

Tilburg University

Consumers' Privacy Choices in the Era of Big Data

Prüfer, Jens; Dengler, Sebastian

Publication date:
2018

Document Version
Early version, also known as pre-print

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
Prüfer, J., & Dengler, S. (2018). *Consumers' Privacy Choices in the Era of Big Data*. (CentER Discussion Paper; Vol. 2018-012). CentER, Center for Economic Research.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Discussion paper

Consumers' Privacy Choices in the Era of
Big Data

By Sebastian Dengler, Jens Prüfer

April, 2018

TILEC Discussion Paper No. 2018-0014
CentER Discussion Paper No. 2018-012

ISSN 2213-9532

ISSN 2213-9419

<http://ssrn.com/abstract=3159028>



Consumers' Privacy Choices in the Era of Big Data*

Sebastian Dengler and Jens Prüfer

Tilburg University[†]

April 5, 2018

Abstract

Recent progress in information technologies provides sellers with detailed knowledge about consumers' preferences, approaching perfect price discrimination in the limit. We construct a model where consumers with less strategic sophistication than the seller's pricing algorithm face a trade-off when buying. They choose between a direct, transaction cost-free sales channel and a privacy-protecting, but costly, anonymous channel. We show that the anonymous channel is used even in the absence of an explicit taste for privacy if consumers are not *too* strategically sophisticated. This provides a micro-foundation for consumers' privacy choices. Some consumers benefit but others suffer from their anonymization.

JEL Codes: L11, D11, D83, D01, L86

Keywords: Privacy, Big Data, Perfect Price Discrimination, Level-k thinking

*We are grateful to seminar audiences at Tilburg, Télécom ParisTech, ZEW Mannheim (14th ICT Conference), Carlos III Madrid (2016 ENTER Jamboree), ETH Zürich, Louvain (IODE workshop), Harvard (ISNIE conference), Tilburg (CLEEN meeting), Toulouse, and the ESNIE spring school (2014). We are especially thankful to Alessandro Acquisti, Jan Boone, Eric van Damme, Clemens Fiedler, Ricard Gil, Víctor H. González-Jiménez, Heiko Karle, Christian König gen. Kersting, Marian W. Moszoro, Valerio Poti, Jan Potters, Renata Rabovič, Christoph Schottmüller, Florian Schütt, Katja Seim, Gyula Seres, Giancarlo Spagnolo, Vatsalya Srivastava, and Birger Wernerfelt, who provided valuable feedback on earlier versions of this paper. All errors are our own.

[†]Dengler: Tilburg Institute for Law, Technology, and Society (TILT), CentER, TILEC; Prüfer: Department of Economics, CentER, TILEC; Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands; s.dengler@tilburguniversity.edu, j.prufer@tilburguniversity.edu.

Two recent technological developments are revolutionizing seller-buyer transactions. First, aided by information and communication technologies (ICTs), sellers have the capability to analyze huge datasets with very detailed information about individual consumers' characteristics and preferences. Second, such data sets are increasingly available, owing to the fact that more economic and social transactions take place supported by ICTs, which easily and inexpensively store the information they produce or transmit.¹ These concurrent developments constitute the rise of big data (Mayer-Schönberger and Cukier 2013). They imply that sellers can make consumers ever more tailored contract offers, which fit their individual preferences or consumption patterns, approaching first-degree (or perfect) price discrimination, as the limit case (Einav and Levin 2014).

Because first-degree price discrimination can deprive consumers of all surplus from the transaction, they may want to protect their privacy and hide their willingness-to-pay (WTP) from sellers with market power by employing anonymization techniques. But anonymization is costly: it can come at an explicit cost or at an opportunity cost. Consumers are at a second disadvantage, compared to sellers, because they “will often be overwhelmed with the task of identifying possible outcomes related to privacy threats and means of protection. [...] Especially in the presence of complex, ramified consequences associated with the protection or release of personal information, our innate bounded rationality limits our ability to acquire, memorize and process all relevant information, and it makes us rely on simplified mental models, approximate strategies, and heuristics” (Acquisti and Grossklags 2007, p.369).

¹Data analytics firms collect and analyze huge commercial databases on consumers, offering help to marketers. For instance, *Axiom's* “database contains information about 500 million active consumers worldwide, with about 1,500 data points per person. That includes a majority of adults in the United States” (The New York Times 2012). Smartphone apps with millions of users, such as *Shopkick*, reward users for checking into stores, scanning products, visiting the dressing rooms, and so forth. *Amazon* recently was issued a patent on a novel *Method and System for Anticipatory Package Shipping* (Patent number US008615473 (December 24, 2013), <http://pdfpiw.uspto.gov/.piw?docid=08615473>). “So Amazon says it may box and ship products it expects customers in a specific area will want – based on previous orders and other factors – but haven't yet ordered” (Wall Street Journal Blog 2014).

In our model, we study the effects of perfect price discrimination on consumers' choices and welfare when anonymization is possible but costly. We explicitly account for the discrepancy between cognitively challenged consumers and a seller whose strategic capabilities outperform them and investigate how limited strategic sophistication affects outcomes.

Our main contribution is to show under which conditions a costly privacy-protective sales channel is used even if consumers do not have an explicit taste for privacy, and how this depends on consumers' sophistication. We thereby provide a micro-foundation for consumers' privacy choices when facing a seller with access to big data.

We construct a model where a mass of consumers with heterogeneous WTP for a product faces a monopolistic seller. Consumers can decide between two channels to buy the product from the seller. The direct channel (D) makes use of all personal information that the seller has about every single consumer. We assume that perfect price discrimination is feasible for the seller in channel D and that this channel economizes on transaction costs, which we normalize to zero. The anonymous channel (A) protects consumers' privacy by hiding individual identities, but comes at a cost, which we denote by s .² As a consequence, perfect price discrimination is infeasible for the seller, who responds best by setting a uniform price for channel A.

Our model therefore represents a situation *after* a long period of consumers *not* using anonymization techniques (due to neglect or lack of suitable technologies). Hence, the seller could acquire data shedding light on individual consumers' preferences, be it via collecting such information in the past (e.g. Amazon) or via buying such information from an intermediary (e.g. Google, Acxiom). However, as consumers decide about anonymization, the seller can neither directly influence their channel choice nor close down the anonymous channel.

²Consumers may need to pay for or install privacy-protective software, experience lower connection speed due to encrypted transmissions, or otherwise increase transaction cost (e.g. by shopping offline with additional costs if done by cash payments).

In the model, consumers first choose between channel D and channel A. Second, the seller sets prices in both channels. Third, every consumer decides whether to buy for the price offered to her, or not. Our analysis is based on a model of limited strategic sophistication, called *level- k thinking*, which was introduced by Stahl and Wilson (1994; 1995) and Nagel (1995). Models with level- k thinking are defined recursively, starting with so-called “naïve” level-0 players which employ a “naïve” (often random) strategy. Level-1 players then best respond to the level-0 strategy, level-2 players to the level-1 strategy, and so forth.³ A sizable literature has developed that explores level- k thinking theoretically and empirically.⁴ The literature has found strong experimental support for level- k thinking and suggests values for k of one or two (Camerer, et al. 2004; Crawford and Iriberry 2007b).

Compared to the behavior-based price discrimination literature, where typically either unlimited strategic sophistication or complete naïveté of consumers is assumed, we zoom in and provide an analysis of behavior when players have *some* strategic sophistication.⁵ We model consumers’ cognitive constraints by their ability to anticipate k strategic iterations. The seller is able to outperform them in strategic thinking (i.e. has a level of $k + 1$) due to superior access to data and computing power. Whether k is relatively low, as suggested by the empirical behavioral literature, or rather high turns out to crucially matter for our results.

We show that the higher consumers’ level of sophistication, the *higher* the optimal price will be in the anonymous channel A. Consumers anonymize if their valuation for the product exceeds the expected price plus the anonymization cost. But when consumers decide about buying, at Stage 3, those anonymization

³While most of this literature analyzes games with symmetric decisions between (e.g. the beauty contest game (Nagel 1995)), we will adapt the concept slightly to the asymmetric situation of our model where the seller has a different set of actions than the consumers.

⁴See Ho et al. (1998), Costa-Gomes et al. (2001), Crawford (2003), Camerer et al. (2004), Costa-Gomes and Crawford (2006), Crawford and Iriberry (2007a), and Goldfarb and Yang (2009), among others.

⁵What we call “unlimited strategic sophistication”, is often referred to as “perfect rationality”. However, players with limited strategic sophistication still act rationally given their (potentially wrong) beliefs, which is why we avoid the terms of “perfect” and “imperfect” rationality.

costs are sunk. Hence, the best response of the seller is to increase the price above the one consumers expected. If the level of sophistication rises in the population, consumers expect to be offered the product for a higher price in the anonymized market. Hence, consumers with medium but not high WTP do not choose channel A at Stage 1, preempting net losses. Consequently, the seller has an incentive to increase the price in the anonymized market even more, because he infers that only consumers with high WTP have chosen channel A at Stage 1.

We further show that, with any positive cost of anonymization, the anonymized market completely unravels for all sophistication levels $k \geq \bar{k}$, where \bar{k} is a finite number. Hence, unlimited strategic sophistication is not a necessary condition for market unravelling. However, if consumers' k is sufficiently low, only a part of the market unravels and the anonymized sales channel can persist, serving consumers with high WTP. Among those who use the anonymous sales channel, some consumers suffer from net losses because prices turn out to be higher than expected, but consumers with a very high WTP may get some surplus. Thereby, this model offers a micro-foundation for consumers' privacy choices: some consumers rationally use costly anonymization techniques even without an exogenous taste for privacy. Because a share of the anonymization cost could be interpreted as a fee that an intermediary can appropriate, this model also suggests that running an anonymous sales channel competing with a channel that tracks individuals and uses all personal data can be a profitable business model when consumers have limited strategic sophistication.

Related Literature: First-degree (or perfect) price discrimination requires complete information of a seller about a specific consumer's WTP and was introduced into the economics literature by Pigou (1920). However, due to the very high information demand and the rather straightforward allocative and distributional implications, perfect price discrimination has not received a lot of

scholarly attention and has been dismissed as a mere theoretical construct.⁶

More prominent are models of “behavior-based price discrimination.” Most of this literature focuses on third-degree price discrimination assuming that a seller learns about the WTP of a (re-)identifiable consumer after the first purchase. The idea is that, if a consumer bought a good at a certain price, the seller learns that this consumer’s WTP must have weakly exceeded that price and raise the price for her in the future. If consumers anticipate this, they may adjust behavior in early periods and postpone purchases to avoid future price increases, or wait for future price cuts (Villas-Boas 2004). In such cases, firms benefit from stricter privacy regulations if they lack commitment power to bind themselves to not increase prices after initial purchases (Taylor 2004).⁷

However, lending support to the early conclusion of Odlyzko, “that in the Internet environment, the incentives towards price discrimination and the ability to price discriminate will be growing” (Odlyzko 2003, p.365), online vendors and other retailers have already gone much further (see Footnote 1) and can approximate fully personalized prices more than ever. It has been shown empirically that “targeted advertising” techniques increase purchases (Luo et al. 2014), prices (Mikians et al. 2012), and sellers’ profits (Shiller 2013). Some consumers, however, feel repelled by this development and want to have control over their personal data back.⁸ Many place a value on their privacy (Tsai et al. 2011).

The early theoretical literature about the economics of privacy, being based on the Chicago school argument that more information available to market participants increases the efficiency of markets, has underlined the negative welfare effects of hiding information from sellers (Posner 1978; Stigler 1980; Posner

⁶For instance, the standard industrial organization textbook, Tirole (1988), spends three of its more than 1100 pages on perfect price discrimination.

⁷For an overview of this strand of literature, see Fudenberg et al. (2006).

⁸Goldfarb and Tucker (2012) study three million observations between 2001 and 2008 and find that refusals to reveal their income in an online survey have risen over time. Tucker (2014) finds in a field experiment that, when Facebook gave users more control over their personally identifiable information, users were twice as likely to click on personalized ads.

1981).⁹ A lot of progress has been made since then: “With so many people making extreme claims in discussions of privacy and related public policy, and with so little understanding of the underlying economics, it is important to identify the fundamental forces clearly. A central fact is that, contrary to the Chicago School argument, the flow of information from one trading partner to the other can reduce ex post trade efficiency when the increase in information does not lead to symmetrically or fully informed parties” (Hermalin and Katz 2006, p.229).

A related issue are the choices of firms that own personal information about consumers and can decide to disclose it to another firm (Taylor 2004; Acquisti and Varian 2005; Casadesus-Masanell and Hervas-Drane 2015). In interactions between an upstream and a downstream firm for whose products consumers’ WTP is positively correlated, the upstream firm will maintain full privacy of its customers if conditions on the upstream firm’s preferences about the downstream firm and on the downstream relationship itself are met (Calzolari and Pavan 2006). However, if one condition is not met, the upstream firm can find it optimal to disclose its customer list to the downstream firm (sometimes even for free), which need not be negative to consumers but could still yield a Pareto improvement (Calzolari and Pavan 2006).

A core question studied in these papers is, what the welfare consequences of privacy or disclosure are, and who should own the property rights of consumers’ personal data (Hermalin and Katz 2006). The answers given have been ambiguous and depend on the specific application of the papers. Recently, the focus has shifted more towards privacy choices of consumers (Conitzer et al. 2012) and the role of platform intermediaries (de Corniere and De Nijs 2014; Bergemann and Bonatti 2015).¹⁰

⁹Even earlier, Warren and Brandeis (1890) study privacy as “right to be let alone”, a point later discussed by Varian (1997) in the context of annoyance from telemarketing.

¹⁰For a recent overview of the growing literature on the economics of privacy, see Acquisti et al. (2016). Related to our approach, Daughety and Reinganum (2010) develop a model of demand for privacy without assuming a direct preference for it. But they study a public goods context and not a product market.

With few exceptions, however, cognitive constraints of consumers have not been incorporated by theoretical studies of markets driven by big data. Taylor (2004), Acquisti and Varian (2005), and Armstrong (2006) assume the existence of a group of unlimitedly sophisticated consumers and a group of naïve consumers. The latter do not foresee that they may want to trade in the future again and, because of this negligence, ignore the negative effects of disclosing personal data. Hence, if consumers are naïve, a seller may oppose stricter regulations as no commitment device is needed (Taylor 2004). In our model, we allow for a more nuanced, marginal analysis of consumers' sophistication.¹¹

The remainder of the paper is organized as follows. In Section 1, we construct a model, which is analyzed in Section 2. Section 3 studies welfare and payoff consequences of changing the level of sophistication k and the anonymization cost s . Section 4 is dedicated to alternative model specifications, covering the beliefs of naïve consumers, heterogeneous costs of anonymization, effects of increasing competition, and a temporal interpretation of the model. Section 5 concludes.

1 Model

We consider an economy where a monopolistic seller of a single consumption good faces a unit mass of atomistic consumers who can buy at most one unit of the good and cannot resell it to each other.¹² Abstracting from potential fixed costs, we assume that the monopolist can produce the good at constant marginal cost $c \geq 0$. Consumers have heterogeneous valuation v for the good, where $v \sim \mathcal{U}[0, 1]$ and can approach the seller in two different ways: directly (referred to as channel D) or after making use of an anonymization technique (channel A). Consumers choosing channel D incur no cost and the seller perfectly

¹¹The need to include cognitive constraints into economic models of privacy is spurred by empirical findings about the so-called *privacy paradox* (Norberg et al. 2007), an apparent discrepancy between people's privacy attitudes and their actual behavior. For a summary of the privacy paradox literature, see Acquisti et al. (2016).

