Buying consumer data

Andreas Bech

USC

April 12, 2025

Abstract. Abstract

Keywords. Keywords.

Contact. Emails. Thanks to.....

1. INTRODUCTION

Markets for information are an increasingly important aspect of economic activity. The ability for market participants to collect and analyze data can have large implications for the profitability of firms and for consumer welfare. Data, in this context, is information that is useful in inferring the privately known preferences of consumers. Consumers' types can be private information but might also be unknown to consumers themselves, for example, as consumers might face some uncertainty about their valuation for a product. For the firm, knowledge of consumer types can be used for price discrimination and product recommendations.

The disclosure of data brings up concerns about privacy. In economics, there isn't a single, agreed-upon way to define privacy. One possibility is to view privacy as the discomfort consumers feel when their personal information is revealed. A more economic approach, however, would look at how the release of private data impacts interactions in the market—for instance, by enabling third-degree price discrimination. A key problem here is the ripple effect caused by individuals sharing their personal details. In a data market that feeds into a product market, one person's choice to disclose can influence what a firm infers about other, supposedly anonymous, consumers. As this effect is not internalized by consumers, counterintuitive outcomes can take place, even from well-intentioned regulation to protect consumers.

The issue is highly relevant to policy and regulatory frameworks. Notably, in 2016, the EU passed the General Data Protection Regulation (GDPR), which effectively gives consumers full rights and control over their personal data. This model gives direct implications for the potential welfare consequences of such regulation. Another example of regulation is the ban on contracting on genetic information in health insurance markets in the US, the Genetic Information Nondiscrimination Act (GINA). While the GDPR allows for contracting on the information, the GINA Act preempts any kind of voluntary disclosure in the health insurance market.

This paper introduces a new model of voluntary disclosure of consumer data. Without a disclosure technology, this model reduces to the classic monopoly problem, and a subset of consumers with gains from trade do not trade in equilibrium. I examine the potential improvements or consequences from a disclosure technology which assumes that consumers have the right and economic claims to their personal data, a situation that would naturally arise under strict consumer privacy regulation in product and data markets. I show that, paradoxically, giving consumers' rights to their own private information does not always improve consumer welfare and can even make consumers worse off in equilibrium.

Below, section 2 discusses related literature. Section 3 introduces the model and equilibrium concept. Section 4 discusses a simple case with a fully revealing type. Section 5 expands on the model and shows how partial disclosure can arise in equilibrium. Finally, section 6 gives a discussion of the results.

2. LITERATURE

The literature on verifiable disclosure begins with Grossman (1981) and Milgrom (1981). Grossman (1981) explores whether a seller with superior information about the quality of their product can credibly disclose this information to potential buyers. In this signaling game, unraveling occurs as the high-quality type chooses to reveal itself, allowing the buyer to update their beliefs about the rest of the population. This induces the second-highest type to disclose, and so on. Similarly, in Milgrom (1981), the most informed party chooses to reveal their information under the assumption that their claims can be verified by the receiver. A similar result is obtained in this model, as the marginal type is induced to disclose; however, in the presence of uncertainty, a partial disclosure result emerges.

Dye (1985) breaks the unravelling by allowing the manager to be uninformed with positive probability. Jovanovic (1982) and Verrecchia (1983) have similar results. More recently, Onuchic and Ramos (2023) explore disclosure in teams and show that equilibria often involve partial disclosure.

In the context of the classical monopoly problem, disclosure has been shown to have similar unravelling results. Sher and Vohra (2015) study a similar model with commitment, whereas this model is without commitment. Glode, Opp, and Zhang (2018), Pram (2021) and Ali et al. (2023) also study a model without commitment, but with "rich" evidence where the consumer is allowed to disclose any subset of the type space that includes their type.

This paper also relates to the literature on online data sharing, privacy and information markets. This strand of literature deals directly with value of data. Acemoglu et al. (2022) show that externalities among consumers, driven by submodularity in the objective of the platform, leads to depressed compensation of consumer data. Bergemann et al. (2022) feature a similar externality but in a general equilibrium model where platform sells data to both consumer and firm. Galperti et al. (2024) uncover a different externality driven by an intermediary "pooling" data records when selling to a firm.

