers who bought good 1 at the price p_1, and a price p_{2D_1} for consumers who declined good 1 at the price p_1. In this case, Firm 2 may independently optimize its revenue from the sale of good 2 to these two sets of consumers.

  First, the set $B_1 \cap B_2$ of consumer types who will buy good 2 at the price p_{2B1} has the form $B_1 \cap \left[\frac{p_{2B_1}}{k}, 1\right]$. Since the consumer is myopic, she will purchase good 1 only on the basis of her valuation θ and hence we know that $B_1 = [p_1, 1]$. We can thus immediately express the revenue from these consumers as $R_2 = p_{2B_1}\left(1 - \max\left\{p_1, \frac{p_{2B_1}}{k}\right\}\right)$. This revenue is maximized (for B_1 consumers) by taking

$$p_{2B_1} = k \cdot \max\left\{p_1, \frac{1}{2}\right\}. \tag{6}$$

  Similarly, the set of consumer types $D_1 \cap B_2$ who will buy good 2 at the price p_{2D_1} has the form $D_1 \cap \left[\frac{p_{2B_1}}{k}, 1\right]$. Again since the consumers are myopic, we have $D_1 = [0, p_1)$. In this case the revenue of Firm 2 can be expressed as $R_2 = p_{2D_1}\left(\max\left\{0, p_1 - \frac{p_{2D1}}{k}\right\}\right)$; and this quantity is maximized for $p_{2D_1} \in [0, k]$ by choosing

$$p_{2D_1} = \frac{kp_1}{2}. \tag{7}$$

  Letting $p_1^* = \max\{p_1, \frac{1}{2}\}$, the total revenue of Firm 2 can be written as

$$\begin{aligned} R_2 &= p_{2B_1} \cdot |B_1 \cap B_2| + p_{2D_1} \cdot |D_1 \cap B_2| - S \\ &= kp_1^* \cdot (1 - p_1^*) + \frac{kp_1}{2} \cdot \frac{p_1}{2} - S \\ &= kp_1^*(1 - p_1^*) + \frac{kp_1^2}{4} - S \end{aligned} \tag{8}$$

- In the case where Firm 2 does not buy the sales information, Firm 2 knows only that consumer types are uniformly distributed in $[0, 1]$. From our analysis of round 5, the myopic consumer of type θ will choose $\theta \in B_2$ only if $k\theta \geq p_2$. Hence the quantity of consumers who buy at price p_2 is $|B_2| = 1 - \frac{p_2}{k}$ and Firm 2's revenue is $R_2 = p_2\left(1 - \frac{p_2}{k}\right)$. This revenue is maximized by choosing

$$p_2 = \frac{k}{2} \tag{9}$$

  achieving the maximum revenue

$$R_2 = \frac{k}{4}. \tag{10}$$

The choice of Firm 2 whether to buy the information from Firm 1 at the price S depends on a comparison between the two corresponding maximum revenues. Comparing Equations (10) and (8), Firm 2 should choose to buy the information whenever

$$\begin{aligned} S &\leq kp_1^*(1 - p_1^*) + \frac{kp_1^2}{4} - \frac{k}{4} \\ &= k\left(p_1^*(1 - p_1^*) - \frac{1 - p_1^2}{4}\right), \end{aligned} \tag{11}$$

where $p_1^* = \max\left\{p_1, \frac{1}{2}\right\}$.

As an illustration, Figure 1 shows the value of consumer information to Firm 2 as a function of p_1. Firm 2 will pay S for consumer sales information relative to the price p_1 as long as the point (p_1, S) lies on or below the curve in the figure. Notice that the value of information is always positive, so then there always exists a profitable price for this exchange of information between the firms.

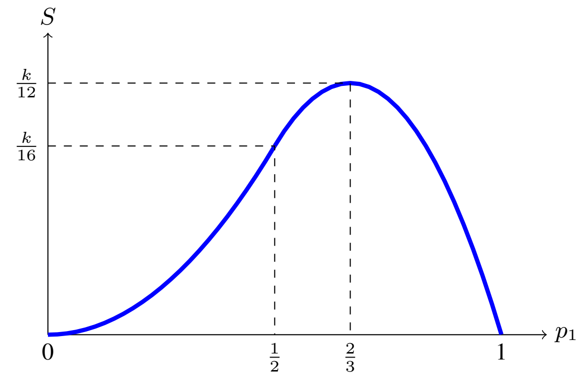

Figure 1: S as a function of p_1, with myopic consumers

4.1.3 Round 3

In round 3, Firm 1 decides a price S to offer Firm 2 for the sales information. By the time Firm 1 makes this choice, the consumers have already made all of their purchase decisions regarding good 1; and the price p_1 and the set B_1 are determined. Hence at this point, Firm 1 is constrained by its utility-maximizing objective to sell the information at the maximum price it can get, provided that the sale generates additional revenue.

From Equation (11), as long as it evaluates to something positive, the price S offered by Firm 1 will be

$$S = k\left(p_1^*(1 - p_1^*) - \frac{1 - p_1^2}{4}\right), \tag{12}$$

where $p_1^* = \max\left\{p_1, \frac{1}{2}\right\}$.