¹²We discuss the case of monopolistic competition in Section 4.3.

knows their individual valuation. Consumers choosing channel A, on the other hand, incur cost $s > 0$ and their individual valuation is hidden from the seller. We assume that consumers do not have any exogenous taste for privacy and that they choose direct channel D in case of indifference between both channels.

After consumers have made choices between the channels, the seller sets prices based on the information he has. In channel D the seller can set personalized prices $p_i(v)$ conditional on each consumer's valuation. However, in channel A, due to the anonymization technique consumers used, the seller can only set a uniform price p_A for all consumers.

Finally, consumers decide whether they want to buy the good at the price the seller posted for them. In case of indifference, we assume they buy the product. Outside options yield zero payoff (except for costs incurred within the game before opting out). The timing of the model is summarized as follows:

- Stage 1 (*Anonymizing*): Consumers choose channel D or channel A and incur costs of 0 or s , respectively. Indifferent consumers choose channel D.
- Stage 2 (*Pricing*): The seller sets prices $p = \{p_i(v), p_A\}$, where $p_i(v)$ are personalized prices in channel D, and p_A is the uniform price in channel A.
- Stage 3 (*Buying*): Consumers decide whether to buy the good for the offered price. Indifferent consumers are assumed to choose buying the good.

The distribution of v (and hence the demand function), the monopolist's cost structure (and hence the supply function), the cost for anonymization s as well as the timing of the game are common knowledge among all players.

Explicitly modeling consumers' cognitive constraints, we assume that all consumers have the same limited level of strategic sophistication, denoted by $k \in \mathbb{Z}_0^+$. The seller, however, outperforms them in terms of sophistication and has a level of $k + 1$. As in Nagel (1995), players with a level of $k > 0$ will generally act as if they believe that all other players had a level of strategic sophistication exactly

one level below their own level. However, Nagel (1995) considers a setting where all players are symmetric and have the same action sets. As our model has one player (the seller) whose action set differs from everyone else's, and whose best response is therefore different, we adapt the concept slightly.

While we maintain that consumers believe that all other consumers are one level less sophisticated, we deviate in assuming that consumers expect the seller to share their level of sophistication. More formally, consumers form the beliefs $\mathbb{E}_i(k_{j \neq i}) = k_i - 1 = k - 1$ for j being a consumer and $\mathbb{E}_i(k_S) = k_i = k$ for S being the seller. Thus, consumers implicitly think of the seller as responding optimally to their own belief about the sophistication of all other consumers. This assumption is in turn based on the atomistic nature and the resulting insignificance of any individual consumer for the seller's choice. The seller whose $k_S = k + 1$, however, forms the belief $\mathbb{E}_S(k_{j \neq S}) = k_S - 1 = k$ for j being any consumer, in line with Nagel (1995).

Their cognitive limitations in belief formation notwithstanding, players still act rationally by pursuing strategies which maximize their utility *given* beliefs. Hence, we can solve the game by backward induction, but our solution concept differs from a Perfect Bayesian Equilibrium (which would be appropriate if all players had unlimited strategic sophistication) because it does not need to be the case that all expectations about others' strategies are eventually confirmed. Notably, we allow consumers' beliefs about the prices in both channels to not coincide with the prices the seller eventually sets. Hence, we do not *impose* that $\mathbb{E}(p|D) = (p|D)$ and $\mathbb{E}(p|A) = (p|A)$, where $(p|D)$ and $(p|A)$ denote the price after having chosen channel D or channel A, respectively. However, we do impose that players restrict their beliefs about possible prices to the support of the distribution of v , i.e. $\mathbb{E}(p|D) \in [0, 1]$, and $\mathbb{E}(p|A) \in [0, 1]$.

A consumer's strategy is a mapping from her valuation for the good v , her level of strategic sophistication k , and the exogenous parameters s and c to her action space $C \times B$, where $C = \{\text{Channel D}, \text{Channel A}\}$ denotes her set of

choices in the anonymizing stage (Stage 1) and $B = \{(Buying|p), (NotBuying|p)\}$ denotes her set of choices in the buying stage (Stage 3), where $p = p_i(v)$ after having chosen channel D and $p = p_A$ after having chosen channel A.

The seller's strategy is a mapping from his level of sophistication $k + 1$, and the exogenous parameters s and c to a set of prices $p = \{p_i(v), p_A\}$, where $p_i(v)$ are personalized prices he can condition on his knowledge about individual consumers in channel D, and p_A is a uniform price for all consumers in channel A.

The game is solved by backward induction. As models with level- k thinking are best solved recursively, the analysis starts with the case where consumers have a strategic sophistication level of $k = 0$ and form a so-called naïve belief. The seller, having a level of $k = 1$, believes (correctly) that all consumers have a level of $k = 0$. Thus, his level-1 best response is also objectively optimal. In later parts of the analysis, consumers have a level of $k > 0$ and believe that all other consumers have a level of $k - 1$, but that the seller employs the level- k best response. The seller, with a level of $k + 1$, will again be the only one with an objectively correct belief and his $k + 1$ -level response is again objectively optimal.

2 Analysis

Stage 3 – Buying: A utility-maximizing consumer buys the good if the price she has to pay does not exceed her valuation, i.e. if, and only if,

$$v \geq p \in \{p_i(v), p_A\}. \quad (1)$$

If she has chosen channel D, the price will be an individualized price $p_i(v)$, and if she has chosen channel A, she will receive the same uniform price p_A as all other consumers who have chosen channel D.

Stage 2 – Pricing: A profit-maximizing seller sets individual prices p_i for all consumers in channel D (denoted by set \mathcal{C}_D) and one uniform price p_A for all

anonymized consumers in channel A (denoted by set \mathcal{C}_A). Knowing v precisely for all consumers in \mathcal{C}_D , and not selling below marginal cost, the seller sets

$$p_i^*(v) = \max\{v, c\} \text{ for all } i \in \mathcal{C}_D, \quad (2)$$

Despite being uninformed about individual valuations v of consumers in \mathcal{C}_A , the seller can infer which consumers are in \mathcal{C}_A due to his higher level of strategic sophistication and set p_A accordingly. We therefore analyze consumers' general Stage 1 behavior first in order to inform the seller's pricing decision in channel A.

Stage 1 – Anonymizing: Consumers use the anonymization technique of channel A if the expected utility of doing so exceeds the expected utility of the direct channel D, i.e. if, and only if, $\mathbb{E}(u_i(A)) > \mathbb{E}(u_i(D))$, where

$$\mathbb{E}(u_i(D)) = \max\{v - \mathbb{E}(p|D), 0\}, \quad (3)$$

$$\mathbb{E}(u_i(A)) = \max\{v - \mathbb{E}(p|A) - s, -s\}. \quad (4)$$

The first value in each set in (3) and (4) reflects the expected payoff the consumer receives if she buys the product at Stage 3. The second value reflects the payoff of subsequently choosing not to buy the product. Although consumers might be limited in their strategic sophistication, they nonetheless understand the nature of the two channels, i.e. they realize that the seller has no incentive to decrease the price below their valuation in channel D and that the seller can only ask for a uniform price in channel A. Hence, consumers, irrespective of their level of strategic sophistication, form the price expectation for channel D

$$\mathbb{E}(p|D) = p_i^*(v) = \max\{v, c\}, \quad (5)$$

correctly expecting to be left with no surplus in channel D. Their exact expectation for the uniform price in channel A, though, still depends on their level of

strategic sophistication. Hence, we only substitute $\mathbb{E}(p|A) = \mathbb{E}(p_A)$ yielding:

$$\mathbb{E}(u_i(D)) = \max\{v - \max\{v, c\}, 0\} = 0, \quad (6)$$

$$\mathbb{E}(u_i(A)) = \max\{v - \mathbb{E}(p_A) - s, -s\}. \quad (7)$$

This implies that consumers choose channel A if, and only if,

$$\max\{v - \mathbb{E}(p_A) - s, -s\} > 0. \quad (8)$$

Because $s > 0$, this can only hold if $v > \mathbb{E}(p_A) + s \equiv \hat{v}$, where \hat{v} denotes the endogenous threshold dividing the population of consumers into \mathcal{C}_D and \mathcal{C}_A .

Lemma 1 (Anonymization Threshold). *There exists a threshold $\hat{v} = \mathbb{E}(p_A) + s$ that denotes the valuation of a consumer who is indifferent between both channels at Stage 1. Consumers with $v > \hat{v}$ choose channel A; consumers with $v \leq \hat{v}$ choose channel D, i.e. $\mathcal{C}_D = [0, \hat{v}]$ and $\mathcal{C}_A = (\hat{v}, 1]$.*

Stage 2 – Pricing (revisited): Having a higher level of strategic sophistication than the consumers, the seller correctly infers \hat{v} and hence knows that $\mathcal{C}_A = (\hat{v}, 1]$. As he further anticipates that consumers will buy the product at Stage 3, if, and only if, $v \geq p_A$, he can easily infer demand $q_A(p_A)$ in channel A:

$$q_A(p_A) = \begin{cases} 0 & \text{if } p_A > 1, \\ 1 - p_A & \text{if } 1 \geq p_A > \hat{v}, \\ 1 - \hat{v} & \text{if } \hat{v} \geq p_A. \end{cases} \quad (9)$$

Charging $p_A = \hat{v}$ dominates all prices $p'_A < \hat{v}$ because any price below \hat{v} decreases profits per unit sold without an increase in quantity to counter the loss. Thus,

by setting $p_A = \hat{v}$, the seller can guarantee himself profits from channel A of:

$$\pi_A(\hat{v}) = q_A(\hat{v})(\hat{v} - c) = (1 - \hat{v})(\hat{v} - c). \quad (10)$$

However, the seller could also charge a price $p_A > \hat{v}$, depending on where \hat{v} lies exactly. Suppose for the moment that the entire consumer population uses channel A (i.e. $\hat{v} = 0$), which is identical to the case of a monopolist unable to engage in price discrimination. Let us denote the globally profit-maximizing price in this case by $p_M = \frac{1+c}{2}$. Then, there are three different cases for the location of the anonymization threshold \hat{v} (shown in Figure 1) compared to p_M :

- (a) The anonymization threshold is below the monopoly price ($\hat{v} < p_M$).
- (b) The anonymization threshold is equal to the monopoly price ($\hat{v} = p_M$).
- (c) The anonymization threshold is above the monopoly price ($\hat{v} > p_M$).

In cases (a) and (b), the globally profit-maximizing price p_M is located within \mathcal{C}_A and hence remains the optimal price to set. The only consumers that are not in \mathcal{C}_A are those that the seller would not have served even if they had anonymized themselves. Only in case (c), where the globally profit-maximizing price p_M is not located within \mathcal{C}_A anymore, any price below the anonymization threshold \hat{v} is at least dominated by setting the price equal to \hat{v} . The seller also has no incentive to raise the price above \hat{v} as profits are strictly decreasing to either side of the global maximum at p_M due to the strict concavity of the profit function. Hence, in this case the optimal price p_A^* is equal to \hat{v} .

Lemma 2 (Optimal Pricing Strategy). *The optimal strategy of the seller consists of a set of prices $\{p_i^*(v), p_A^*\}$ in channel D and channel A, respectively, where $p_i^*(v) = \max\{v, c\}$ and $p_A^* = \max\{\hat{v}, p_M = \frac{1+c}{2}\}$.*

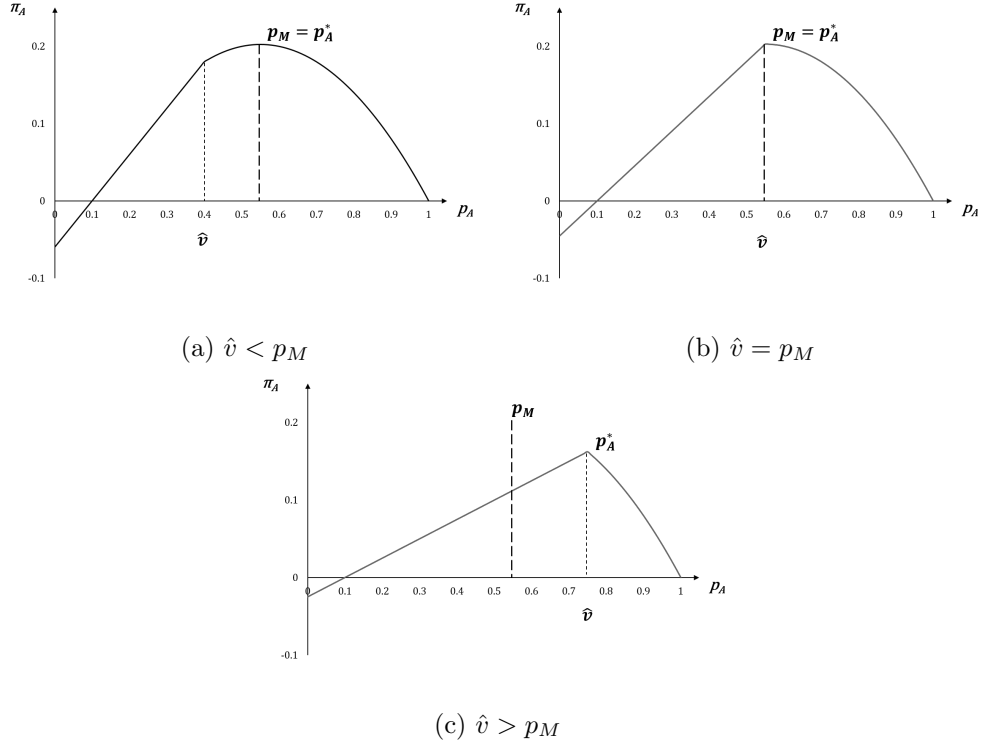


Figure 1: Profits in channel A for different locations of \hat{v} with parameters $v \sim \mathcal{U}[0, 1]$, $c = 0.1$

This implies that the seller sets a higher price than consumers had expected:

$$p_A^* \geq \hat{v} = \mathbb{E}(p_A) + s > \mathbb{E}(p_A). \quad (11)$$

With unlimited strategic sophistication, $\mathbb{E}(p_A) = p_A^*$ would be required in equilibrium, leading to a contradiction. Because only beliefs about off-equilibrium paths can be wrong in a Perfect Bayesian Equilibrium, we conclude that if all players had unlimited strategic sophistication, channel A would remain unused.

However, with limited strategic sophistication such discrepancy is possible. This is due to the fact that s will be a sunk cost for consumers at Stage 3, which the seller can exploit via increasing the price by exactly s , compared to their expectations. Consumers, due to their limited strategic sophistication, cannot anticipate the seller's strategic response.

Stage 1 – Anonymizing (revisited): The last missing piece to fully characterize the solution is the formation of consumers’ price expectations in channel A, $\mathbb{E}(p_A)$, in Stage 1. We determine these recursively and will start with consumers with sophistication $k = 0$, which are referred to as “naïve” consumers: they naïvely expect the monopolist to engage in regular monopoly pricing in channel A, i.e. $\mathbb{E}_0(p_A) = p_M$.¹³ Thereby, they ignore the fact that the very choice of channel A might be signaling a high valuation to the seller. For channel D, we have already assumed that naïve consumers foresee perfect price discrimination in channel D as it does not require iterative thinking about other consumers.