Voluntary disclosure complements other ways in which consumer information is transmitted to the firm. A different approach is through information design where the firm receives a signal from an intermediary which can be used to segment the market. Bergemann, Brooks and Morris (2015) show what payoffs can be achieved by the intermediary in this case. Ichihashi (2020) also uses information design type signals to model the interaction between firm and consumer.

3. MODEL

A population consists of a unit mass of consumers. Each consumer has a type, which we denote by x. Consumers have unit demand for a product sold by a single firm. The distribution of the consumer's value for the product, $v \in [\underline{v}, \bar{v}]$, is a function of the type, x, such that the conditional distribution of v is $F(v|x)$.

There is a disclosure technology available to the firm whereby it can learn the consumer's valuation. That is, the firm can offer to buy the consumer's information for a price, r, by which the firm learns v. The idea is that the consumer can choose to share his consumer characteristics, x, with the firm, but the firm has access to consumer data and statistical models by which it can learn v very accurately. The consumer on the other hand does not have enough information about the product and thus only partially knows the valuation through $F(v|x)$. a consumer of type x faces a binary choice of either sharing information or not, denoted by $a(x) \in \{0, 1\}$. The game takes place in four stages

1. Firm sets price $r \geq 0$ for the data of consumers
2. Each consumer decides whether to share data
3. Firm sets prices $p(x)$ for each type that sells information, and a price p for the subset of non-disclosing consumers.
4. Consumers buy if their valuation exceeds the price they face in product market.

Note that in the first stage the consumer only knows his type x and not v. The solution concept is PBE. Absent the disclosure technology the model reduces to the classic monopoly problem. The firm is assumed not to have commitment in setting prices. If the firm could commit to a policy $p(x)$ (and p for non-disclosure) and a price r, it is technically a mechanism design problem, but a rather trivial one, since it can set p prohibitively high and make all consumers disclose.

3.1. Equilibrium

The expected surplus at stage 1 of a consumer of type x is

$$\int_p^{\bar{v}} (v - p) dF(v|x)$$

when facing a price p in equilibrium. This is the quantity the consumer compares to r, the surplus from selling information.

The firm has to form beliefs about the subset of consumers who don't disclose in equilibrium, that is, the firm believes the types who don't disclose are distributed according to $x \sim G$. Furthermore, strategies have to be optimal given these beliefs. A profile $(a(x), G, r, p(x), p)$ forms a PBE if

1. Consumer optimally chooses to sell data given p and r,

$$a(x) = 1 \text{ if } r \geq \int_p^{\bar{v}} (v - p) dF(v|x)$$

2. Firm sets optimal non-disclosure price, p, given belief G,

$$p = \arg\max_{p'} p'(1 - G(p'))$$

3. Firm sets optimal data price, r, given $a(x)$ and prior F

4. Firm belief, G, is consistent with $a(x)$

In essence the game reduces to the firm optimally choosing r; every decision and belief formation follows from the choice of r.

4. FULLY REVEALING SIGNAL

Suppose v is perfectly correlated with x such that the consumer knows v in the first stage. If the consumer shares his data, the firm can, knowing v, perfectly price discriminate the consumer and the consumer receives no payoff in the product market, but only receives r for selling his data. We are thus only concerned with the price p that the firm sets for non-disclosing types. Consumers disclosing get payoff r while consumers not disclosing receive payoff $v - p$.

I will now show that in the unique perfect Bayesian Equilibrium of this game the monopolist offers $r = 0$ and all consumers disclose their information.