Note that for any such best response price S, the resulting revenue of Firm 2 will be exactly

$$R_2 = \frac{k}{4}. \tag{13}$$

Note from Figure 1 that the maximum value of S occurs when $p_1 = \frac{2}{3}$. This can be explained by considering how p_1 affects Firm 2's market segmentation. As illustrated in Figure 2, Firm 2's revenue is composed of two rectangular regions under the demand curve. The size of these regions will vary for different values of p_1, but the price $p_1 = \frac{2}{3}$ allows the rectangles to have equal widths, which maximizes their total area.

4.1.4 Round 2

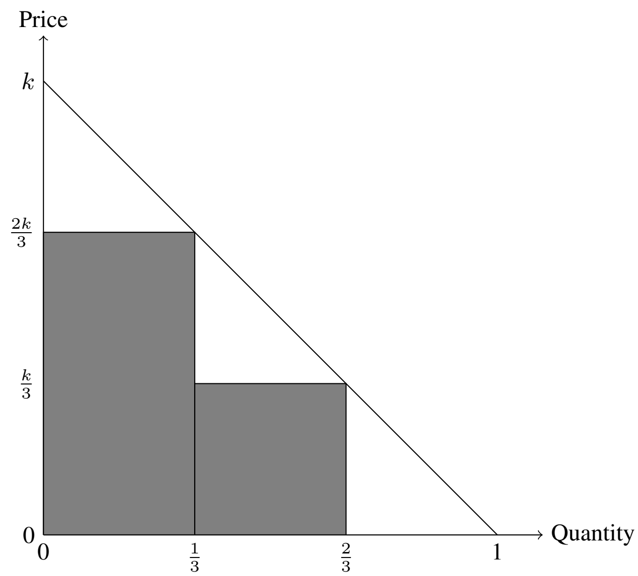

Figure 2: The shaded region gives the revenue of Firm 2, with high price $\frac{2k}{3}$ and low price $\frac{k}{3}$.

In Round 2, the myopic consumer decides whether to buy good 1 at the price p_1 in order to maximize her immediate payoff given by Equation (3). A myopic consumer of type θ will choose to buy good 1 at price p_1 just in case $\theta \geq p_1$, which yields the following aggregate consumer choice set.

$$B_1 = [p_1, 1]. \tag{14}$$

4.1.5 Round 1

In round 1, Firm 1 chooses a price p_1 to offer consumers in exchange for good 1. Since we have determined that Firm 1 will eventually sell its information, its revenue (using Equation (1)), can be written as $R_1 = p_1 \cdot |B_1| + S$; and using Equation (14) this reduces to $R_1 = p_1(1-p_1) + S$.

Note that $p_1 = \frac{1}{2}$ is the price that maximizes revenue from the sale of good 1; and that the sales revenue from good 1 is increasing for every $p_1 < \frac{1}{2}$. Note also from Figure 1 (and the corresponding equation) that the information value S is also increasing for every $p_1 < \frac{1}{2}$. Hence the price that maximizes the sum of these revenues must be at least $\frac{1}{2}$. So for this maximizing price, we will have $p_1^* = \max\left\{p_1, \frac{1}{2}\right\} = p_1$.

Substituting the value of S from Equation (12) we obtain

$$\begin{aligned} R_1 &= p_1(1-p_1) + S \\ &= p_1(1-p_1) + kp_1^*(1-p_1^*) + \frac{kp_1^2}{4} - \frac{k}{4} \\ &= p_1(1-p_1) + kp_1 - \frac{3kp_1^2}{4} - \frac{k}{4}. \end{aligned} \tag{15}$$

Optimizing for p_1 using calculus, we obtain Firm 1's price:

$$p_1 = \frac{2+2k}{4+3k}. \tag{16}$$

This price yields a maximum revenue:

$$R_1 = \frac{(2+k)^2}{4(4+3k)}. \tag{17}$$

4.2 Myopic Consumer Equilibrium

Now that we have determined the optimal price for Firm 1 to choose in round 1, we may use the analysis from earlier rounds to deduce the equilibrium choices and utilities for all other players in the game.

4.2.1 Equilibrium Choices

LEMMA 4.1. *If consumers are myopic in the disclosure regime, then for any $k > 0$, Firm 1 always chooses to sell its consumer sales information to Firm 2; Firm 2 always chooses to purchase the information; and the equilibrium prices are unique, with the following values.*

$$p_1 = \frac{2+2k}{4+3k} \tag{18}$$

$$B_1 = \left[\frac{2+2k}{4+3k}, 1\right] \tag{19}$$

$$S = \frac{k(2+3k)(2+k)}{4(4+3k)^2} \tag{20}$$

$$p_{2B_1} = \frac{2k(1+k)}{4+3k} \tag{21}$$

$$p_{2D_1} = \frac{k(1+k)}{4+3k} \tag{22}$$

$$B_1 \cap B_2 = \left[\frac{2k(1+k)}{4+3k}, 1\right] \tag{23}$$

$$D_1 \cap B_2 = \left[\frac{k(1+k)}{4+3k}, \frac{2k(1+k)}{4+3k}\right] \tag{24}$$

PROOF. Each of the above choice values may be derived explicitly by evaluating equations in the analysis above at the equilibrium value of p_1 given in Equation (16). $\square$