Lemma 3 (Solution with Level-0). *For any non-prohibitively high anonymization cost $s > 0$ and production cost $c \geq 0$, and with strategically “naïve” consumers ($k = 0$), there is a unique solution with the following characteristics:*

- *Consumers form the 0-beliefs $\mathbb{E}_0(p_D) = p_i^*(v)$ and $\mathbb{E}_0(p_A) = p_M = \frac{1+c}{2}$.*
- *Consumers anonymize if, and only if, $v > \hat{v}_0 = \mathbb{E}_0(p_A) + s = p_M + s$, separating into the sets $\mathcal{C}_D = [0, \hat{v}_0]$ and $\mathcal{C}_A = (\hat{v}_0, 1]$.*
- *The seller forms the 1-beliefs $\mathbb{E}_1(\mathcal{C}_D) = [0, \hat{v}_0]$ and $\mathbb{E}_1(\mathcal{C}_A) = (\hat{v}_0, 1]$.*
- *The seller sets prices $p_i^*(v) = \max\{v, c\}$ and $p_{A_0}^* = \hat{v}_0 = p_M + s$.*
- *All consumers in \mathcal{C}_D with $v \geq c$ buy the product at the price offered.*
- *All consumers in \mathcal{C}_A buy the product at the price offered.*

Lemma 3 shows that naïve consumers in channel A pay a premium of s as compared to their expectations ($p_{A_0}^* - \mathbb{E}_0(p_A) = s$). They do not anticipate that the seller can infer that only consumers with a valuation of at least $p_M + s$ choose the anonymous channel. Given this lower bound on the valuations in \mathcal{C}_A , the seller can ignore that anonymized consumers spent s on top, and extract the lower bound’s full valuation. This divergence between expected price and

¹³Alternative starting point assumptions for naïve consumers are discussed in Section 4.1.

realized price, in turn, informs us about the way in which consumers form their price expectations for higher levels of strategic sophistication, $k > 0$.

If instead of being naïve, all consumers are capable of one iteration of strategic reasoning ($k = 1$), they anticipate that the seller's best response to the 0-belief of naïve consumers is $p_{A_0}^* = p_M + s$. Recall, that consumers with $k > 0$ form the beliefs $\mathbb{E}_i(k_{j \neq i}) = k - 1$ for all other consumers and $\mathbb{E}_i(k_S) = k$ for the seller. Therefore, consumers with $k = 1$ assume that the seller responds optimally to a population of consumers with $k = 0$, and adjust their price expectation. As consumers are atomistic, their own anonymization choice is inconsequential for the seller's best response. Accordingly, they form the 1-belief $\mathbb{E}_1(p_A) = p_{A_0}^* = p_M + s$ leading to $\hat{v}_1 = p_M + 2s$, to which the seller's actual best response is $p_{A_1}^* = p_M + 2s$ (as reasoned above). This, in turn, would be the expected price in channel A by consumers with a level of $k = 2$, thus forming the 2-belief $\mathbb{E}_2(p_A) = p_{A_1}^* = p_M + 2s$, and so forth. More generally, we can write $\mathbb{E}_k(p_A) = p_{A_{k-1}}^*$ for all $k > 0$, which in combination with $\mathbb{E}_0(p_A) = p_M$ leads to:

$$\mathbb{E}_k(p_A) = p_M + ks, \quad (12)$$

$$p_{A_k}^* = p_M + (k + 1)s = \hat{v}_k. \quad (13)$$

At every additional level of strategic sophistication, consumers incorporate the sunk cost once more than before, which induces the seller to raise the price further. Consequently, \hat{v}_k increases and \mathcal{C}_A shrinks in size as k increases, i.e., the more strategically sophisticated the consumer population is, the fewer consumers anonymize. When \hat{v}_k matches or exceeds the highest valuation no consumer does so anymore, channel A remains unused and the anonymous market breaks down completely. We denote the threshold level of strategic sophistication from which onwards this is the case by \bar{k} and define:

$$\bar{k} \equiv \min\{k \in \mathbb{Z}_0^+ | \hat{v}_k \geq 1\}. \quad (14)$$

The inequality in (14) can hold with equality as indifferent consumers opt for channel D, by assumption. Using (13) in (14) and solving for \bar{k} yields:

$$\bar{k} \geq \frac{1-c}{2s} - 1 \Rightarrow \bar{k} = \left\lceil \frac{1-c}{2s} - 1 \right\rceil. \quad (15)$$

This shows that channel A breaks down at a *finite* level of strategic sophistication, in turn implying that unlimited strategic sophistication, while *sufficient*, is *not necessary* for a breakdown of channel A.

Lemma 4 (Usage of Channel A). *For any non-prohibitively high anonymization cost $s > 0$ and production cost $c \geq 0$, the anonymous channel is used if, and only if, consumers are not too strategically sophisticated, i.e. if $k < \bar{k} = \left\lceil \frac{1-c}{2s} - 1 \right\rceil$.*

That channel A breaks down at a finite level of sophistication \bar{k} has consequences for the belief formation of consumers when $k > \bar{k}$. While for $k \leq \bar{k}$ belief formation according to (12) does not violate that all players restrict price expectations to $p \in [0, 1]$, this is not the case for $k > \bar{k}$. Denoting any level of consumer sophistication $k > \bar{k}$ by \bar{k}^+ , we specify beliefs $\mathbb{E}_{\bar{k}^+}(p_A)$ to meet this condition in (16). In line with Lemma 4, any belief $\mathbb{E}_{\bar{k}^+}(p_A)$ also has to render the choice of channel D optimal for consumers regardless of their valuation, leading to (17):

$$\mathbb{E}_{\bar{k}^+}(p_A) \in [0, 1] \quad \Rightarrow \quad \mathbb{E}_{\bar{k}^+}(p_A) \leq 1, \quad (16)$$

$$\hat{v}_{\bar{k}^+} = \mathbb{E}_{\bar{k}^+}(p_A) + s \geq 1 \quad \Rightarrow \quad \mathbb{E}_{\bar{k}^+}(p_A) \geq 1 - s. \quad (17)$$

Both conditions are satisfied for any belief $\mathbb{E}_{\bar{k}^+}(p_A) \in [1 - s, 1]$. Hence, multiple beliefs are possible when $k > \bar{k}$, but all imply that channel A remains unused. For any level of consumer sophistication where channel A remains unused (including $k = \bar{k}$), the seller forms $k+1$ -beliefs $\mathbb{E}_{k+1}(\mathcal{C}_D) = [0, 1]$ and $\mathbb{E}_{k+1}(\mathcal{C}_A) = \emptyset$. Therefore, setting p_A is an action on an unreached branch of the game tree and the seller can set any price $p_{A_{\bar{k}^+}}^* \in [0; 1]$ (restricted only by the support of the distribution of v). We summarize the analysis for general level- k in Proposition 1.

Proposition 1 (Solution with Level- k). *For any non-prohibitively high anonymization cost $s > 0$ and production cost $c \geq 0$, it holds that:*

1. *If consumers have a level of strategic sophistication of $k \leq \bar{k} = \lceil \frac{1-c}{2s} - 1 \rceil$, there is a unique solution with the following characteristics:*

- *Consumers form k -beliefs $\mathbb{E}_k(p_D) = p_i^*(v)$ and $\mathbb{E}_k(p_A) = p_M + ks = \frac{1+c}{2} + ks$.*
- *Consumers anonymize if, and only if, $v > \hat{v}_k = p_M + (k+1)s$, separating into the sets $\mathcal{C}_D = [0, \hat{v}_k]$ and $\mathcal{C}_A = (\hat{v}_k, 1]$ (where $\mathcal{C}_A = \emptyset$ if $k = \bar{k}$).*
- *The seller forms $k+1$ -beliefs $\mathbb{E}_{k+1}(\mathcal{C}_D) = [0, \hat{v}_k]$ and $\mathbb{E}_{k+1}(\mathcal{C}_A) = (\hat{v}_k, 1]$.*
- *If $k < \bar{k}$, the seller sets $p_i^*(v) = \max\{v, c\}$ and $p_{A_k}^* = \hat{v}_k = p_M + (k+1)s$.*
- *If $k = \bar{k}$, the seller sets $p_i^*(v) = \max\{v, c\}$ and any $p_{A_{\bar{k}}}^* \in [0, 1]$.*
- *All consumers in \mathcal{C}_D with $v \geq c$ buy the product at the price offered.*
- *All consumers in \mathcal{C}_A buy the product at the price offered.*

2. *If consumers have a level of strategic sophistication of $k > \bar{k} = \lceil \frac{1-c}{2s} - 1 \rceil$, there are multiple solutions with the following characteristics:*

- *Consumers form k -beliefs $\mathbb{E}_{\bar{k}+}(p_D) = p_i^*(v)$ and $\mathbb{E}_{\bar{k}+}(p_A) \in [1-s, 1]$.*
- *No consumer anonymizes as $\hat{v}_{\bar{k}+} \in [1, 1+s]$ and hence $v \leq \hat{v}_{\bar{k}+}$ for all v , leading to the sets $\mathcal{C}_D = [0, 1]$ and $\mathcal{C}_A = \emptyset$.*
- *The seller forms the $k+1$ -beliefs $\mathbb{E}_{k+1}(\mathcal{C}_D) = [0, 1]$ and $\mathbb{E}_{k+1}(\mathcal{C}_A) = \emptyset$.*
- *The seller sets $p_i^*(v) = \max\{v, c\}$ and any $p_{A_k}^* \in [0, 1]$.*
- *All consumers in \mathcal{C}_D with $v \geq c$ buy the product at the price offered.*
- *No consumer buys the product via channel A.*

In Proposition 1.1, consumers with high valuations ($v > \hat{v}_k$) choose the anonymous channel A, consumers with low valuations ($v \leq \hat{v}_k$) choose the direct channel D and are perfectly price discriminated against. Consumers with very low valuations ($v < p_M$) choose the direct channel D irrespectively of k and s as

they cannot possibly hope to get a uniform price that is affordable to them via channel A. These are the consumers that are not served in monopolistic markets without possibility for perfect price discrimination. The multiplicity of solutions in Proposition 1.2 depends only on the multiplicity of possible beliefs about unreached paths of the game. But all solutions feature the same behavior, where no consumer anonymizes and the seller charges individualized prices $p_i^*(v)$ to all.

3 Welfare

Different levels of consumer sophistication k lead to different anonymization behavior, which has consequences for consumer surplus (CS), profits (π), and total welfare (W). We will first take a look at consumer surplus and profits for both channels separately. Total welfare, for which we employ the customary definition, $W = CS + \pi$, will only be included in our final aggregate analysis. Throughout the entire section, though, Figure 2 might serve as illustration of the effects of increasing k when comparing Figure 2a and Figure 2b, which we will explain in detail below. In the comparative statics analysis the discreteness of k is taken into account by calculating changes as differences rather than derivatives. Additionally, due to the potential non-linearity when increasing k from $\bar{k} - 1$ to \bar{k} , these differences only hold for $k + 1 < \bar{k}$.¹⁴

3.1 Channel D

Consumer Surplus and Profits in Channel D

As the seller engages in perfect price discrimination for consumers in \mathcal{C}_D ,

$$CS_{D_k} = 0, \tag{18}$$

¹⁴Recall that \bar{k} is usually the result of rounding (unless $\frac{1-c}{2} - 1 \in \mathbb{Z}_0^+$) and hence the last change in the composition of \mathcal{C}_A and \mathcal{C}_D is usually of different size than s . When increasing consumer sophistication from $\bar{k} - 1$ to \bar{k} , the increase of \mathcal{C}_D is bounded from above by s as *all remaining* consumers switch to channel D. Introducing separate cases in all difference equations is avoided for legibility, but addressed in the text where necessary.

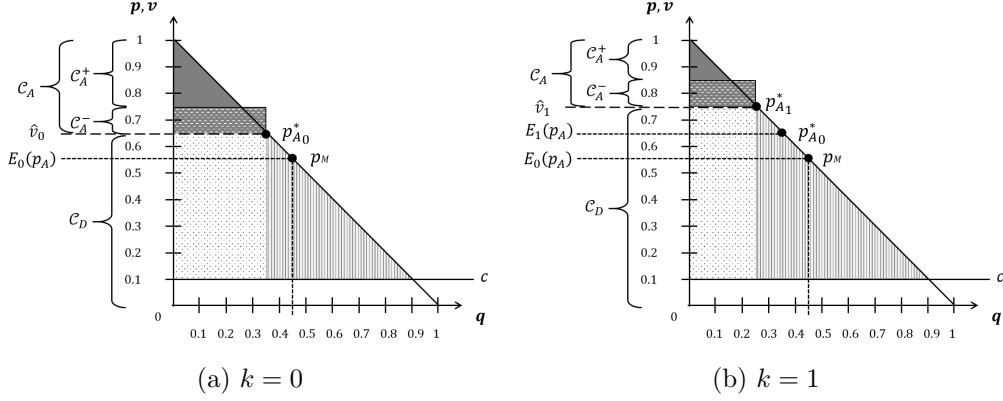


Figure 2: Welfare analysis with parameters $v \sim \mathcal{U}[0, 1]$, $c = 0.1$, $s = 0.1$

whereas the seller appropriates the entire surplus in channel D:

$$\pi_{D_k} = \frac{(\hat{v}_k - c)^2}{2} = \frac{1}{8}(1 - c)^2 + \frac{1 - c}{2}(k + 1)s + \frac{(k + 1)^2}{2}s^2. \quad (19)$$

π_{D_k} corresponds to the vertically striped (lower right) triangle in Figure 2.

Comparative Statics for k in Channel D

Recalling $\mathcal{C}_D = [0, \hat{v}_k]$ and $\hat{v} = p_M + (k + 1)s$, we note first that increasing k raises \hat{v} and $\mathcal{C}_D = [0, \hat{v}_k]$ grows. Let $\Delta CS_{D_k} \equiv CS_{D_{k+1}} - CS_{D_k}$ and $\Delta \pi_{D_k} \equiv \pi_{D_{k+1}} - \pi_{D_k}$ denote the effects of increasing consumer sophistication on consumer surplus and profits in channel D. Using (18) and (19), it follows that, for $k < \bar{k} - 1$:

$$\Delta CS_{D_k} = 0, \quad (20)$$

$$\Delta \pi_{D_k} = (\hat{v}_{k+1} - c)s - \frac{s^2}{2} = \frac{1 - c}{2}s + \frac{2k + 3}{2}s^2. \quad (21)$$

Due to perfect price discrimination, consumer surplus in channel D, unsurprisingly, does not change when consumer sophistication rises. Profits in channel D, though, increase because the set of consumers which the seller can perfectly discriminate, \mathcal{C}_D , grows. This can also be seen by comparing Figure 2a and Figure 2b where the larger bracket along the vertical axis shows the increasing size

of channel D and the larger striped triangle the increase in profits. When consumer sophistication rises from $\bar{k} - 1$ to \bar{k} profits still increase (bounded from above by the expression in (21)) and comes to a halt from there onwards as all consumers are then in \mathcal{C}_D .