Proposition 1. *All consumers sell data (market unravels) and are worse off with disclosure technology.*

Proof. First, I will prove that partial disclosure cannot take place. Given $r > 0$, suppose there is partial (or no) disclosure with the subset of consumers with $r < v - p$ not disclosing. Clearly, there is a threshold $\hat{v}$ where the consumer is indifferent such that

$$r = \hat{v} - p$$

and all consumers below the threshold choose to disclose. In this case the firm will set a price $p \geq \hat{v}$ as setting a lower price is non-optimal. But now the marginal consumer can profitably deviate to disclosing for payoff r. This process continues until all consumers disclose, and this is what is meant by unraveling.

Partial disclosure with $r = 0$ is not an equilibrium either, as the firm can deviate to offer $r > 0$, however small, and capture the full surplus of the market.

Finally, it remains to check that the proposed equilibrium is a PBE. On path, all consumer disclose at price $r = 0$ and get payoff 0. The firm can choose any belief off path as a measure 0 of consumers do not disclose. Setting belief $P(v = \bar{v}) = 1$, the firm charges price $p = \bar{v}$ and no consumer has a profitable deviation. The firm has no profitable deviation either as raising r will only lower profit. This concludes the proof that the proposed outcome is an equilibrium and is unique. □

When consumer do not face uncertainty about their own type, the market unravels and consumers get zero payoff. The model highlights a paradoxical consequence of giving consumers "property rights" to their data. An option to sell data to the firm makes them strictly worse off than in the absence of the disclosure technology. To see an improvement for consumers, consumers need a way to group together such that they are not perfectly price discriminated against. Thus a form of partial disclosure is needed. Glode, Opp and Zhang (2018), Pram (2021), Ali et al. (2023) allow consumers to disclose that they belong to a closed subset of the type space, thus giving the option of not fully revealing their type.

5. CONSUMERS DO NOT KNOW V

The previous section highlights a potentially extreme outcome with the disclosure technology. The following section maintains some uncertainty for consumer during the first stage, which will give the marginal consumer a strictly positive payoff in expectation.

The consumer knows his binary characteristic, x, but does not know v. That is, consumers face some uncertainty about their valuation for the firms product. Again, if the consumer discloses the firm learns v perfectly and so the consumer gets no surplus in product market if he sells information. A consumer of any type gets a strictly positive expected surplus in the product market (assuming full support of the distribution of v) and therefore it must be that $r > 0$ in equilibrium. Assume also that the firm has marginal cost c.

There are two types, a high and low valued consumer, $x \in \{x_L, x_H\}$. A consumer of type x chooses to sell data when facing non-disclosure price p if

$$r \geq \int_p^1 (v-p)f(v|x)dv$$

Assumption 1. *$F(v|x_H)$ FOSD $F(v|x_L)$*

Assumption 1 insures that types are ordered in the expected surplus they get in the product market when not selling information. Under Assumption 1 the low valued consumer decides to sell data if a high value consumer does, but not vise-versa. The firm has to choose a price, p, in the product market for the subset for non-disclosing types. Since there are only two types, the firm chooses p under full disclosure and p_H if only low value consumers choose to disclose in equilibrium. The firm faces a trade-off where it has to pay a higher price r in order to discriminate both types in the product market.

The high types expect to get more surplus in the product market given p. In the case where

$$\int_p^1 (v-p)f(v|x_H)dv \geq r \geq \int_{p_H}^1 (v-p_H)f(v|x_H)dv$$

the high types would prefer not to sell data facing the price, p, prevailing in the absence of a disclosure technology, but prefer to switch to selling data when facing price p_H. The low types thus exert an externality on the high types by allowing the firm to infer the types that choose not to sell data in equilibrium.

Proposition 2. *Under certain parametrizations the model can have partial disclosure, where low types disclose and high types do not.*

The firm sets r to maximize its profit in equilibrium

$$\Pi = \begin{cases} \int vdF(v) - r & \text{Full disclosure} \\ \lambda(\int vdF(v|x_L) - r) + (1-\lambda)\int_{p_H}^1 p_H dF(v|x_H) & \text{Only low types disclose} \\ \int_p^1 pdF(v) & \text{No disclosure} \end{cases}$$

where λ is the mass of low types. The firm will set r as low as possible to induce a given level of disclosure, leaving only 3 possible values for r.