4.2.2 Consumer Surplus

LEMMA 4.2. *The consumer surplus for myopic consumers in the disclosure regime is given by*

$$CS = \frac{(1+k)(4+5k+2k^2)}{2(4+3k)^2} \tag{25}$$

PROOF. The average surplus from good 1 for a buy-1 consumer is $\frac{1-p_1}{2}$, while the number of buy-1 consumers is $1-p_1$. Hence the consumer surplus from the sale of good 1 in equilibrium can be computed as

$$\frac{(1-p_1)^2}{2} = \frac{1}{2}\left(1 - \frac{2+2k}{4+3k}\right)^2 = \frac{(2+k)^2}{2(4+3k)^2} \tag{26}$$

There are $1-p_1$ consumers who purchase good 2 at the high equilibrium price $p_2 = kp_1$; and the average surplus for such a consumer is $\frac{k-kp_1}{2}$. So the consumer surplus from the sale of good 2 at the high price in equilibrium is

$$\frac{k(1-p_1)^2}{2} = \frac{k(2+k)^2}{2(4+3k)^2}. \tag{27}$$

The number of consumers who purchase good 2 at the low equilibrium price $p_{2B_1} = \frac{kp_1}{2}$ is $\frac{p_1}{2}$; and the average surplus from among

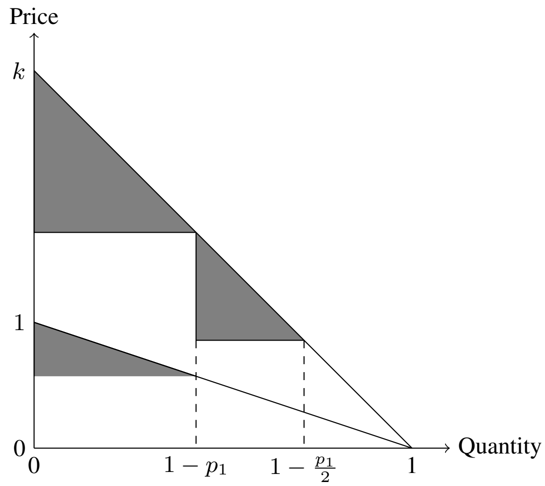

Figure 3: The shaded regions show consumer surplus in equilibrium.

these consumers is $\frac{kp_1}{4}$. So the consumer surplus from the sale of good 2 at the low price in equilibrium is

$$\frac{kp_1^2}{8} = \frac{k(2+2k)^2}{8(4+3k)^2} = \frac{k(1+k)^2}{2(4+3k)^2}. \tag{28}$$

Adding these components together, we obtain a total consumer surplus of

$$\frac{(2+k)^2}{2(4+3k)^2} + \frac{k(2+k)^2}{2(4+3k)^2} + \frac{k(1+k)^2}{2(4+3k)^2}$$
$$= \frac{(1+k)(4+5k+2k^2)}{2(4+3k)^2}. \tag{29}$$

Figure 3 illustrates the three regions of consumer surplus. The diagonal lines are the demand curves for the two goods, and the total consumer surplus is the sum of the shaded regions. The top triangle is the consumer surplus for high-type consumers buying good 2 at the higher price. The middle triangle is the consumer surplus from mid-type consumers buying good 2 at the lower price; and the bottom triangle is the consumer surplus from high-type consumers buying good 1.

□

4.2.3 *Welfare*

LEMMA 4.3. *The wellfare in the disclosure regime with myopic consumers is given by*

$$\frac{1+k}{2}\left(\frac{12+19k+8k^2}{16+24k+9k^2}\right). \tag{30}$$

PROOF. Suppose the player choices are at the unique equilibrium. Then the welfare is computed as

$$\begin{aligned}
&= R_1 + R_2 + \text{Consumer Surplus} \\
&= \left(\frac{k(1+k)}{4(4+3k)} + \frac{1}{4}\right) + \left(\frac{k}{4}\right) + \left(\frac{(1+k)(4+5k+2k^2)}{2(4+3k)^2}\right) \\
&= \frac{1+k}{4}\left(\frac{k}{4+3k} + 1 + \frac{2(4+5k+2k^2)}{(4+3k)^2}\right) \\
&= \frac{1+k}{4}\left(1 + \frac{8+14k+7k^2}{16+24k+9k^2}\right) \\
&= \frac{1+k}{2}\left(\frac{12+19k+8k^2}{16+24k+9k^2}\right)
\end{aligned} \tag{31}$$

□

4.3 Strategic Consumer Best Responses

For this section, we assume that each consumer in round 2 is strategic, and that she believes the firms will exchange her purchase history in the rounds between her two purchases. Our solution concept will be pure strategy perfect Bayesian equilibrium.

4.3.1 *Round 5*

In round 5, each consumer must choose whether or not to buy good 2 at price p_2. Since this is the last round of the game, each consumer is constrained by her own utility-maximizing objective to buy good 2 based on her valuation. So as in the myopic case, a consumer of type θ will maximize her utility by making the following choice:

$$\begin{cases} \theta \in B_2 & \text{if } k\theta \geq p_2 \\ \theta \in D_2 & \text{if } k\theta < p_2, \end{cases} \tag{32}$$

where p_2 is used generically to refer to the price charged by Firm 2 to this consumer.

4.3.2 *Intervals of Consumer Types*

At this point, it is useful to elaborate on the structural properties of sets of consumer types that are defined by choices in a pure strategy equilibrium of the game.