Lemma 5 (Effects of Changing Consumer Sophistication (Channel D)). *Raising sophistication of consumers from k to $k + 1$ increases the usage of channel D for all $k < \bar{k}$ (and is maximal for $k \geq \bar{k}$). Consumer surplus in channel D is zero ($CS_{D_k} = 0$) and independent of k ($\Delta CS_{D_k} = 0$). Seller's profits from channel D are positive ($\pi_D > 0$) and increasing in k for all $k < \bar{k}$ (and maximal for $k \geq \bar{k}$).*

3.2 Channel A

Consumer Surplus and Profits in Channel A

In channel A, consumer surplus consists of two parts: the benefit from consumption of the good after the transaction at Stage 3 (denoted by $CS_{A_k}^+$) and the cost of anonymization incurred at Stage 1 (denoted by $CS_{A_k}^-$):

$$CS_{A_k}^+ = \frac{(1 - \hat{v}_k)^2}{2} = \frac{1}{8}(1 - c)^2 - \frac{1 - c}{2}(k + 1)s + \frac{1}{2}(k + 1)^2 s^2, \quad (22)$$

$$CS_{A_k}^- = (1 - \hat{v}_k)s = \frac{1 - c}{2}s - (k + 1)s^2. \quad (23)$$

In Figure 2, $CS_{A_k}^+$ corresponds to the solid grey (upper) triangle, whereas the dashed rectangle that partially overlaps this triangle represents $CS_{A_k}^-$. Net consumer surplus ($CS_{A_k} \equiv CS_{A_k}^+ - CS_{A_k}^-$) in channel A then amounts to:

$$CS_{A_k} = \frac{(1 - \hat{v}_k)^2}{2} - (1 - \hat{v}_k)s = \frac{1}{8}(1 - c)^2 - \frac{1 - c}{2}(k + 2)s + \frac{(k + 1)(k + 3)}{2}s^2. \quad (24)$$

Additionally, note that only some consumers in channel A end up with positive net surplus (those in $\mathcal{C}_A^+ = [\hat{v}_k + s, 1]$), whereas others end up with negative net surplus (those in $\mathcal{C}_A^- = (\hat{v}_k, \hat{v}_k + s)$). The seller's profits in channel A correspond

to the dotted white rectangle in Figure 2 and are given by

$$\pi_{A_k} = (1 - \hat{v}_k)(\hat{v}_k - c) = \frac{1}{4}(1 - c)^2 - (k + 1)^2 s^2. \quad (25)$$

Comparative Statics for k in Channel A

Recalling that $\mathcal{C}_A = (\hat{v}_k, 1]$ and $\hat{v}_k = p_M + (k + 1)s$, we note first that increasing k to $k + 1$ raises \hat{v}_k and hence decreases the size of $\mathcal{C}_A = (\hat{v}_k, 1]$. Let $\Delta CS_{A_k} \equiv CS_{A_{k+1}} - CS_{A_k}$ and $\Delta \pi_{A_k} \equiv \pi_{A_{k+1}} - \pi_{A_k}$ denote the effects of increasing consumer sophistication on consumer surplus and profits in channel A. Using (24) and (25), it follows that, for $k < \bar{k} - 1$:

$$\Delta CS_{A_k} = - \left((1 - \hat{v}_{k+1})s + \frac{s^2}{2} \right) + s^2 = -\frac{1-c}{2}s + \frac{2k+5}{2}s^2, \quad (26)$$

$$\Delta \pi_{A_k} = (1 - \hat{v}_{k+1})s - (\hat{v}_k - c)s = -(2k+3)s^2. \quad (27)$$

While the first term in (26) stems from the reduction of consumer surplus at Stage 3, the second term captures the gain from fewer consumers incurring the anonymization cost. When moving from Figure 2a to Figure 2b, the first effect is represented by the shrinking dark grey triangle, and the second effect by the shrinking dashed rectangle.¹⁵ Which of these effects dominates determines the net effect from increasing k on consumer surplus in channel A. Denoting the threshold level where consumer surplus stops decreasing by $\bar{k}_{\Delta CS}$, we define:

$$\bar{k}_{\Delta CS} \equiv \min\{k \in \mathbb{Z}_0^+ \mid \Delta CS_{A_k} \geq 0\}. \quad (28)$$

Using (26) in (28), addressing the discreteness of k as before, and solving yields:

$$\bar{k}_{\Delta CS} \geq \frac{1-c}{2s} - \frac{5}{2} \Rightarrow \bar{k}_{\Delta CS} = \left\lceil \frac{1-c}{2s} - \frac{5}{2} \right\rceil. \quad (29)$$

¹⁵The dark grey triangle shrinks by a trapezoid composed of the rectangle of area $(1 - \hat{v}_{k+1})s$ and the triangle of area $\frac{s^2}{2}$, whereas the dashed rectangle has height s and shrinks in width by s , making for a decrease in area of s^2 .

Recalling $\bar{k} = \lceil \frac{1-c}{2s} - 1 \rceil$ allows to pin down this threshold's relative location:

$$\bar{k} - \bar{k}_{\Delta CS} = \left\lceil \frac{1-c}{2s} - 1 \right\rceil - \left\lceil \frac{1-c}{2s} - \frac{5}{2} \right\rceil = \left\lceil \frac{1-c}{2s} \right\rceil - \left\lceil \frac{1-c}{2s} - \frac{1}{2} \right\rceil + 1 \in \{1, 2\}. \quad (30)$$

This reveals that consumer surplus stops decreasing already one or two levels of sophistication before channel A breaks down. While this seems counterintuitive at first, it is helpful to recall that $\mathcal{C}_A = \mathcal{C}_A^- \cup \mathcal{C}_A^+$ and that \mathcal{C}_A^- is situated below \mathcal{C}_A^+ . Hence, as k increases, \mathcal{C}_A^+ ceases to contain consumers before \mathcal{C}_A^- does, which means that consumer surplus eventually turns negative. Denoting the additional thresholds \bar{k}_{CS} , where consumer surplus turns negative, and $\bar{k}_{\mathcal{C}_A^+}$, where no consumer in channel A gets net surplus from the transaction, we define:

$$\bar{k}_{CS} \equiv \min\{k \in \mathbb{Z}_0^+ \mid CS_{A_k} = 0\}, \quad (31)$$

$$\bar{k}_{\mathcal{C}_A^+} \equiv \min\{k \in \mathbb{Z}_0^+ \mid \mathcal{C}_A^+ = \emptyset\}. \quad (32)$$

Using (24) in (31) and the definition of $\mathcal{C}_A^+ = (\hat{v}_k + s, 1]$ in (32) and solving yields:

$$\bar{k}_{CS} \geq \frac{1-c}{2s} - 3 \Rightarrow \bar{k}_{CS} = \left\lceil \frac{1-c}{2s} - 3 \right\rceil, \quad (33)$$

$$\bar{k}_{\mathcal{C}_A^+} \geq \frac{1-c}{2s} - 2 \Rightarrow \bar{k}_{\mathcal{C}_A^+} = \left\lceil \frac{1-c}{2s} - 2 \right\rceil. \quad (34)$$

Similarly, these thresholds can be put in relation to \bar{k} :

$$\bar{k} - \bar{k}_{CS} = \left\lceil \frac{1-c}{2s} - 1 \right\rceil - \left\lceil \frac{1-c}{2s} - 3 \right\rceil = \left\lceil \frac{1-c}{2s} \right\rceil - \left\lceil \frac{1-c}{2s} \right\rceil + 2 = 2, \quad (35)$$

$$\bar{k} - \bar{k}_{\mathcal{C}_A^+} = \left\lceil \frac{1-c}{2s} - 1 \right\rceil - \left\lceil \frac{1-c}{2s} - 2 \right\rceil = \left\lceil \frac{1-c}{2s} \right\rceil - \left\lceil \frac{1-c}{2s} \right\rceil + 1 = 1. \quad (36)$$

As (35) shows, the combined cost of anonymization incurred by all consumers in \mathcal{C}_A outweighs the combined surplus from purchasing the good at the penultimate level before the breakdown of channel A, while (36) shows that at the last level before the breakdown of channel A there are no consumers left in channel A

making net surplus. Jointly, they form the set derived in (30) for the sophistication level at which consumer surplus stops decreasing. Hence, we can resolve the counterintuitive result that consumer surplus can stop decreasing already at $\bar{k} - 2$ by having shown that this is only possible because consumer surplus is 0, at best, at this point and negative at $\bar{k} - 1$ the latest. Due to the discreteness of k , the minimum can be attained at either level (indicated by the result of (30)). In any case, raising sophistication from $\bar{k} - 1$ to \bar{k} leads to an increase in consumer surplus because channel A remains unused and consumer surplus jumps to 0 as all consumers are being perfectly price discriminated in channel D.

To summarize our discussion of consumer surplus: consumers lose surplus the more strategically sophisticated they become until everyone “gives in” to the seller’s price discrimination practices in the direct channel D.

Profits in channel A, however, generally decrease in consumer sophistication, as (27) shows. Contrary to consumer surplus, there are no thresholds to be determined as profits in channel A decrease until channel A is not used at all.

Lemma 6 (Changing Consumer Sophistication (Channel A)). *Raising the sophistication of consumers from k to $k + 1$ decreases the usage of channel A for all $k < \bar{k}$ (for $k \geq \bar{k}$ it remains unused). Consumer surplus (CS_A) decreases for all $k < \bar{k}_{\Delta CS} = \lceil \frac{1-c}{2s} - \frac{5}{2} \rceil \in \{\bar{k} - 2; \bar{k} - 1\}$ and becomes non-positive at $\bar{k}_{CS} = \lceil \frac{1-c}{2s} - 3 \rceil = \bar{k} - 2$. Additionally, at $\bar{k}_{C_A^+} = \lceil \frac{1-c}{2s} - 2 \rceil = \bar{k} - 1$ all consumers in channel A incur a net loss. The seller’s profits from channel A (π_A) are positive but decreasing in k for all $k < \bar{k}$ (and zero for all $k \geq \bar{k}$).*

3.3 Aggregate Market (Channel A & Channel D)

Consumer Surplus, Profits, and Welfare

Defining $CS_k \equiv CS_{D_k} + CS_{A_k}$, $\pi_k \equiv \pi_{D_k} + \pi_{A_k}$, and $W_k \equiv CS_k + \pi_k$ leads to the following results (combining (18) and (24) in (37), (19) and (25) in (38), and, ultimately, (37) and (38) in (39):

$$CS_k = \frac{(1 - \hat{v}_k)^2}{2} - (1 - \hat{v}_k)s = \frac{1}{8}(1 - c)^2 - \frac{1 - c}{2}(k + 2)s + \frac{k^2 + 4k + 3}{2}s^2, \quad (37)$$

$$\pi_k = \frac{(\hat{v}_k - c)^2}{2} + (1 - \hat{v}_k)(\hat{v}_k - c) = \frac{3}{8}(1 - c)^2 + \frac{1 - c}{2}(k + 1)s - \frac{k^2 + 2k + 1}{2}s^2, \quad (38)$$

$$W_k = \frac{(1 - c)^2}{2} - (1 - \hat{v}_k)s = \frac{1}{2}(1 - c)^2 - \frac{1 - c}{2}s + (k + 1)s^2. \quad (39)$$

Like total consumer surplus and profits, total welfare can be identified graphically in Figure 2. The first term in (39), $\frac{(1-c)^2}{2}$, corresponds to the whole area between the demand curve and the marginal cost curve, while the second term, $(1 - \hat{v}_k)s$, corresponds to the dashed rectangle. Although the market outcome of Stage 3 is efficient, because every consumer with $v \geq c$ buys the product, total welfare is reduced by the cost of consumers' anonymization behavior as long as $\hat{v}_k < 1$ or, equivalently, $k < \bar{k}$. For any $k \geq \bar{k}$, a fully efficient outcome ensues.

Comparative Statics for k for the Aggregate Market

We derive the effects on the aggregated quantities as differences due to the discrete nature of changes in consumer sophistication. For $k < \bar{k} - 1$:

$$\Delta CS_k \equiv CS_{k+1} - CS_k = -(1 - \hat{v}_{k+1})s + \frac{s^2}{2} = -\frac{1 - c}{2}s + \frac{2k + 5}{2}s^2, \quad (40)$$

$$\Delta \pi_k \equiv \pi_{k+1} - \pi_k = (1 - \hat{v}_k)s + \frac{s^2}{2} = \frac{1 - c}{2}s + \frac{2k + 3}{2}s^2, \quad (41)$$

$$\Delta W_k \equiv W_{k+1} - W_k = s^2 = s^2. \quad (42)$$

As consumer surplus in channel D is equal to zero independent of k , the effect of changing k on aggregate consumer surplus is identical to the effect in channel A: it decreases in k until the threshold level $\bar{k}_{\Delta CS}$ is reached. Recognizing the similarity of (40) and (41), we define an additional threshold level of consumer

sophistication where profits stop increasing, \bar{k}_π :

$$\bar{k}_\pi \equiv \min\{k \in \mathbb{Z}_0^+ \mid \Delta\pi_k \leq 0\}. \quad (43)$$

Substituting (41) in (43) and solving for the threshold level yields:

$$\bar{k}_\pi \leq \frac{1-c}{2s} - \frac{3}{2} \Rightarrow \bar{k}_\pi = \left\lceil \frac{1-c}{2s} - \frac{3}{2} \right\rceil. \quad (44)$$

Also, we locate this threshold in relation to \bar{k} :

$$\bar{k} - \bar{k}_\pi = \left\lceil \frac{1-c}{2s} - 1 \right\rceil - \left\lceil \frac{1-c}{2s} - \frac{3}{2} \right\rceil = \left\lceil \frac{1-c}{2s} \right\rceil - \left\lceil \frac{1-c}{2s} - \frac{1}{2} \right\rceil \in \{0, 1\}. \quad (45)$$

As (45) indicates, profits stop increasing either at the last level before the breakdown of channel A or exactly at \bar{k} . Recalling, however, that all comparative statics difference equations, including (41), only apply to $k < \bar{k} - 1$, we have to examine this case closer because $\bar{k}_\pi \in \{\bar{k} - 1, \bar{k}\}$. Recall further that \mathcal{C}_D increases until $k = \bar{k}$ and that the seller appropriates all surplus from consumers in channel D, but only receives a share of the surplus generated in channel A. It follows that profits are still increasing when consumers' sophistication changes from $\bar{k} - 1$ to \bar{k} . Hence, we have to adjust (44) and (45) to (46) and (47), respectively:

$$\bar{k}_\pi = \left\lceil \frac{1-c}{2s} - 1 \right\rceil, \quad (46)$$

$$\bar{k} - \bar{k}_\pi = 0. \quad (47)$$

While increasing k has negative effects on consumer surplus and positive effects on profits, welfare is generally increasing in k as (42) shows (including the change from $\bar{k} - 1$ to \bar{k}). Thresholds cannot even be determined as the change is independent of k . This result is driven by the fact that increasing the level of sophistication leads to fewer anonymized consumers, corresponding to less cost of anonymization being incurred. Independently of k the surplus from transact-

ing the good stays constant at the maximum due to perfect price discrimination in channel D (raising k simply shifts the surplus from consumers to the seller).

We summarize this analysis in Proposition 2 and Propostion 3.