Inserting values for r:

$$\Pi = \begin{cases} \int vdF(v) - \int_{p_H}^1 (v-p_H)f(v|x_H)dv & \text{Full disclosure} \\ \lambda(\int vdF(v|x_L) - \int_{p_H}^1 (v-p_H)f(v|x_L)dv) + (1-\lambda)\int_{p_H}^1 p_H dF(v|x_H) & \text{Only low types disclose} \\ \int_p^1 pdF(v) & \text{No disclosure} \end{cases}$$

To see this, consider the three possible scenarios that can arise in equilibrium, depending on the price, r, set by the firm. Either no type sells information, all types sell, or only low types choose to sell. It is then for the firm to choose the optimal price of information, r, given the expected profit in each scenario. Notice that the optimal r for the firm depends on the marginal cost, c, and the set of distributions $F(v|x)$, and the distribution of x — these are the parameters of the model.

If no type sells then, for any x, r is such that

$$r \leq \int_{p}^{1} (v - p) f(v|x) dv$$

Clearly, the firm can achieve this scenario by setting $r = 0$.

If every type shares then r is such that, for any x,

$$r \geq \int_{p_H}^{1} (v - p_H) f(v|x) dv$$

as the firm interprets a deviation as a high type in equilibrium. In this case the optimal price for information is $r = \int_{p_H}^{1} (v - p_H) f(v|x_H) dv$ and the firm perfectly discriminates everyone in the product market.

Finally, if only low types share then r is such that

$$\int_{p_H}^{1} (v - p_H) f(v|x_H) dv \geq r \geq \int_{p_H}^{1} (v - p_H) f(v|x_L) dv$$

in which case firm optimally sets $r = \int_{p_H}^{1} (v - p_H) f(v|x_L) dv$. Here the firm can only price discriminate low types and sets the non-disclosure price p_H, since the firm believes that any non-disclosure is from a high type.

Thus, the firm chooses between three values for r. The optimal r depends on F and the firm's cost, c. Offering a higher disclosure price gives the firm more data and a better ability to price discriminate. Partial disclosure arises when the firm finds it optimal only to pay for the low type's data. This in turn depends on the amount of surplus that can be extracted from high types, respectively, low types, and also on the size of r that will induce the high types to sell data. The example below illustrates that partial disclosure arises, in particular, in the uniform example.

5.1. Example

Consider an example where $f(v|x_H) = 2v$ and $f(v|x_L) = 2(1 - v)$, such that v is unconditionally uniform. Assume also zero marginal cost and that there is an equal mass of each type of consumer. Thus, a high type chooses to disclose given r and p if

$$r \geq \int_p^1 (v-p)2vdv$$

In order to find the optimal price of information, r, the firm will compare the three scenarios discussed in the previoius section. In scenario 1, no type shares and the firm does not pay for information. In the absence of disclosure the firm makes a monopoly profit of $1/4$ and consumers earn an aggregate consumer surplus of $1/8$. The expected surplus of the low type is $1/24$ and the expected surplus of the high type is $5/24$.

In scenario 2, both types sell their data and the firm sets the lowest price such that both consumers disclose, $r = \int_{p_H}^1 (v - p_H)2vdv = \frac{2}{3} - \frac{8}{9\sqrt{3}}$. In the product market the firms knows v for every consumer and has profit $1/2$. Overall profit is $\frac{8}{9\sqrt{3}} - \frac{1}{6}$, or approximately 0.35. The average consumer payoff is $r \approx 0.15$, which is higher than the average in scenario 1. Note however that the payoff to high type is lower than the expected payoff in the absence of the disclose technology.

Finally, in scenario 3, only the low types disclose. The firm pays $r = \frac{2}{3} - \frac{10}{9\sqrt{3}}$ to the low types (measure $1/2$). It makes profit $1/6$ of the low types and profit $\frac{1}{3\sqrt{3}}$ from the high types. The overall profit for the firm is identical to the profit in scenario 2. The low types get a payoff of $r \approx 0.025$ whereas the high types face price $p_H = 1/\sqrt{3}$ in the product market and get an expected surplus of about 0.15. The high types are thus indifferent between scenario 2 and 3, whereas the low types prefer scenario 2 where they are lumped together with the high types.