To begin, it is clear that for any k and for any fixed set of choices $\{p_1, S, p_{2B_1}, p_{2D_1}\}$, the best response sets $B_1 \cap B_2$, $B_1 \cap D_2$, $D_1 \cap B_1$, and $D_1 \cap D_2$ are all intervals (possibly empty), because their definitions involve the conjunction of simple linear inequalities in θ. See Tables 2 and 3 for explicit construction of these intervals.

Table 2: Interval constraints for $B_1 \cap B_2$ and $D_1 \cap D_2$

Type θ	Interval Constraint for Consumer Type θ
$B_1 \cap B_2$	$\left[\max\left\{p_1 + p_{2B_1} - p_{2D_1}, \frac{p_{2B_1}}{k}, \frac{p_1+p_{2B_1}}{1+k}\right\}, 1\right]$
$D_1 \cap D_2$	$\left[0, \min\left\{\frac{p_1+p_{2B_1}}{1+k}, \frac{p_{2D_1}}{k}, p_1\right\}\right]$

Thus $B_1 = (B_1 \cap B_2) \cup (B_1 \cap D_2)$ can be written as the disjoint union of two (possibly-empty) intervals. The following lemma tells us that in any pure strategy equilibrium, the best response set B_1 is an interval (possibly empty) that is bounded above by 1. This representation will help us to express in a succinct manner all best responses that involve the set B_1 under a pure-strategy equilibrium assumption.

LEMMA 4.4. *In any pure strategy equilibrium, the best response set B_1 of consumers in Round 2 is an interval of the form $[\theta_1, 1]$ for some $\theta_1 \in [0, 1]$.*

Table 3: Interval Constraints for $D_1 \cap B_2$ and $B_1 \cap D_2$

Type θ	Boundary Constraints for Defining Interval
$0 < k < 1$	
$D_1 \cap B_2$	$\left[\frac{p_{2D_1}}{k}, \min\left\{p_1 + p_{2B_1} - p_{2D_1}, \frac{p_1 - p_{2D_1}}{1-k}\right\}\right]$
$B_1 \cap D_2$	$\left[\max\left\{\frac{p_1 - p_{2D_1}}{1-k}, p_1\right\}, \frac{p_{2B_1}}{k}\right]$
$k = 1$	
$D_1 \cap B_2$	$[p_{2D_1}, p_1 + p_{2B_1} - p_{2D_1}]$ and $p_{2D_1} \leq p_1$
$B_1 \cap D_2$	$[p_1, p_{2B_1}]$ and $p_1 \leq p_{2D_1}$
$k > 1$	
$D_1 \cap B_2$	$\left[\max\left\{\frac{p_{2D_1}}{k}, \frac{p_{2D_1} - p_1}{k-1}\right\}, p_1 + p_{2B_1} - p_{2D_1}\right]$
$B_1 \cap D_2$	$\left[p_1, \min\left\{\frac{p_{2D_1} - p_1}{k-1}, \frac{p_{2B_1}}{k}\right\}\right]$

PROOF. In a pure strategy equilibrium, there is a single set of price choices $p_1, S, p_{2B_1}, p_{2D_1}$ along the equilibrium path. Moreover, the consumer must believe in round 2 that these prices are the correct future choices. So let these price choices be fixed, and consider each consumer's choice in round 2.

First consider the case $k < 1$. The aggregate utility over all game rounds for a consumer whose type θ is in B_1 can be expressed as

$$u_{B_1}(\theta) = (\theta - p_1) + \max\{0, k\theta - p_{2B_1}\}$$

Similarly, the aggregate utility for a consumer whose type θ is in D_1 can be expressed as

$$u_{D_1}(\theta) = \max\{0, k\theta - p_{2D_1}\}.$$

Let

$$w(\theta) = u_{B_1}(\theta) - u_{D_1}(\theta).$$

We claim that w is necessarily increasing. To see this, note that the first part, $\theta - p_1$, is strictly increasing with slope 1; the second part, $\max\{0, k\theta - p_{2B_1}\}$, is non-decreasing; and the last part, $-\max\{0, k\theta - p_{2D_1}\}$, has slope at least $-k$. Since $k < 1$, the overall sum of these components must be increasing for every $\theta \in [0, 1]$.

Since $B_1 = \{\theta : w(\theta) > 0\}$, the previous argument shows B_1 is closed under increases, which in turn implies it must be an interval that contains 1 (or is empty, although this would lead to other contradictions). Note that the idea behind this proof (although not the notation) would largely carry through even within the context of a mixed strategy equilibrium.

Next we consider the case $k \geq 1$. Here we argue by contradiction. Suppose that B_1 is not an interval. Since it is the union of $B_1 \cap B_2$ and $B_1 \cap D_2$, these two intervals must both be non-empty and separated by some non-empty set X. Since $D_1 \cap D_2$ contains 0 and is closed under taking decreases, it cannot be X; and so the only possibility for X is $D_1 \cap B_2$. We thus have a total ordering of the following four sub-intervals of $[0, 1]$:

$$D_1 \cap D_2 < B_1 \cap D_2 < D_1 \cap B_2 < B_1 \cap B_2,$$

with the above juxtaposed neighboring intervals adjacent to each other also on $[0, 1]$.

We will obtain a contradiction to the notion that the highest three intervals are non-empty and in the specified order. First, since $k > 1$, the property that $B_1 \cap D_2$ is adjacent to and below $D_1 \cap B_2$ implies that the upper bound of $B_1 \cap D_2$ is $\frac{p_{2D_1} - p_1}{k-1}$.