Proposition 2 (Welfare). *For any non-prohibitively high anonymization cost $s > 0$, production cost $c \geq 0$, and any finite level of consumer sophistication k , aggregated consumer surplus (CS_k), profits (π_k), and welfare (W_k) exhibit the following characteristics:*

- $CS_k > 0$ for $k < \bar{k}_{CS}$, $CS_k \leq 0$ for $\bar{k}_{CS} \leq k < \bar{k}$, and $CS_k = 0$ for $k \geq \bar{k}$, where $\bar{k}_{CS} = \lceil \frac{1-c}{2s} - 3 \rceil$ and $\bar{k} - \bar{k}_{CS} = 2$.
- $\pi_k > 0$ for $k < \bar{k}$, and $\pi_k = W_k$ for $k \geq \bar{k}$.
- $W_k > 0$ for $k < \bar{k}$, and $W_k = W^*$ for $k \geq \bar{k}$, where $W^* = \frac{(1-c)^2}{2}$ is the first-best outcome.

Proposition 3 (Effects of Changing Consumer Sophistication). *Raising the level of strategic sophistication of consumers from k to $k + 1$ has the following effects on consumer surplus, profits, and welfare (ceteris paribus):*

- $\Delta CS_k < 0$ for $k < \bar{k}_{\Delta CS}$, $\Delta CS_k \geq 0$ for $\bar{k}_{\Delta CS} \leq k < \bar{k}$, and $\Delta CS_k = 0$ for $k \geq \bar{k}$, where $\bar{k}_{\Delta CS} = \lceil \frac{1-c}{2s} - \frac{5}{2} \rceil$ and $\bar{k} - \bar{k}_{\Delta CS} \in \{1, 2\}$.
- $\Delta \pi_k > 0$ for $k < \bar{k}$, and $\Delta \pi_k = 0$ for $k \geq \bar{k}$.
- $\Delta W_k > 0$ for $k < \bar{k}$, and $\Delta W_k = 0$ for $k \geq \bar{k}$.

Corollary 1 (Individual Surplus). *As long as consumers are not too sophisticated ($k < \bar{k}_{C_A^+} = \lceil \frac{1-c}{2s} - 2 \rceil = \bar{k} - 1$), some consumers who anonymize themselves (those in C_A^+) end up with positive net surplus, whereas others (those in C_A^-) end up with negative net surplus.*

Proposition 3 shows that (except for boundary cases) the strategic sophistication of consumers works to their disadvantage at an aggregated level and can

break down the market for anonymous shopping. By contrast, the seller benefits from this stepwise breakdown, a development that would also be appreciated by a total welfare maximizer. The reason for this preference is, interestingly, not based on allocation: Due to perfect price discrimination in the direct channel, all consumers with a valuation above the marginal cost of production get the product independent of the existence of the anonymous channel. If the big data technologies driving channel D are already in place, it is inefficient to incur the cost of anonymization. The less channel A is used, the smaller this inefficiency.

Corollary 1 zooms into anonymizing consumers. See Figure 2, where three groups of consumers are distinguished: \mathcal{C}_D , \mathcal{C}_A^- , and \mathcal{C}_A^+ . While the first denotes those consumers who chose channel D, \mathcal{C}_A^- contains those consumers in channel A who make a net loss (because they do not anticipate the seller's incentive to increase the price by s fully), whereas those in \mathcal{C}_A^+ end up with a net benefit.

Comparative Statics for s

Here we analyze the effects of changes in the cost of anonymization. This analysis may inform whether policy efforts to reduce the cost of anonymizing techniques are desirable. Before delving into the analysis, we first identify the threshold where anonymization becomes prohibitively costly for channel A to be used at all (the equivalent of \bar{k}) as the upper bound for our analysis. Denoting this threshold by \bar{s} , we define:

$$\bar{s} \equiv \min\{s \in \mathbb{R}_0^+ | \mathcal{C}_A = \emptyset\}. \quad (48)$$

Recalling that $\mathcal{C}_A = (\hat{v}_k; 1]$ and $\hat{v}_k = \frac{1+c}{2} + (k+1)s$, yields:

$$\mathcal{C}_A = \emptyset \Leftrightarrow \hat{v}_k \geq 1 \Leftrightarrow s \geq \frac{1-c}{2} \frac{1}{(k+1)} \Rightarrow \bar{s} = \frac{1-c}{2} \frac{1}{(k+1)}. \quad (49)$$

Since s is a continuous variable, we do not need to take special cases into account and can use derivatives (instead of differences) of (37), (38), and (39). For $s \leq \bar{s}$:

$$\frac{\partial CS_k}{\partial s} = -\frac{1-c}{2}(k+2) + (k+1)(k+3)s = (k+1)s - (k+2)(1-\hat{v}_k), \quad (50)$$

$$\frac{\partial \pi_k}{\partial s} = \frac{1-c}{2}(k+1) - (k+1)^2 s = (k+1)(1-\hat{v}_k), \quad (51)$$

$$\frac{\partial W}{\partial s} = -\frac{1-c}{2} + 2(k+1)s = (k+1)s - (1-\hat{v}_k). \quad (52)$$

Defining thresholds similarly as above and limiting the analysis to $s \leq \bar{s}$ yields:

$$\frac{\partial CS_k}{\partial s} \begin{cases} < 0 & \text{if } s < \frac{1-c}{2} \frac{k+2}{(k+1)(k+3)} \equiv \bar{s}_{CS}, \\ \geq 0 & \text{if } s \geq \frac{1-c}{2} \frac{k+2}{(k+1)(k+3)} \equiv \bar{s}_{CS}, \end{cases} \quad (53)$$

$$\frac{\partial \pi_k}{\partial s} \begin{cases} > 0 & \text{if } s < \frac{1-c}{2} \frac{1}{k+1} \equiv \bar{s}_\pi, \\ \leq 0 & \text{if } s \geq \frac{1-c}{2} \frac{1}{k+1} \equiv \bar{s}_\pi, \end{cases} \quad (54)$$

$$\frac{\partial W_k}{\partial s} \begin{cases} < 0 & \text{if } s < \frac{1-c}{2} \frac{1}{2k+2} \equiv \bar{s}_W, \\ \geq 0 & \text{if } s \geq \frac{1-c}{2} \frac{1}{2k+2} \equiv \bar{s}_W. \end{cases} \quad (55)$$

It can further be shown that

$$\bar{s} = \bar{s}_\pi > \bar{s}_{CS} > \bar{s}_W, \quad (56)$$

which reveals that the seller's profits increase in s until the prohibitive level \bar{s} is reached. With increasing anonymization cost, more consumers choose channel D instead of channel A, such that the seller appropriates their entire valuation. The effects on consumer surplus and total welfare, on the other hand, are ambiguous and depend on the initial level of s . This is due to the changing effects of raising s on the composition of \mathcal{C}_A : At first, \mathcal{C}_A^- increases in size as s increases. But when there are no consumers in \mathcal{C}_A^+ anymore, a further increase will reduce the size of

\mathcal{C}_A^- again. The respective second derivatives provide further insights:

$$\frac{\partial^2 CS_k}{\partial s^2} = (k+1)(k+3) > 0, \quad (57)$$

$$\frac{\partial^2 W_k}{\partial s^2} = 2(k+1) > 0. \quad (58)$$

Both, consumer surplus and total welfare, are convex in s , implying that they reach a local maximum at $s = \bar{s}$. Moreover, on the lower end of the distribution, i.e. for $s \rightarrow 0$, both consumer surplus and total welfare have a supremum (not a maximum because we defined $s > 0$). Relaxing this constraint for the remainder of this section allows us to study the case of costless anonymization, where $s = 0$.¹⁶ Substituting $s = 0$ and $s = \bar{s}$ in (37), (38), and (39) yields consumer surplus, profits, and welfare at either extreme case:

$$CS_k(s=0) = \frac{1}{8}(1-c)^2, \quad CS_k(s=\bar{s}) = 0, \quad (59)$$

$$\pi_k(s=0) = \frac{3}{8}(1-c)^2, \quad \pi_k(s=\bar{s}) = \frac{1}{2}(1-c)^2, \quad (60)$$

$$W_k(s=0) = \frac{1}{2}(1-c)^2, \quad W_k(s=\bar{s}) = \frac{1}{2}(1-c)^2. \quad (61)$$

For $s = 0$, the difference between consumers' expectations and seller's optimal price disappears. Hence, failing to correctly anticipate the seller's reaction to their anonymization decision does not matter anymore. No consumer in channel A incurs a net loss. As the seller optimally sets the price that consumers expect, there is no change in expectations with increasing consumer sophistication. Hence, if anonymization is costless, the seller's advantage from big data technologies is irrelevant for consumers with high valuation. Those with low valuation get the product but all surplus is extracted by the seller.

Lemma 7 (Solution with Level-k and Costless Anonymization). *With costless anonymization, $s = 0$, and for any non-prohibitive production cost $c \geq 0$, there*

¹⁶If we had assumed $s \geq 0$ already in Section 1, we would have had to distinguish among cases for $s > 0$ and $s = 0$ throughout the analysis, sacrificing clarity.

is a unique solution with the following characteristics:

- Consumers form the k -beliefs $\mathbb{E}_k(p_D) = p_i^*(v)$ and $\mathbb{E}_k(p_A) = p_M = \frac{1+c}{2}$.
- Consumers anonymize if, and only if, $v > \hat{v}_k = p_M$, separating into the sets $\mathcal{C}_D = [0, p_M]$ and $\mathcal{C}_A = (p_M, 1]$.
- The seller forms the $k+1$ -beliefs $\mathbb{E}_{k+1}(\mathcal{C}_D) = [0, p_M]$ and $\mathbb{E}_{k+1}(\mathcal{C}_A) = (p_M, 1]$.
- The seller sets $p_i^*(v) = \max\{v, c\}$ and $p_{A_0}^* = p_M$.
- All consumers in \mathcal{C}_D with $v \geq c$ buy the product at the price offered.
- All consumers in \mathcal{C}_A buy the product at the price offered.

Additionally, channel A does not break down for any level of consumer sophistication. An efficient outcome ensues irrespective of k . Proposition 4 summarizes the comparative statics analysis for s . Figure 3 visualizes it.

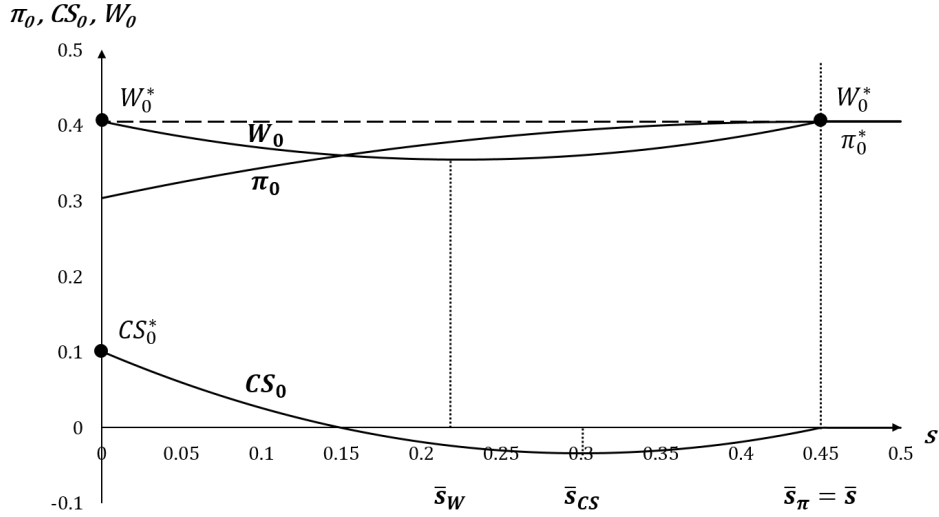


Figure 3: Consumer surplus, profits and welfare as functions of s with parameters $v \sim \mathcal{U}[0, 1]$, $c = 0.1$, $k = 0$

Proposition 4 (Anonymization Cost and Welfare). *For any non-prohibitively high level of consumer sophistication $k < \bar{k}$ and production cost $c \geq 0$, anonymization is prohibitively costly for $s \geq \bar{s} = \frac{1-c}{2} \frac{1}{k+1}$. Then, channel A remains unused. As long as channel A is used, aggregated consumer surplus (CS_k), profits (π_k), and welfare (W_k) exhibit the following characteristics:*

- CS_k is maximal at $s = 0$, decreases in s to its minimum (which is negative) at $s = \bar{s}_{CS}$, then increases in s back to 0 at $s = \bar{s}$.
- π_k is minimal at $s = 0$ and increases in s to its maximum at $s = \bar{s}_{CS} = \bar{s}$.
- W_k is maximal at $s = 0$, decreases in s to its minimum at $s = \bar{s}_W$, then increases in s to another maximum at $s = \bar{s}$. Both maxima lead to the first-best outcome $W_k^* = \frac{(1-c)^2}{2}$.

Proposition 4 shows that higher cost of anonymization is negative for consumers despite the fact that consumer surplus increases in s for relatively high values, which becomes apparent from the fact that consumer surplus is maximal when anonymization is costless. The seller, on the other hand, unambiguously benefits from higher cost of anonymization and prefers prohibitively high cost of $s = \bar{s}$. This maximizes his profits as he can extract all consumer surplus via channel D. A total welfare maximizer, focusing on the welfare-deteriorating role of s , can choose either extreme to prevent consumers from incurring the cost: To achieve an efficient outcome, anonymization should be either costless ($s = 0$) or prohibitively costly ($s = \bar{s}$). Both options are welfare-maximizing, but lead to different surplus allocations. While the seller makes positive profits in either welfare-maximizing scenario, consumers only receive positive surplus if $s = 0$.

4 Alternative Model Specifications

4.1 Beliefs of “naïve” Consumers

In our model we have assumed that “naïve” consumers expect the price in channel A to be equal to the unconditional monopoly price p_M . Some other applications of level- k thinking employ a random distribution as starting point for players with $k = 0$. If the “naïve” consumers in our model were to make their anonymization decision randomly, the seller would correctly infer this and set the price accordingly. Depending on the location of the valuation of anonymized consumers expected by the seller ($\mathbb{E}(v|i \in \mathcal{C}_A)$), three cases can be distinguished:

(a) $\mathbb{E}(v|i \in \mathcal{C}_A) < p_M$,

(b) $\mathbb{E}(v|i \in \mathcal{C}_A) = p_M$,

(c) $\mathbb{E}(v|i \in \mathcal{C}_A) > p_M$.

These cases are equivalent to the cases for the anonymization threshold \hat{v} in Section 2. As discussed there, the seller’s best response in cases (a) and (b) is to charge the unconditional monopoly price p_M . This would require our analysis to include one additional first step of strategic iteration, such that consumers would expect p_M for channel A if they had sophistication $k = 1$. In case (c), however, the seller’s best response is to charge $p_A = \mathbb{E}(v|i \in \mathcal{C}_A) + s$, essentially responding in the same way as before by increasing the price by s above the cutoff valuation. Depending on the exact distance from p_M , this would reduce the number of steps until the complete breakdown of channel A, but not change the underlying mechanism of iterations from there onwards. Hence, while the choice of p_M as a starting point for our analysis pins the analysis to a particular point, it does not crucially affect the model’s analysis.

If anonymization is costless, however, changing the beliefs of naïve consumers has a larger influence. As the iterative process is suspended, expectations do not

change after the initial change from $k = 0$ to $k = 1$, which only affects the seller’s best response in case (a). But the change works in the same fashion as discussed above for $s > 0$. This allows for any price $p_A \in [p_M, 1]$ to be expected by consumers and to be set by the seller, which then becomes the threshold valuation \hat{v} that influences the surplus distribution. Thus, while the resulting solution is not necessarily efficient for $k = 0$ anymore, it is for any $k \geq 1$ and hence does not constitute a crucial departure from our model either.