There are two equilibria with the same profit for the firm, scenario 2 and 3. The equilibrium in scenario 2 Pareto dominates and is more efficient as all gains from trade are exhausted. However, even though all types disclose, both consumer types get a positive payoff in equilibrium. This is in contrast to the fully revealing signal where consumers don't have uncertainty about their type. In that case the firms do not have to compensate for disclosure in equilibrium.

5.2. Continuum of types

Consider a continuum of types $x \sim H$ with each type's valuation distributed with density $f(v|x)$. Given a price for disclosure, r, the marginal consumer, $\hat{x}$, satisfies

$$r = \int_{p_{ND}}^1 (v - p_{ND}) f(v|\hat{x}) dv$$

The existence of a marginal type relies on Assumption 1 which orders the types according to their expected payoff in the product market. Types with $x > \hat{x}$ choose not to sell their data and vice versa. Note that the right-hand side is always positive, as long as $p_{ND} < 1$, which breaks the unraveling for $r > 0$. An equilibrium with partial disclosure (interior solution) is thus possible, depending on the parameters of the model.

If r is such that the cutoff is $\hat{x}$ then the distribution of v for the segment that does not disclose is

$$f_{ND}(v) = \frac{1}{1 - H(\hat{x})} \int_{\hat{x}}^{\bar{x}} f(v|x)h(x)dx$$

and p_{ND} solves $\max_p (p - c)(1 - F_{ND}(p))$. That is, given r, p_{ND} and $\hat{x}$ are determined together.

5.2.1. Uniform example

Let $x \sim U[\delta, 1 - \delta]$ and $v \sim U[x - \delta, x + \delta]$. The marginal consumer is characterized by

$$r = \int_{p_{ND}}^{\hat{x}+\delta} (v - p_{ND}) \frac{1}{2\delta} dv$$

6. EXTENSIONS

6.1. Optimality of posted price

7. DISCUSSION

In this paper I have examined the consequences of consumers having control over their personal data. The equilibria arising from consumer ownership of personal data are ambiguous as firms can make inferences about consumer's types both from what they disclose and what they don't, as is a common theme in disclosure models. I show that when consumers are fully informed of their valuations for the product, the ownership of data affects them negatively in equilibrium as the information market unravels. Consumers are not compensated for their disclosure and the option not to disclose is negatively affected by the disclosure of other types in equilibrium. This result is a natural extension of the unravelling results of Grossman (1981) and Milgrom (1981) in the monopoly problem.

I show further than when consumers are faced with uncertainty about their valuations the unravelling result is broken and partial disclosure can arise in equilibrium. Uncertainty for the consumer leaves strictly positive expected payoff for the marginal anonymous consumer which stops the unravelling that results in the full information case. Contrary to the full information case consumer can benefit from being compensated for disclosure of information. This goes especially for low types that have gains from trade but do not buy the product in the market without the disclosure technology.

The results in this paper yield further questions as to the economic implications of consumer privacy laws like GDPR in the EU. The economic benefits for consumers of this kind of privacy protection is unclear and is shown to depend critically on the assumptions made on information available to participants in the market.

References

Viral V Acharya, Peter DeMarzo, and Ilan Kremer. Endogenous Information Flows and the Clustering of Announcements. *American Economic Review*, 101(7):2955–2979, December 2011.

Alessandro Acquisti, Curtis Taylor, and Liad Wagman. The Economics of Privacy. *Journal of Economic Literature*, 54(2):442–492, June 2016.

Alessandro Acquisti and Hal R. Varian. Conditioning Prices on Purchase History. *Marketing Science*, 24(3):367–381, August 2005.

S Nageeb Ali, Greg Lewis, and Shoshana Vasserman. Voluntary Disclosure and Personalized Pricing. *The Review of Economic Studies*, 90(2):538–571, March 2023.