Second, the property that $D_1 \cap B_2$ is adjacent to and below $B_1 \cap B_2$ implies that the lower boundary for $B_1 \cap B_2$ is $p_1 + p_{2B_1} - p_{2D_1}$. This in turn implies that in Round 4, Firm 2 will choose the maximum value of p_{2B_1} such that all consumers with types in $B_1 \cap B_2$ purchase at that price. This necessitates $p_{2B_1} = p_1 + p_{2B_1} - p_{2D_1}$ which in turn implies $p_1 = p_{2D_1}$ which in turn implies that the upper bound of $B_1 \cap D_2$ is zero. This gives a contradiction since $B_1 \cap D_2$ was the second-highest-ordered interval in a chain of three non-empty intervals from $[0, 1]$. $\square$

4.3.3 *Round 4*

Having sufficiently structured consumer buying types, we may turn our attention to the choice of Firm 2 in the next round. In round 4, Firm 2 decides the price p_2 to charge for good 2, based on the information it has available about consumers. Again we must consider separate cases depending on whether Firm 2 chooses to buy the consumer sales information from Firm 1.

- In the case where Firm 2 buys the sales information, it will choose two prices: a price p_{2B_1} for consumers who bought good 1 at the price p_1, and another price p_{2D_1} for consumers who declined good 1 at the price p_1. Again Firm 2 may independently optimize its revenue from the sale of good 2 to these two sets of consumers.

  First, the set $B_1 \cap B_2$ of consumer types who will buy good 2 at the price p_{2B1} has the form $B_1 \cap \left[\frac{p_{2B_1}}{k}, 1\right]$. Since B_1 is an interval of the form $[\theta_1, 1]$, (and Firm 2 believes this) the revenue from these consumers can be expressed as $R_2 = p_{2B_1}\left(1 - \max\left\{\theta_1, \frac{p_{2B_1}}{k}\right\}\right)$ and is maximized (for B_1 consumers) by taking

  $$p_{2B_1} = k \cdot \max\left\{\theta_1, \frac{1}{2}\right\}. \tag{33}$$

  Similarly, the set of consumer types $D_1 \cap B_2$ who will buy good 2 at the price p_{2D_1} has the form $D_1 \cap \left[\frac{p_{2B_1}}{k}, 1\right]$. From the form of B_1 given by Lemma 4.4, D_1 must also be an interval bounded below by zero, so the revenue of Firm 2 can be expressed as $R_2 = p_{2D_1}\left(\max\left\{0, \theta_1 - \frac{p_{2D1}}{k}\right\}\right)$; and this quantity is maximized for $p_{2D_1} \in [0, k]$ by choosing

  $$p_{2D_1} = \frac{k\theta_1}{2}. \tag{34}$$

  Letting $\theta_1^* = \max\{\theta_1, \frac{1}{2}\}$, the total revenue of Firm 2 can be written as

  $$\begin{aligned} R_2 &= p_{2B_1} \cdot |B_1 \cap B_2| + p_{2D_1} \cdot |D_1 \cap B_2| - S \\ &= k\theta_1^* \cdot (1 - \theta_1^*) + \frac{k\theta_1}{2} \cdot \frac{\theta_1}{2} - S \\ &= k\theta_1^*(1 - \theta_1^*) + \frac{k\theta_1^2}{4} - S \end{aligned} \tag{35}$$

- The case in which Firm 2 does not buy the sales information is completely identical to the myopic consumer case. The revenue is maximized by choosing

  $$p_2 = \frac{k}{2} \tag{36}$$

  achieving the maximum revenue

  $$R_2 = \frac{k}{4}. \tag{37}$$

The choice of Firm 2 whether to buy the information from Firm 1 at the price S depends on a comparison between the two corresponding maximum revenues. Comparing Equations (35) and (37), Firm 2 should choose to buy the information whenever

$$\begin{aligned} S &\leq k\theta_1^*(1-\theta_1^*) + \frac{k\theta_1^2}{4} - \frac{k}{4} \\ &= k\left(\theta_1^*(1-\theta_1^*) - \frac{1-\theta_1^2}{4}\right), \end{aligned} \tag{38}$$

where $\theta_1^* = \max\left\{\theta_1, \frac{1}{2}\right\}$.

4.3.4 Round 3

In round 3, Firm 1 decides a price S to offer Firm 2 for the sales information. By the time Firm 1 makes this choice, the consumers have already made all of their purchase decisions regarding good 1; and the price p_1 and the set B_1 are determined. Hence at this point, Firm 1 is constrained by its utility-maximizing objective to sell the information at the maximum price it can get, provided that the sale generates additional revenue.

From Equation (38), as long as it evaluates to something positive, the price S offered by Firm 1 will be

$$S = k\left(\theta_1^*(1-\theta_1^*) - \frac{1-\theta_1^2}{4}\right), \tag{39}$$

where $\theta_1^* = \max\left\{\theta_1, \frac{1}{2}\right\}$.