4.2 Heterogeneous Cost of Anonymization

We have assumed that all consumers find it equally costly to use the anonymous channel A. However, it is easy to imagine that some people might find it less cumbersome to discover and make use of privacy-protecting technologies such as deleting cookies, activating “do not track” browser options, or installing various plugins.¹⁷ Additionally, heterogeneity in s can stem from differing exogenous tastes for privacy in the consumer population, which would reduce the experienced disutility of using channel A. Heterogeneous values of s could be seen as the result of aggregating both effects.

If consumers have heterogeneous anonymization costs s_i , where $s_i \sim \mathcal{U}[\underline{s}, \bar{s}]$ (we redefine \bar{s} for this section), they anonymize for $v > \mathbb{E}_k(p_A) + s_i$. As this expression now depends on two individually heterogeneous variables, there is no uniform cutoff valuation \hat{v} separating the sets \mathcal{C}_D and \mathcal{C}_A . Rather, three segments of consumers’ valuations v need to be distinguished (cf. Figure 4):

- (a) Consumers with $v > \mathbb{E}_k(p_A) + \bar{s}$,
- (b) Consumers with $v \in (\mathbb{E}_k(p_A) + \underline{s}, \mathbb{E}_k(p_A) + \bar{s}]$,
- (c) Consumers with $v \leq \mathbb{E}_k(p_A) + \underline{s}$.

¹⁷For instance, TOR is a “free software and an open network that helps [users] defend against traffic analysis, a form of network surveillance that threatens personal freedom and privacy, confidential business activities and relationships, and state security” (<https://www.torproject.org/>). A more detailed list can be found in Sellenart’s “A paranoid’s toolbox for browsing the web”: <http://pierre.senellart.com/talks/cerre-20160915.pdf>.

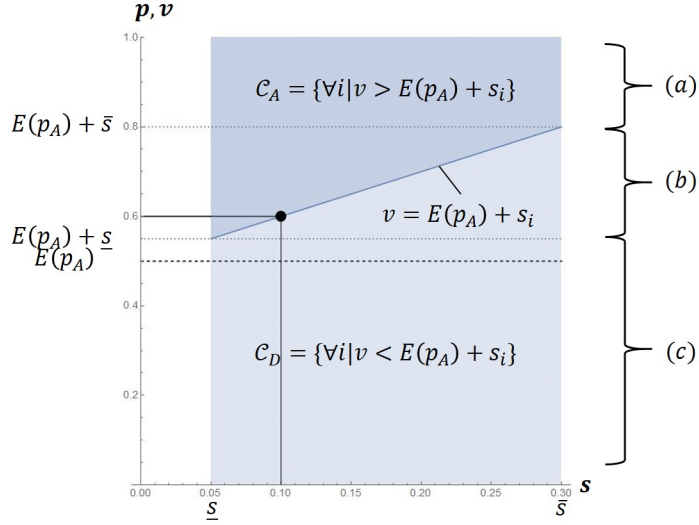


Figure 4: Composition of sets \mathcal{C}_D and \mathcal{C}_A depending on v and s_i with parameters $v \sim \mathcal{U}[0, 1]$, $s_i \sim \mathcal{U}[0.05, 0.3]$, $c = 0$, $k = 0$

Given their price expectation, consumers in segment (c) have a valuation v so low that they choose channel D even for the lowest possible cost of anonymization \underline{s} . Vice versa, consumers in segment (a) have a sufficiently high valuation v such that they choose channel A even if they face the highest possible cost of anonymization \bar{s} . For consumers in segment (b), however, the precise level of their anonymization cost s_i matters for their anonymization choice. Given any valuation $v \in (\mathbb{E}_k(p_A) + \underline{s}, \mathbb{E}_k(p_A) + \bar{s})$ only those whose cost of anonymization s_i is sufficiently low choose channel A, while others with the same valuation v but a higher cost s_i choose channel D. Figure 4 exemplifies this for consumers with a valuation of $v = 0.6$: among these consumers those with anonymization cost $s_i < 0.1$ choose channel A, but those with $s_i \geq 0.1$ choose channel D.

The new composition of the sets \mathcal{C}_D and \mathcal{C}_A implies that demand in both channels is now defined differently for all three segments and hence becomes a piecewise (but still continuous) function of p . As the seller still perfectly price-discriminates in channel D, we focus on the implications of this change for channel A. There, demand is still linear for prices in segment (a) and constant for

prices in segment (c). For price levels in segment (b), however, the uniformly distributed cost of anonymization leads to quadratic demand. We derive the demand function of channel A in general in Appendix A.1. For the specific parameters of Figure 4, this leads to the following demand in channel A:

$$q_A(p_A) = \begin{cases} 0 & \text{if } p_A \geq 1, \\ 1 - p_A & \text{if } 0.8 < p_A < 1, \\ 1 - 0.625 - \frac{(p_A - 0.55)^2}{0.5} & \text{if } 0.55 < p_A \leq 0.8, \\ 1 - 0.625 & \text{if } p_A \leq 0.55. \end{cases} \quad (62)$$

The general derivation of demand in Appendix A.1 confirms that demand is always quadratic for $p \in (\mathbb{E}[p_A] + \underline{s}, \mathbb{E}[p_A] + \bar{s}]$. However, we show in Appendix A.2 that $p_{A_k}^* > \mathbb{E}_k(p_A) + \underline{s}$ as long as $\mathcal{C}_A \neq \emptyset$, i.e. the seller's optimal price is no longer equal to the valuation at the lower bound of \mathcal{C}_A . This in turn implies that there are some consumers in \mathcal{C}_A that do not buy the product.

Continuing the example from above, this result is illustrated in Figure 5. There, the left panel depicts the demand function $q_A(p_A)$ as well as the optimal price in channel A, given by $p_A^* = 0.6629$ with the chosen parameters, whereas the right panel replicates Figure 4 to highlight the mapping from \mathcal{C}_A to $q_A(p_A)$, but replacing the example point with the optimal price p_A^* . Both panels of Figure 5 show that the optimal price p_A^* now exceeds the lowest valuation in \mathcal{C}_A . Now, consumers in the white area between $q_A(p_A^*)$ and $q_A(p_A)$ are not buying the product, despite having a valuation of at least p_M (the unconditional monopoly price). Contrary to our baseline model, the seller is now willing to forgo profits from some consumers because the density of consumers in \mathcal{C}_A across valuations is not uniform in the neighborhood of the lower bound of \mathcal{C}_A anymore.

If we increase consumers' strategic sophistication from $k = 0$ to $k = 1$, consumers form the expectation $\mathbb{E}_1(p_A) = 0.6629$ and make their anonymization choice accordingly. The difficulty of finding an analytical closed-form solution for

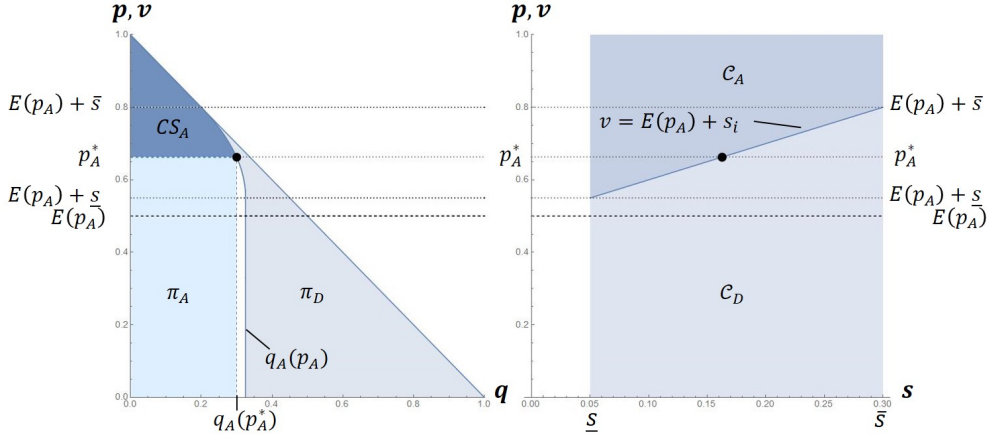


Figure 5: Consumers' anonymization choice as a function of v and s_i with parameters $v \sim \mathcal{U}[0, 1]$, $s_i \sim \mathcal{U}[0.05, 0.3]$, $c = 0$, $k = 0$

p_A^* for any k transmits to finding the cutoff level \bar{k} , from which onwards channel A remains unused. Figure 6 illustrates that profit-maximizing prices $p_A^*(k)$ do not linearly increase in k , unlike in our baseline model (cf. (13)).

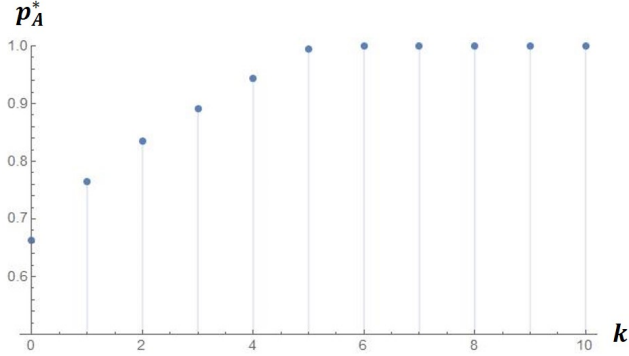


Figure 6: Optimal price in channel A p_A^* for $k = 0, 1, 2, \dots, 10$ with parameters $v \sim \mathcal{U}[0, 1]$, $s_i \sim \mathcal{U}[0.05, 0.3]$, $c = 0$

The qualitative result, that channel A is used less for higher levels of consumers' sophistication and eventually remains unused, however, is replicated: Once the optimal price falls within the interval $(0.95, 1]$, channel A is unused at the next higher k . Notably, this only holds if anonymization is costly for all consumers, i.e. if $\underline{s} > 0$. If $\underline{s} = 0$, we show in Appendix A.3 that $p_{A_k}^* < 1$ for all finite

k ; then a full breakdown of channel A is not achieved for finite k , anymore.¹⁸

Summing up, we conclude that despite some quantitative changes, the general pattern of a gradual breakdown of the anonymous channel and the corresponding effects are not altered qualitatively if anonymization costs are heterogeneous.

4.3 Increasing Competition

Many markets where sellers have access to large amounts of data on consumers' preferences or characteristics, a prerequisite for perfect price discrimination, are dominated by one firm.¹⁹ But to which extent would such a dominant firm, or a monopolist in a market niche, adapt behavior if consumers had access to an (imperfect) substitute product, thereby increasing competition? Assume a rival R offers a product competing with the monopolist's product. A consumer's net value of the rival's product is

$$v^R \equiv \sigma v - p_R, \quad (63)$$

where $\sigma \in [0, 1)$ denotes the degree of substitutability of the products and p_R denotes the price of the rival's product. Alternately, σ proxies the intensity of competition. As any $p_R > 0$ can be reflected by a lower intensity of competition σ , we assume $p_R = 0$ and focus on changes in σ to study the effects of increasing competition. In this scenario, consumers buy from the "monopolist" M if, and only if, $v - p \geq v^R$, i.e. if

$$v - v^R = (1 - \sigma)v \geq p, \quad (64)$$

¹⁸This reinstates unlimited sophistication as a *necessary* condition for a complete breakdown of channel A in the particular case of $\underline{s} = 0$.

¹⁹Prüfer and Schottmüller (2017) explain this development in theoretical terms and cite empirical work to support the statement.

where $p \in \{p_i(v), p_A\}$. Knowing v precisely for all consumers in \mathcal{C}_D , M sets

$$p_i^* = \max\{(1 - \sigma)v, c\} \text{ for all } i \in \mathcal{C}_D, \quad (65)$$

thus guaranteeing that no consumer in channel D buys the rival's product as long as M can earn a profit from this consumer. Consumers anonymize if $\mathbb{E}(u(A)) > \mathbb{E}(u(D))$, i.e. if, and only if,

$$v - \mathbb{E}_k(p_A) - s > v - p_i^* \Leftrightarrow \quad (66)$$

$$v > \frac{\mathbb{E}_k(p_A) + s}{1 - \sigma} \equiv \tilde{v}_k, \quad (67)$$

where \tilde{v}_k denotes the cutoff valuation akin to $\hat{v}_k = \mathbb{E}_k(p_A) + s$ in our main model.

Lemma 8 (Anonymization and Monopolistic Competition). *For any given price expectation of consumers for channel A, $\mathbb{E}_k(p_A)$, the presence of a rival selling a product of substitutability $\sigma \in [0, 1)$ raises the anonymization threshold from \hat{v}_k to $\tilde{v}_k = \frac{\hat{v}_k}{1 - \sigma}$.*

Understanding this increase in the anonymization threshold, M might consider to also increase his price, as in the baseline model, and set $p_{A_k}^* = \tilde{v}_k$. But consumers in channel A might still buy from R. Thus, M faces the same participation constraint in channel A as in channel D: to leave every consumer with at least a net surplus of σv . It follows that pricing at \tilde{v}_k is infeasible. M has to decrease the price to fulfill:

$$(1 - \sigma)\tilde{v}_k \geq p_{A_k}^* \quad (68)$$

which yields, in combination with (67),

$$p_{A_k}^* = (1 - \sigma) \frac{\hat{v}_k}{1 - \sigma} = \hat{v} = \mathbb{E}_k(p_A) + s. \quad (69)$$

Thus, the seller cannot capitalize on the increased anonymization threshold as a

consequence of increased competition. Even though every consumer in channel A now ends up with a positive net surplus from the transaction (depending on the size of s compared to the guaranteed benefit of σv), some consumers are still worse off, having chosen channel A instead of channel D. Consumers with $k = 1$, however, do not adjust their expectation based on forgone surplus but simply update their price expectation to the price that would be M’s best response to consumers with $k = 0$, just as in the baseline model.

Summarizing, consumer surplus increases with competitive pressure. Therefore, there is less anonymization for any given price expectation. Prices in channel D decrease to account for consumers’ improved outside option but the price in channel A is unaffected by competition.

4.4 Temporal Interpretation with “Naïve” Updating

The iterative reasoning structure underlying the belief formation in our model also lends itself to a temporal interpretation. Consider infinitely many repetitions of the presented three-stage game in discrete time where all consumers are of a “naïve” type and form the same initial “naïve” belief about the price in channel A as in our main model specification

$$\mathbb{E}_0(p_{A_0}) = p_M, \tag{70}$$

which they update over time according to the following “naïve” heuristic:

$$\mathbb{E}_t(p_{A_t}) = p_{A_{t-1}}. \tag{71}$$

Similarly, we maintain the assumption from the baseline model, that the seller correctly infers consumers’ price expectations. Depending on the seller’s time horizon and competition, this reduced-form dynamic model generates different insights.

If we assume the seller to be myopic, i.e. to optimize only *within*, but not *across* different periods, we obtain the same result as in the baseline model. We just need to replace k by t , leading to unraveling of the anonymous channel A in a finite number of periods

$$\bar{t} = \lceil \frac{1-c}{2s} - 1 \rceil. \quad (72)$$

Similarly to Corollary 1, then there are some consumers in channel A which end up with a positive net surplus and some with a negative net surplus. However, due to their belief updating, no consumer receives a negative net surplus twice. Should the seller, additionally to being fully myopic, also face an equally myopic rival competing with a product of a degree of substitutability $\sigma \in [0, 1)$, the results of Section 4.3 replicate, again replacing k by t .