Note that for any such best response price S, the resulting revenue of Firm 2 will be exactly

$$R_2 = \frac{k}{4}. \tag{40}$$

4.3.5 Round 2

Each consumer is offered a price p_1 for good 1 and must choose independently of other consumers whether or not to purchase at this price. From the structure of B_1, we know that there is a single threshold type θ_1 such that a consumer of this type is indifferent between buying and not buying. This constraint implies that

$$\theta_1 - p_1 + \max\{0, k\theta_1 - p_{2B_1}\} = \max\{0, k\theta_1 - p_{2D_1}\} \tag{41}$$

We can now solve for θ_1 by including the two additional constraints from the derivations of p_{2B_1} and p_{2D_1} in round 4. The first such constraint, $p_{2B_1} = k \cdot \max\left\{\theta_1, \frac{1}{2}\right\}$ implies that $\max\{0, k\theta_1 - p_{2B_1}\} = 0$; while the second constraint, $p_{2D_1} = \frac{k\theta_1}{2}$, implies that $\max\{0, k\theta_1 - p_{2D_1}\} = \frac{k\theta_1}{2}$. Combining these constraints yields

$$\begin{aligned} \theta_1 - p_1 + 0 &= \frac{k\theta_1}{2} \\ \theta_1\left(1 - \frac{k}{2}\right) &= p_1 \\ \theta_1 &= \frac{2p_1}{2-k} \end{aligned} \tag{42}$$

4.3.6 Round 1

Here Firm 1 chooses a price $p_1 \in [0,1]$ to charge for good 1. The consumer will buy at this price only if her type is at least $\theta_1 = \frac{2p_1}{2-k}$, and Firm 1 will later sell the purchase history B_1 at the equilibrium price

$$S = k\left(\theta_1^*(1-\theta_1^*) - \frac{1-\theta_1^2}{4}\right)$$

Note from Equation (42) that $\theta_1 \geq p_1$ for every $k \in [0,2]$, which implies that $\theta_1 \geq \frac{1}{2}$ whenever Firm 1 chooses p_1 to maximize its revenue, implying that $\theta_1^* = \theta_1$. Thus

$$\begin{aligned} S &= k\left(\theta_1(1-\theta_1) - \frac{1-\theta_1^2}{4}\right) \\ &= k\left(\theta_1 - \frac{3\theta_1^2}{4} - \frac{1}{4}\right) \\ &= k\left(\frac{2p_1}{2-k} - \frac{3\left(\frac{2p_1}{2-k}\right)^2}{4} - \frac{1}{4}\right) \\ &= \frac{2kp_1}{2-k} - \frac{3kp_1^2}{(2-k)^2} - \frac{k}{4} \end{aligned}$$

The revenue of Firm 1 can be written in terms of p_1 and k as

$$\begin{aligned} R_1 &= p_1(1-\theta_1) + S \\ &= p_1\left(1 - \frac{2p_1}{2-k}\right) + \frac{2kp_1}{2-k} - \frac{3kp_1^2}{(2-k)^2} - \frac{k}{4} \end{aligned}$$

Taking the derivative and setting it to zero yields the equilibrium price

$$p_1 = \frac{4-k^2}{2(4+k)} \tag{43}$$

giving a maximum revenue of

$$R_1 = \frac{1}{4+k}. \tag{44}$$

4.4 Strategic Consumer Equilibrium

Now that we have determined the optimal price for Firm 1 to choose in round 1, we may use the analysis from earlier rounds to deduce the equilibrium choices and utilities for all other players in the game.

4.4.1 Equilibrium Choices

LEMMA 4.5. *If consumers are strategic in the disclosure regime, then for any $k > 0$, Firm 1 always chooses to sell its customer information to Firm 2; Firm 2 always chooses to purchase the information; and the pure strategy equilibrium prices have the following values.*

$$p_1 = \frac{4-k^2}{2(4+k)} \tag{45}$$

$$B_1 = \left[\frac{2+k}{4+k}, 1\right] \tag{46}$$

$$S = \frac{k(1+k)}{(4+k)^2} \tag{47}$$

$$p_{2B_1} = \frac{k(2+k)}{4+k} \tag{48}$$

$$p_{2D_1} = \frac{k(2+k)}{2(4+k)} \tag{49}$$

$$B_1 \cap B_2 = \left[\frac{k(2+k)}{4+k}, 1\right] \tag{50}$$

$$D_1 \cap B_2 = \left[\frac{k(2+k)}{2(4+k)}, \frac{k(2+k)}{4+k}\right] \tag{51}$$

PROOF. Each of the above choice values may be derived explicitly by evaluating equations in the analysis above at the equilibrium value of p_1. $\square$

4.4.2 *Consumer Surplus*

LEMMA 4.6. *The consumer surplus for myopic consumers in the disclosure regime is given by*

$$CS = \frac{4+8k+k^2}{8(4+k)} \tag{52}$$

PROOF. The average surplus from good 1 for a buy-1 consumer is $\frac{1+\theta_1}{2} - p_1$, while the number of buy-1 consumers is $1 - \theta_1$. Hence the consumer surplus from the sale of good 1 in equilibrium can be computed as

$$\begin{aligned}
&\left(\frac{1+\theta_1}{2} - p_1\right)(1-\theta_1) \\
&\left(\frac{1+\frac{2+k}{4+k}}{2} - \frac{4-k^2}{2(4+k)}\right)\left(1 - \frac{2+k}{4+k}\right) \\
&= \frac{(4+k)+(2+k)-(4-k^2)}{2(4+k)} \cdot \frac{2}{4+k} \\
&= \frac{(4+k)+(2+k)-(4-k^2)}{2(4+k)} \cdot \frac{2}{4+k} \\
&= \frac{2+2k+k^2}{(4+k)^2}
\end{aligned}$$

There are $1 - \theta_1$ consumers who purchase good 2 at the high equilibrium price $p_{2B_1} = k\theta_1$; and the average surplus for such a consumer is $\frac{k-k\theta_1}{2}$. So the consumer surplus from the sale of good 2 at the high price in equilibrium is

$$\begin{aligned}
&\frac{k(1-\theta_1)^2}{2} \\
&= \frac{k\left(1-\left(\frac{2+k}{4+k}\right)\right)^2}{2} \\
&= \frac{k}{2}\left(\frac{2}{4+k}\right)^2 \\
&= \frac{2k}{(4+k)^2}
\end{aligned}$$

The number of consumers who purchase good 2 at the low equilibrium price $p_{2B_1} = \frac{k\theta_1}{2}$ is $\frac{\theta_1}{2}$; and the average surplus from among these consumers is $\frac{k\theta_1}{4}$. So the consumer surplus from the sale of good 2 at the low price in equilibrium is

$$\begin{aligned}
&\frac{k\theta_1^2}{8} \\
&= \frac{k\left(\frac{2+k}{4+k}\right)^2}{8} \\
&= \frac{k(2+k)^2}{8(4+k)^2}
\end{aligned}$$

Adding these components together, total consumer surplus is

$$\begin{aligned}
&\frac{2+2k+k^2}{(4+k)^2} + \frac{2k}{(4+k)^2} + \frac{k(2+k)^2}{8(4+k)^2} \\
&= \frac{16+16k+8k^2+16k+4k+4k^2+k^3}{8(4+k)^2} \\
&= \frac{16+36k+12k^2+k^3}{8(4+k)^2} \\
&= \frac{4+8k+k^2}{8(4+k)}.
\end{aligned} \tag{53}$$

□

4.4.3 *Welfare*

LEMMA 4.7. *The welfare in the disclosure regime with myopic consumers is given by*

$$\frac{12+16k+3k^2}{8(4+k)} \tag{54}$$

PROOF. Suppose the player choices are at the unique equilibrium. Then the Welfare is given by

$$\begin{aligned}
&R_1 + R_2 + \text{Consumer Surplus} \\
&= \frac{1}{4+k} + \frac{k}{4} + \frac{4+8k+k^2}{8(4+k)} \\
&= \frac{8+2k(4+k)+4+8k+k^2}{8(4+k)} \\
&= \frac{12+16k+3k^2}{8(4+k)}
\end{aligned} \tag{55}$$

□

4.5 Comparisons with the Privacy Regime

We are interested in how Firm revenue, consumer surplus, and welfare in the disclosure regime compare to the same economic quantities in the privacy regime.

In the privacy regime, the consumer surplus from the sale of good 1 is $\frac{1}{8}$; and consumer surplus from the sale of good 2 is $\frac{k}{8}$. This gives a total consumer surplus in the privacy regime of

$$\frac{1+k}{8} \tag{56}$$

Similarly, welfare in the privacy regime can be computed as

$$\begin{aligned}
&R_1 + R_2 + \text{Consumer Surplus} \\
&= \frac{1}{4} + \frac{k}{4} + \frac{1+k}{8} = \frac{3(1+k)}{8}.
\end{aligned} \tag{57}$$

LEMMA 4.8. *In the myopic consumer case, for any $k > 0$, consumer surplus is lower in the disclosure regime compared to the privacy regime.*

PROOF. From Equation (29), consumer surplus in the disclosure regime is

$$\frac{(1+k)(4+5k+2k^2)}{2(4+3k)^2}.$$

For $k > 0$, we have

$$\begin{aligned}
&\frac{(1+k)(4+5k+2k^2)}{2(4+3k)^2} \\
&= \frac{1+k}{2} \cdot \frac{4+5k+2k^2}{16+24k+9k^2} \\
&< \frac{1+k}{2} \cdot \frac{1}{4} = \frac{1+k}{8}.
\end{aligned}$$

Thus consumer surplus is lower in the disclosure regime when consumers are myopic.

□

LEMMA 4.9. *In the strategic consumer case, for $0 < k < 2$, consumer surplus is higher in the disclosure regime compared to the privacy regime.*

PROOF. From Equation (53), consumer surplus in the disclosure regime is

$$\frac{4 + 8k + k^2}{8(4 + k)}.$$

For $0 < k < 2$, we have

$$\begin{aligned} &\frac{4 + 8k + k^2}{8(4 + k)} \\ &= \frac{(4 + k)(1 + k) + 3k}{8(4 + k)} \\ &= \frac{1 + k}{8} + \frac{3k}{8(4 + k)} \\ &> \frac{1 + k}{8}. \end{aligned}$$

Thus consumer surplus is higher in the disclosure regime when the consumers are strategic.

□

LEMMA 4.10. *In the myopic consumer case, for any $k > 0$, welfare is higher in the disclosure regime, compared to the privacy regime.*

PROOF. From Equation (31), equilibrium welfare in the disclosure regime is given by

$$\begin{aligned} &\frac{1 + k}{2}\left(\frac{12 + 19k + 8k^2}{16 + 24k + 9k^2}\right) \\ &> \frac{3(1 + k)}{8} \end{aligned}$$

Thus welfare is always higher in the disclosure regime when consumers are myopic.