The results from Section 4.1 apply in a similar fashion. If consumers start from a higher initial price expectation, they would stop to use channel A earlier, and vice versa. In addition to the analysis of $\mathbb{E}(v|i \in \mathcal{C}_A)$ relative to p_M in Section 4.1, the starting point in the temporal interpretation could additionally be shifted upwards if consumers collectively had a higher level of strategic sophistication. Such a higher level of strategic sophistication might in turn imply a more strategically sophisticated fashion of belief updating, further reducing the number of periods until channel A remains unused.

However, if the seller is not only able to anticipate strategic decisions *within*, but also *across* periods, in the absence of competitive pressure, he would strategically forgo some profits from consumers in channel A in the first period $t = 0$ and set the price $p_{A_0} = 1 - s$. In this way, the seller exploits the “naïve” updating of consumers and reaches full disclosure by the consumer population from $t = 1$ onwards. Because consumers beliefs $\mathbb{E}_1(p_A) = p_{A_0} = 1 - s$ leading to $\hat{v}_1 = \mathbb{E}_1(p_A) + s = 1$, no consumer chooses channel A.²⁰

²⁰The seller, of course, only engage in this different dynamic pricing scheme if the forgone profits from selling less in $t = 0$ are smaller than the difference in profits generated from full

5 Discussion and Conclusion

This paper started from the empirical observation that the technological developments that are behind the “rise of big data” have led to asymmetries on markets for consumer goods (Mayer-Schönberger and Cukier 2013). Sellers making use of datasets on choices of large masses of consumers can tailor prices to individual characteristics more and more and thereby appropriate a lion’s share of the surplus created by market transactions. The specific model of perfect price discrimination used here is just the limit case that will be approximated when prediction methods, mostly using machine learning, improve in power. On top of the sheer amount of information that is available to sellers, consumers are at a second disadvantage. They face cognitive constraints regarding strategic sophistication (Acquisti and Grossklags 2007), while the seller’s data processing capabilities enable him to find best responses to consumers’ behavior.

We have taken these developments seriously and constructed a model to study their implications on prices, consumption choices, and consumers’ incentives to use anonymization technologies protecting their privacy. We have shown that under certain conditions, most notably under the assumption of imperfect strategic sophistication of consumers, a costly privacy-protective sales channel is used even if consumers do not have an explicit taste for privacy. In our model, consumers want to restore their privacy (i.e. choose channel A) solely based on their valuation of the good and their price expectation. We thereby provide a micro-foundation for consumers’ privacy choices, to which the existence of a privacy-protective sales channel can cater.

Our model shows that unlimited strategic sophistication is a sufficient but not a necessary condition for the breakdown of the anonymous sales channel if anonymization is equally costly to all consumers. Allowing for heterogeneity in anonymization cost, sources of which can be different technological savviness

market coverage from $t = 1$ to $t = \bar{t}$ as compared to the profits generated across the otherwise slowly decreasing demand in channel A until $t = \bar{t}$.

but also differing preferences for privacy, can reinstate the necessity of unlimited strategic sophistication for a complete breakdown of the anonymous channel.

In general, the use of big data technologies by sellers improves total welfare by avoiding the deadweight loss usually associated with a monopoly: In contrast to uniform monopoly pricing, consumers with low valuations, $v < p_M$, can purchase the product now. This increases efficiency but not consumer surplus as the seller appropriates the entire surplus from these additional transactions. We have further shown that using the anonymous channel backfires and leads to a net loss for at least some (and under certain conditions all) anonymized consumers (forming the set \mathcal{C}_A^-). We have shown that increasing consumer sophistication leads to a reduction in consumer surplus but to an increase in profits and total welfare. Analyzing different anonymization cost levels, consumer surplus is largest in the extreme case of costless anonymization, $s = 0$, and profits are maximal in the extreme case of anonymization being prohibitively costly, $s = \bar{s}$.

Total welfare, however, is maximal at either extreme, $s = 0$ or $s = \bar{s}$. The two cases differ, however, in the way in which they ensure a first-best result. If $s = 0$, consumers with high valuation anonymize for free and receive positive surplus, whereas all others choose the direct channel, where they get perfectly discriminated against and are left with zero surplus. If $s = \bar{s}$, consumers choose the direct channel irrespective of their valuation for the good. Allocation is efficient because anonymization is too costly and the seller can appropriate all surplus.

Because the distribution of surplus is highly asymmetric, however, policy makers may want to have a second look. A consumer-oriented welfare maximizer should try to eliminate anonymization cost, whereas a seller-oriented welfare maximizer may seek to increase anonymization costs to a prohibitive level. A policy maker with a preference for consumer surplus could, for example, require marketers and online platforms serving as matchmakers between sellers and buyers of consumer goods to set anonymous shopping technologies as *default*. This would then require consumers to *opt in* to non-anonymous shopping instead of

today's standard, where full tracking of consumers' choices is the default and a few providers offer *opt out* technologies. This proposal is also discussed by Acquisti et al. (2016). Those consumers willing to reveal their characteristics to sellers (our model suggests these are those with low valuations) would log in to some service and receive the product for a price related to their WTP. Consumers with higher valuations would stay in the (now default) anonymous channel and pay a higher price, but still retain some surplus.

On the theory side, future research could shed light on the effects of heterogeneity in the level of strategic sophistication amongst consumers, relying on a more elaborate cognitive hierarchy model than this first attempt we undertook here.²¹ This is a complex undertaking, however, because it is not only necessary to specify a distribution of k -levels across the population of consumers (and how it might be related to their WTP). It also requires to specify every consumer's belief about other consumers' level(s) of sophistication and the seller's response to them as a function of that consumer's own sophistication.

To test our theory empirically, we consider it most promising to conduct laboratory experiments where subjects could be assigned specific valuations and a perfectly price-discriminating algorithm could actually be implemented. Subjects could indicate their respective anonymization choices given their valuations and known cost of anonymization. The implied thresholds for anonymization would correspond to a certain level of strategic sophistication according to our model, which in turn could be compared to other measures of strategic sophistication spawned from the level- k literature. Using measures of differences between belief interactions of subjects could inform whether the current model, which neglects more complex cognitive hierarchies, is a fair representation of subjects' behavior or whether efforts to generalize our theory are needed.

²¹In a cognitive hierarchy model as introduced by Camerer et al. (2004) higher-level players would form beliefs about the distribution of all levels of sophistication lower than their own.

References

- Acquisti, Alessandro, Curtis Taylor, and Liad Wagman. 2016. "The Economics of Privacy." *Journal of Economic Literature* 54 (2): 442–92.
- Acquisti, Alessandro, and Jens Grossklags. 2007. "What Can Behavioral Economics Teach Us About Privacy" In *Digital Privacy: Theory, Technologies and Practices*, edited by Alessandro Acquisti, Stefanos Gritzalis, Costas Lambrinouidakis, and Sabrina De Capitani di Vimercati, 363–77. New York & London: Auerbach Publications.
- Acquisti, Alessandro, and Hal R. Varian. 2005. "Conditioning Prices on Purchase History." *Marketing Science* 24 (3): 367–81.
- Argenton, Cédric, and Jens Prüfer. 2012. "Search Engine Competition with Network Externalities." *Journal of Competition Law and Economics* 8 (1): 73–105.
- Armstrong, Mark. 2006. "Recent Developments in the Economics of Price Discrimination." In *Advances in Economics and Econometrics: Theory and Applications: Ninth World Congress*, edited by R. Blundell, W. K. Newey, and T. Persson, 2: 97–141. Cambridge, UK: Cambridge University Press.
- Bergemann, Dirk, and Alessandro Bonatti. 2015. "Selling Cookies." *American Economic Journal: Microeconomics*, 7 (3): 259-94.
- Calzolari, Giacomo, and Alessandro Pavan. 2006. "On the Optimality of Privacy in Sequential Contracting." *Journal of Economic Theory* 130 (1): 168–204.
- Camerer, Colin F., Teck-Hua Ho, and Juin-Kuan Chong. 2004. "A Cognitive Hierarchy Model of Games." *The Quarterly Journal of Economics* 119 (3): 861–98.
- Casadesus-Masanell, Ramon, and Andres Hervas-Drane. 2015. "Competing with Privacy." *Management Science* 61 (1): 229–46.
- Conitzer, Vincent, Curtis R. Taylor, and Liad Wagman. 2012. "Hide and Seek: Costly Consumer Privacy in a Market with Repeat Purchases." *Marketing Science* 31 (2): 277–92.

- Costa-Gomes, Miguel A., and Vincent P. Crawford. 2006. "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study." *The American Economic Review* 96 (5): 1737–68.
- Costa-Gomes, Miguel, Vincent P. Crawford, and Bruno Broseta. 2001. "Cognition and Behavior in Normal-Form Games: An Experimental Study." *Econometrica* 69 (5): 1193–1235.
- Crawford, Vincent P. 2003. "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions." *The American Economic Review* 93 (1): 133–49.
- Crawford, Vincent P., and Nagore Iriberri. 2007a. "Level-k Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?" *Econometrica* 75 (6): 1721–70.
- Crawford, Vincent P., and Nagore Iriberri. 2007b. "Fatal Attraction: Saliency, Naïveté, and Sophistication in Experimental 'Hide-and-Seek' Games." *The American Economic Review* 97 (5): 1731–50.
- Daughety, Andrew F., and Jennifer F. Reinganum. 2010. "Public Goods, Social Pressure, and the Choice between Privacy and Publicity." *American Economic Journal: Microeconomics*, 2 (2): 191–221.
- de Cornière, Alexandre, and Romain de Nijs. 2016. "Online Advertising and Privacy." *The RAND Journal of Economics* 47 (1): 48–72.
- Einav, Liran and Jonathan Levin. 2014. "The Data Revolution and Economic Analysis." *Innovation Policy and the Economy* 14: 1–24.
- Fudenberg, Drew, J. Miguel Villas-Boas, and T. Hendershott. 2006. "Behavior-Based Price Discrimination and Customer Recognition." *Handbook on Economics and Information Systems* 1: 377–436.
- Goldfarb, Avi, and Catherine Tucker. 2012. "Shifts in Privacy Concerns." *The American Economic Review* 102 (3): 349–53.

- Goldfarb, Avi, and Botao Yang. 2009. "Are All Managers Created Equal?" *Journal of Marketing Research* 46 (5): 612–22.
- Hermalin, Benjamin E., and Michael L. Katz. 2006. "Privacy, Property Rights and Efficiency: The Economics of Privacy as Secrecy." *Quantitative Marketing and Economics* 4 (3): 209–39.
- Ho, Teck-Hua, Colin Camerer, and Keith Weigelt. 1998. "Iterated Dominance and Iterated Best Response in Experimental 'P-Beauty Contests.'" *The American Economic Review* 88 (4): 947–69.
- Luo, Xueming, Michelle Andrews, Zheng Fang, and Chee Wei Phang. 2014. "Mobile Targeting." *Management Science* 60 (7): 1738–56.
- Mayer-Schönberger, Viktor, and Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work and Think*. John Murray.
- Mikians, Jakub, László Gyarmati, Vijay Erramilli, and Nikolaos Laoutaris. 2012. "Detecting Price and Search Discrimination on the Internet." In *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*, 79–84. New York, NY, USA: ACM.
- Nagel, Rosemarie. 1995. "Unraveling in Guessing Games: An Experimental Study." *The American Economic Review* 85 (5): 1313–26.
- Norberg, Patricia A., Daniel R. Horne, and David A. Horne. 2007. "The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors." *Journal of Consumer Affairs* 41 (1): 100–126.
- Odlyzko, Andrew. 2003. "Privacy, Economics, and Price Discrimination on the Internet." In *Proceedings of the 5th International Conference on Electronic Commerce*, 355–66. ICEC 2003. New York, NY, USA: ACM.
- Pigou, Arthur C. 1920. *The Economics of Welfare*. London: Macmillan and Co., limited.
- Posner, Richard A. 1978. "An Economic Theory of Privacy." *Regulation* 2: 19–26.

- Posner, Richard A. 1981. "The Economics of Privacy." *The American Economic Review* 71 (2): 405–9.
- Prüfer, Jens, and Christoph Schottmüller. 2017. "Competing with Big Data." CentER Discussion Paper No. 2017-007.
- Shiller, Benjamin Reed. 2013. "First Degree Price Discrimination Using Big Data." Working Paper 58. Working Paper Series. Brandeis University, Department of Economics and International Business School.
- Stahl, Dale O., and Paul W. Wilson. 1994. "Experimental Evidence on Players' Models of Other Players." *Journal of Economic Behavior & Organization* 25 (3): 309–27.
- Stahl, Dale O., and Paul W. Wilson. 1995. "On Players' Models of Other Players: Theory and Experimental Evidence." *Games and Economic Behavior* 10 (1): 218–54.
- Stigler, George J. 1980. "An Introduction to Privacy in Economics and Politics." *The Journal of Legal Studies* 9 (4): 623–44.
- Taylor, Curtis R. 2004. "Consumer Privacy and the Market for Customer Information." *The RAND Journal of Economics* 35 (4): 631–50.
- The New York Times. 2012. "Acxiom, the Quiet Giant of Consumer Database Marketing." *The New York Times*, June 16, sec. Technology.
- Tirole, Jean. 1988. *The Theory of Industrial Organization*. 1st ed. MIT Press.
- Tsai, Janice Y., Serge Egelman, Lorrie Cranor, and Alessandro Acquisti. 2011. "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study." *Information Systems Research* 22 (2): 254–68.
- Tucker, Catherine E. 2014. "Social Networks, Personalized Advertising, and Privacy Controls." *Journal of Marketing Research* 51 (5): 546–62.
- Varian, Hal R. 1997. "Economic Aspects of Personal Privacy." Privacy and Self-Regulation in the Information Age. U.S. Department of Commerce - National Telecommunications and Information Administration.

Villas-Boas, J. Miguel. 2004. "Price Cycles in Markets with Customer Recognition" *RAND Journal of Economics* 35 (3): 486-501.

Wall Street Journal Blog. 2014. "Amazon Wants to Ship Your Package Before You Buy It." *Digits – Tech News & Analysis From the WSJ*. January 17. <http://blogs.wsj.com/digits/2014/01/17/amazon-wants-to-ship-your-package-before-you-buy-it>.

Warren, Samuel D., and Louis D. Brandeis. 1890. "The Right to Privacy." *Harvard Law Review* 4 (5): 193-203.

For Online Publication

A Appendix: Further Results for Heterogenous s

A.1 General Form of Demand in Channel A with $s_i \sim \mathcal{U}[\underline{s}, \bar{s}]$

Assume a monopolistic seller producing at constant marginal cost $c \geq 0$ and consumers valuing the good heterogeneously with v , where $v \sim \mathcal{U}[0, \bar{v}]$. We further assume that consumers face heterogeneous anonymization cost $s_i \sim \mathcal{U}[\underline{s}, \bar{s}]$, where $0 \leq \underline{s} < \bar{s} \leq \bar{v}$, if they choose channel A, and zero cost, if they choose channel D. Consumers choose channel A if and only if $v > \mathbb{E}_k(p_A) + s_i$. As long as $\mathcal{C}_A \neq \emptyset$ (i.e. $\bar{v} > \mathbb{E}_k(p_A) + \underline{s}$), three cases can be distinguished, depending on the relation of the maximum valuation \bar{v} to $\mathbb{E}_k(p_A)$, \underline{s} , and \bar{s} :

- (1) $\bar{v} > \mathbb{E}_k(p_A) + \bar{s}$,
- (2) $\bar{v} = \mathbb{E}_k(p_A) + \bar{s}$,
- (3) $\mathbb{E}_k(p_A) + \bar{s} > \bar{v} > \mathbb{E}_k(p_A) + \underline{s}$.