□

LEMMA 4.11. *In the strategic consumer case, for any $k > 0$, welfare is higher in the disclosure regime, compared to the privacy regime.*

PROOF. From Equation (55), equilibrium welfare in the disclosure regime is given by

$$\begin{aligned} &\frac{12 + 16k + 3k^2}{8(4 + k)} \\ &= \frac{(3 + 3k)(4 + k) + k}{8(4 + k)} \\ &= \frac{3(1 + k)}{8} + \frac{k}{8(4 + k)} \\ &> \frac{3(1 + k)}{8} \end{aligned}$$

Thus welfare is always higher in the disclosure regime when consumers are strategic.

□

5. DISCUSSION AND CONCLUSION

As markets for personal information develop, firms increasingly view the data they collect from customers as a primary source of revenue. While the potential for business innovations is great, the overall impact this transformation will have on consumers remains unclear. Our model provides a stylized tool for exploring the economic implications of data sharing between two firms. We find that myopic consumers are made worse off when firms are allowed to share information across all ranges of our parameter space. Forward-looking consumers, on the other hand, regulate their behavior to limit the value of their personal information. As a result, the downstream firm cannot price discriminate to the same extent and firm profits are actually lower than they would be in the privacy regime. Moreover, consumer surplus and overall welfare are both higher when firms are allowed to share information.

Our model predicts that the disclosure regime will yield higher welfare than the privacy regime. This can be understood as a version of the usual neoclassical efficiency result. Our model is stylized in the sense that welfare depends only on how much of each good is sold to consumers, without complicating factors like risk aversion, information gathering costs, and so on. In this environment, perfect information would result in maximum welfare - but zero consumer surplus. To the extent that such a lopsided result is undesirable, our utility functions could be changed to include a penalty for unfairness without changing the basic functioning of the model.

It is worth noting that the model framework by Taylor is similar to ours in that full information would maximize welfare [17]. Nevertheless, Taylor finds conditions under which information sharing decreases welfare. These can be understood as local effects, however. As Hermalin and Katz note, intermediary increases in information can sometimes decrease welfare, even when full-information maximizes it [13]. Similarly, we could specify non-linear demand functions in our framework for which the privacy regime yields greater welfare than the disclosure regime.

While our model assumes that consumer tastes are perfectly correlated between two goods, we believe that similar, albeit weaker, effects would hold for partial correlations. Information sharing gives Firm 2 a signal about what part of the market a potential customer is from. Under partial correlation, the probability distribution on consumer types inferred by Firm 2 would no longer be uniform, but the firm could still leverage this information to extract more revenue from the market. Future research may shed further light on this possibility.

Further extensions to our model may consider more than two firms, or an infinite sequence of firms. Such a framework could allow us to explore, for example, whether sharing of multiple purchase decisions further erodes consumer welfare as firms are able to segment their market with increasing granularity. We may also explore directional results for wider classes of demand functions beyond our linear one.

6. REFERENCES

[1] `http://www.freshplum.com`.

[2] `http://www.aislelabs.com`.

[3] `http://www.nomi.com`.

[4] `http://www.retailnext.net`.

[5] G. Acar, C. Eubank, S. Englehardt, M. Juarez, A. Narayanan, and C. Diaz. The web never forgets: Persistent tracking mechanisms in the wild. In *Proceedings of CCS 2014*, November 2014.

[6] Alessandro Acquisti. Privacy in electronic commerce and the

economics of immediate gratification. In *Proceedings of the 5th ACM conference on Electronic commerce*, pages 21–29. ACM, 2004.

[7] Alessandro Acquisti. The economics of personal data and the economics of privacy. *Background Paper for OECD Joint WPISP-WPIE Roundtable*, 1, 2010.

[8] Alessandro Acquisti and Jens Grossklags. What can behavioral economics teach us about privacy. *Digital privacy*, page 329, 2007.

[9] Alessandro Acquisti and Hal R Varian. Conditioning prices on purchase history. *Marketing Science*, 24(3):367–381, 2005.

[10] Giacomo Calzolari and Alessandro Pavan. On the optimality of privacy in sequential contracting. *Journal of Economic Theory*, 130(1):168–204, 2006.

[11] J. Grossklags and A. Acquisti. When 25 cents is too much: an experiment on willingness-to-sell and willingness-to-protect personal information. In *Proceedings (online) of the Sixth Worskhop on Economics of Information Security (WEIS'07)*, Pittsburgh, PA, 2007.

[12] Il-Horn Hann, Kai-Lung Hui, Sang-Yong T Lee, and Ivan PL Png. Consumer privacy and marketing avoidance: A static model. *Management Science*, 54(6):1094–1103, 2008.

[13] Benjamin E Hermalin and Michael L Katz. Privacy, property rights and efficiency: The economics of privacy as secrecy. *Quantitative Marketing and Economics*, 4(3):209–239, 2006.

[14] M. Kuneva. Roundtable on online data collection, targeting and profiling: Keynote speech, March 2009. Available from `http://europa.eu/rapid/press-release_SPEECH-09-156_en.htm`.

[15] Andrew Odlyzko. Privacy, economics, and price discrimination on the internet. In *Proceedings of the 5th international conference on Electronic commerce*, pages 355–366. ACM, 2003.

[16] Curtis R Taylor. Privacy in competitive markets. Technical report, 2003.

[17] Curtis R Taylor. Consumer privacy and the market for customer information. *RAND Journal of Economics*, pages 631–650, 2004.