These cases are depicted in Figure A.1. The area of \mathcal{C}_A , i.e. maximal demand in channel A, for all three cases is given by:

$$q_A^{max} = \frac{(\bar{x} + \bar{y})}{2} \cdot \frac{\bar{z}}{\bar{s} - \underline{s}}, \quad (\text{A.1})$$

with

$$\bar{x} = \max\{\bar{v} - [\mathbb{E}_k(p_A) + \underline{s}], 0\}, \quad (\text{A.2})$$

$$\bar{y} = \max\{\bar{v} - [\mathbb{E}_k(p_A) + \bar{s}], 0\}, \quad (\text{A.3})$$

$$\bar{z} = \max\{\min\{\bar{v}, [\mathbb{E}_k(p_A) + \bar{s}]\} - [\mathbb{E}_k(p_A) + \underline{s}], 0\}. \quad (\text{A.4})$$

While \bar{x} and \bar{y} measure the range of valuations, \bar{z} measures the share of consumers present in channel A at any given valuation, which requires normalization by

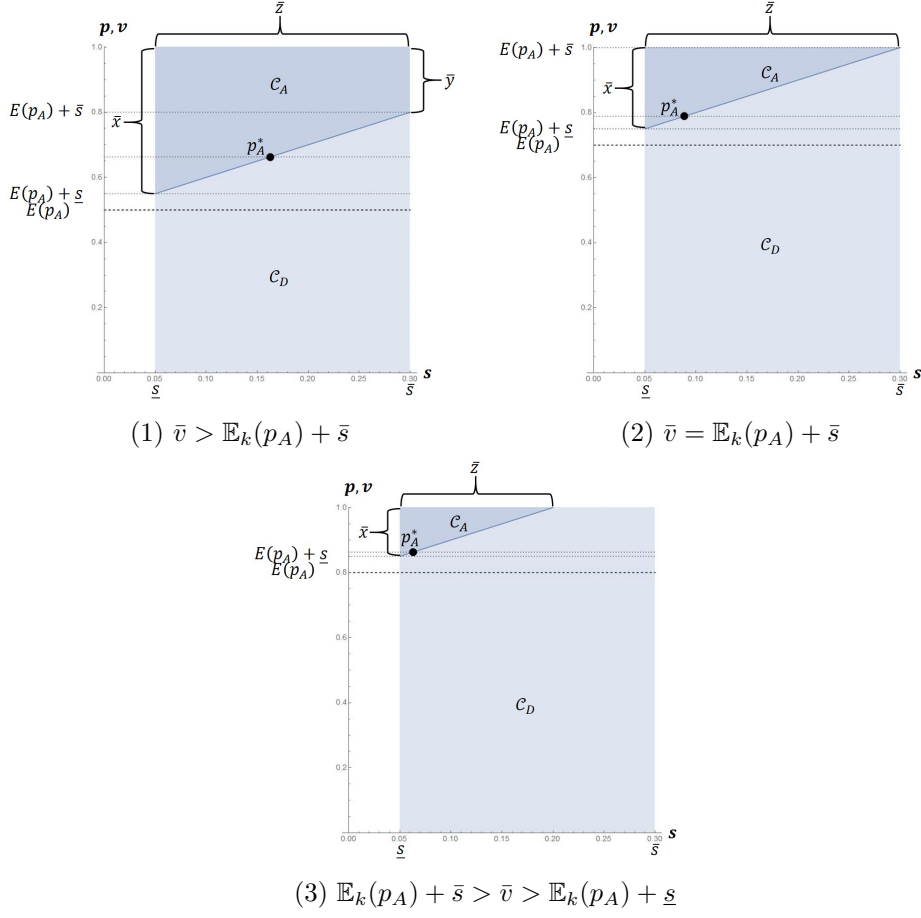


Figure A.1: Composition of sets \mathcal{C}_D and \mathcal{C}_A depending on v and s_i with parameters $v \sim \mathcal{U}[0, 1]$, $s_i \sim \mathcal{U}[0.05, 0.3]$

factor $\frac{1}{\bar{s}-\underline{s}}$. Maximal demand is achieved for any $p_A \leq \mathbb{E}_k(p_A) + \underline{s}$. However, any $p_A > \mathbb{E}_k(p_A) + \underline{s}$ reduces demand in channel A such that q_A^{max} is reduced by q_A^{sub} , where

$$q^{sub}(p_A) = \frac{(x(p_A) + y(p_A))}{2} \cdot \frac{z(p_A)}{\bar{s} - \underline{s}}, \quad (\text{A.5})$$

with

$$x(p_A) = \max\{p_A - [\mathbb{E}_k(p_A) + \underline{s}], 0\}, \quad (\text{A.6})$$

$$y(p_A) = \max\{p_A - [\mathbb{E}_k(p_A) + \bar{s}], 0\}, \quad (\text{A.7})$$

$$z(p_A) = \max\{\min\{p_A - [\mathbb{E}_k(p_A) + \underline{s}], \max\{\min\{\bar{v}, \mathbb{E}_k(p_A) + \bar{s}\} - [\mathbb{E}_k(p_A) + \underline{s}], 0\}\}, 0\}. \quad (\text{A.8})$$

Demand in channel A is then given by subtracting (A.5) from (A.1) and takes the following general form (matched to the segments in Figure 4 in the main manuscript):

$$q(p_A) = q_A^{max} - q^{sub}(p_A) = \begin{cases} 0 & \text{for } p_A \geq \bar{v}, \\ \bar{v} - p_A & \text{for } p_A \text{ in (a),} \\ \bar{v} - [\mathbb{E}_k(p_A) + \frac{\bar{s} + \underline{s}}{2}] - \frac{(p_A - [\mathbb{E}_k(p_A) + \underline{s}])^2}{2(\bar{s} - \underline{s})} & \text{for } p_A \text{ in (b),} \\ \bar{v} - [\mathbb{E}_k(p_A) + \frac{\bar{s} + \underline{s}}{2}] & \text{for } p_A \text{ in (c).} \end{cases} \quad (\text{A.9})$$

This results in the following piecewise demand functions for the three different cases:

Case 1: $\bar{v} > \mathbb{E}_k(p_A) + \bar{s}$

$$q(p_A) = \begin{cases} 0 & \text{if } p_A \geq \bar{v}, \\ \bar{v} - p_A & \text{if } \mathbb{E}_k(p_A) + \bar{s} \leq p_A < \bar{v}, \\ \bar{v} - [\mathbb{E}_k(p_A) + \frac{\bar{s} + \underline{s}}{2}] - \frac{(p_A - [\mathbb{E}_k(p_A) + \underline{s}])^2}{2(\bar{s} - \underline{s})} & \text{if } \mathbb{E}_k(p_A) + \underline{s} < p_A < \mathbb{E}_k(p_A) + \bar{s}, \\ \bar{v} - [\mathbb{E}_k(p_A) + \frac{\bar{s} + \underline{s}}{2}] & \text{if } p_A \leq \mathbb{E}_k(p_A) + \underline{s}. \end{cases} \quad (\text{A.10})$$

Case 2: $\bar{v} = \mathbb{E}_k(p_A) + \bar{s}$

$$q(p_A) = \begin{cases} 0 & \text{if } p_A \geq \bar{v}, \\ \frac{\bar{s}-s}{2} - \frac{(p_A - [\mathbb{E}_k(p_A) + s])^2}{2(\bar{s}-s)} & \text{if } \mathbb{E}_k(p_A) + s < p_A < \mathbb{E}_k(p_A) + \bar{s}, \\ \frac{\bar{s}-s}{2} & \text{if } p_A \leq \mathbb{E}_k(p_A) + s. \end{cases} \quad (\text{A.11})$$

Case 3: $\mathbb{E}_k(p_A) + \bar{s} > \bar{v} > \mathbb{E}_k(p_A) + s$

$$q(p_A) = \begin{cases} 0 & \text{if } p_A \geq \bar{v}, \\ \frac{(\bar{v} - [\mathbb{E}_k(p_A) + s])^2}{2(\bar{s}-s)} - \frac{(p_A - [\mathbb{E}_k(p_A) + s])^2}{2(\bar{s}-s)} & \text{if } \mathbb{E}_k(p_A) + s < p_A < \bar{v}, \\ \frac{(\bar{v} - [\mathbb{E}_k(p_A) + s])^2}{2(\bar{s}-s)} & \text{if } p_A \leq \mathbb{E}_k(p_A) + s. \end{cases} \quad (\text{A.12})$$

A.2 Proof that $p_{A_k}^* > \mathbb{E}_k(p_A) + s$

Claim. If consumers' anonymization cost are $s_i \sim \mathcal{U}[\underline{s}, \bar{s}]$ with $0 \leq s < \bar{s} \leq \bar{v}$, then $p_{A_k}^* > \mathbb{E}_k(p_A) + s$ as long as $\mathcal{C}_A \neq \emptyset$. It suffices to show that $\pi_{A_k}(\mathbb{E}_k(p_A) + s + \varepsilon) > \pi_{A_k}(\mathbb{E}_k(p_A) + s)$, for some $\varepsilon > 0$ without determining $p_{A_k}^*$ exactly.

Proof. Recall that demand in segment (b) (just above $\mathbb{E}_k(p_A) + s$) is of the form

$$q(p_A) = q_A^{max} - \frac{(p_A - [\mathbb{E}_k(p_A) + s])^2}{2(\bar{s} - s)} \quad (\text{A.13})$$

in all three possible cases, with $q_A^{max} = q(\mathbb{E}_k(p_A) + s)$, i.e.

$$q_A^{max} = \begin{cases} \bar{v} - [\mathbb{E}_k(p_A) + \frac{\bar{s}+s}{2}] & \text{if } \bar{v} > \mathbb{E}_k(p_A) + \bar{s} \text{ (Case 1),} \\ \frac{\bar{s}-s}{2} & \text{if } \bar{v} = \mathbb{E}_k(p_A) + \bar{s} \text{ (Case 2),} \\ \frac{(\bar{v} - [\mathbb{E}_k(p_A) + s])^2}{2(\bar{s}-s)} & \text{if } \mathbb{E}_k(p_A) + \bar{s} > \bar{v} > \mathbb{E}_k(p_A) + s \text{ (Case 3).} \end{cases} \quad (\text{A.14})$$

Thus, profits in all three cases can be written as

$$\pi_{A_k}(p_A) = \left(q_A^{max} - \frac{(p_A - [\mathbb{E}_k(p_A) + \underline{s}])^2}{2(\bar{s} - \underline{s})} \right) (p_A - c). \quad (\text{A.15})$$

The corresponding first order derivative in p_A is given by

$$\frac{\partial \pi_{A_k}}{\partial p_A} = q_A^{max} - \frac{3p_A^2 - p_A(2c + 4[\mathbb{E}_k(p_A) + \underline{s}]) + [\mathbb{E}_k(p_A) + \underline{s}](2c + [\mathbb{E}_k(p_A) + \underline{s}])}{2(\bar{s} - \underline{s})}. \quad (\text{A.16})$$

Evaluating at $p_A = \mathbb{E}_k(p_A) + \underline{s} + \varepsilon$ with $\varepsilon > 0$ gives

$$\pi'_{A_k}(\mathbb{E}_k(p_A) + \underline{s} + \varepsilon) = q_A^{max} - \varepsilon \cdot \frac{[\mathbb{E}_k(p_A) + \underline{s}] + \frac{3}{2}\varepsilon - c}{\bar{s} - \underline{s}}. \quad (\text{A.17})$$

Approaching $p_A = \mathbb{E}_k(p_A) + \underline{s}$ from above, the limit is given by

$$\lim_{\varepsilon \rightarrow 0} q_A^{max} - \varepsilon \cdot \frac{[\mathbb{E}_k(p_A) + \underline{s}] + \frac{3}{2}\varepsilon - c}{\bar{s} - \underline{s}} = q_A^{max}. \quad (\text{A.18})$$

From $q_A^{max} > 0$ in all three possible cases it follows that

$$\frac{\partial \pi_{A_k}}{\partial p_A} > 0 \text{ for at least some } p_A = \mathbb{E}_k(p_A) + \underline{s} + \varepsilon \text{ with } \varepsilon > 0 \quad (\text{A.19})$$

Hence, $\pi_{A_k}(\mathbb{E}_k(p_A) + \underline{s} + \varepsilon) > \pi_{A_k}(\mathbb{E}_k(p_A) + \underline{s})$ for some $\varepsilon > 0$ and $p_{A_k}^* > \mathbb{E}_k(p_A) + \underline{s}$. \square

A.3 Proof that $p_{A_k}^* < 1$ for all finite k if $\underline{s} = 0$

Claim. If consumers' anonymization cost are $s_i \sim \mathcal{U}[0, \bar{s}]$ with $0 < \bar{s} \leq \bar{v}$, then $p_{A_k}^* < \bar{v}$ for all finite k .

Proof.

Suppose not. Then, there must be a finite $k = \tilde{k}$ at which the seller sets $p_{A_k}^* = \bar{v}$ for the first time (when k increases). Because consumers indifferent between

channel A and channel D are assumed to choose channel D, it follows that $\mathcal{C}_{A_{\tilde{k}}} = \emptyset$. Otherwise, the seller forgoes profits from consumers in $\mathcal{C}_{A_{\tilde{k}}}$ with $v < \bar{v}$ by setting $p_{A_{\tilde{k}}}^* = \bar{v}$.

Recall further that, with level- k thinking, for any $k > 0$ it holds that

$$\mathbb{E}_k(p_A) = p_{A_{k-1}}^*. \quad (\text{A.20})$$

and that consumers choose channel A if, and only if $v > \mathbb{E}(p_A) + s_i$. As $\underline{s} = 0$ by assumption, it follows that the condition that $\mathcal{C}_{A_{\tilde{k}}} = \emptyset$ requires

$$\mathbb{E}_{\tilde{k}}(p_A) + \underline{s} = \mathbb{E}_{\tilde{k}}(p_A) \geq v \text{ for all } v. \quad (\text{A.21})$$

Combining (A.20) and (A.21), gives

$$\mathbb{E}_{\tilde{k}}(p_A) = p_{A_{\tilde{k}-1}}^* \geq v \text{ for all } v, \quad (\text{A.22})$$

implying that already at $\tilde{k} - 1$ the seller sets $p_{A_{\tilde{k}-1}}^* = \bar{v}$ (as the seller is restricted to set prices within the support of the demand), leading to a contradiction with the assumption that the seller sets $p_{A_{\tilde{k}}}^* = \bar{v}$ at $k = \tilde{k}$ for the first time (when k increases).

By transitivity, this further implies that also $p_{A_0}^* = \bar{v}$ which is only possible if $p_M = \frac{\bar{v}+c}{2} = \bar{v}$, which itself is ruled out by the assumption underlying the model that c is not prohibitively costly.

Therefore, there can be no finite $k = \tilde{k}$ at which the seller sets $p_{A_{\tilde{k}}}^* = \bar{v}$.

Thus, it has to hold that $p_{A_k}^* < 1$ for all finite k if $\underline{s} = 0$. □