Dirk Bergemann and Alessandro Bonatti. The economics of social data.

Dirk Bergemann and Alessandro Bonatti. Selling Cookies. *American Economic Journal: Microeconomics*, 7(3):259–294, August 2015.

Dirk Bergemann, Alessandro Bonatti, and Alex Smolin. The Design and Price of Information. *American Economic Review*, 108(1):1–48, January 2018.

Dirk Bergemann, Benjamin Brooks, and Stephen Morris. The Limits of Price Discrimination. *American Economic Review*, 105(3):921–957, March 2015.

Dirk Bergemann, Tibor Heumann, and Stephen Morris. Screening with Persuasion, December 2022. arXiv:2212.03360 [econ].

Dirk Bergemann, Tibor Heumann, and Michael C. Wang. A Unified Approach to Second and Third Degree Price Discrimination, January 2024. arXiv:2401.12366 [cs, econ].

Dirk Bergemann and Stephen Morris. Bayes correlated equilibrium and the comparison of information structures in games: Bayes correlated equilibrium. *Theoretical Economics*, 11(2):487–522, May 2016.

Jeremy Bertomeu and Davide Cianciaruso. Verifiable disclosure. *Economic Theory*, 65(4):1011–1044, June 2018.

Alessandro Bonatti and Gonzalo Cisternas. Consumer Scores and Price Discrimination. *The Review of Economic Studies*, 87(2):750–791, March 2020.

David Bounie, Antoine Dubus, and Patrick Waelbroeck. Selling strategic information in digital competitive markets. *The RAND Journal of Economics*, 52(2):283–313, June 2021.

Markus Konrad Brunnermeier, Rohit Lamba, and Carlos Segura-Rodriguez. Inverse Selection. *SSRN Electronic Journal*, 2020.

Vincent Conitzer, Curtis R. Taylor, and Liad Wagman. Hide and Seek: Costly Consumer Privacy in a Market with Repeat Purchases. *Marketing Science*, 31(2):277–292, March 2012.

Kimon Drakopoulos and Ali Makhdoumi. Providing Data Samples for Free. *Management Science*, 69(6):3536–3560, June 2023.

Piotr Dworczak and Giorgio Martini. The Simple Economics of Optimal Persuasion. *journal of political economy*.

Ronald A Dye. Disclosure of Nonproprietary Information. 2024.

Matthew Elliott and Andrea Galeotti. Market Segmentation Through Information. *SSRN Electronic Journal*, 2019.

Michael J. Fishman and Kathleen M. Hagerty. The Optimal Amount of Discretion to Allow in Disclosure. *The Quarterly Journal of Economics*, 105(2):427, May 1990.

Simone Galperti, Aleksandr Levkun, and Jacopo Perego. The Value of Data Records. *Review of Economic Studies*, 91(2):1007–1038, March 2024.

Simone Galperti and Jacopo Perego. Competitive Markets for Personal Data. *SSRN Electronic Journal*, 2022.

Vincent Glode, Christian C. Opp, and Xingtan Zhang. Voluntary disclosure in bilateral transactions. *Journal of Economic Theory*, 175:652–688, May 2018.

Jerry R. Green and Jean-Jacques Laffont. Partially Verifiable Information and Mechanism Design. *The Review of Economic Studies*, 53(3):447, July 1986.

Sergiu Hart, Ilan Kremer, and Motty Perry. Evidence Games: Truth and Commitment. *American Economic Review*, 107(3):690–713, March 2017.

Andreas Haupt and Zoë Hitzig. Contextually Private Mechanisms, April 2024. arXiv:2112.10812 [econ].

Shota Ichihashi. Online Privacy and Information Disclosure by Consumers. *American Economic Review*, 110(2):569–595, February 2020.

Shota Ichihashi and Alex Smolin. Data Provision to an Informed Seller, March 2023. arXiv:2204.08723 [econ].

Hao Li and Xianwen Shi. Discriminatory Information Disclosure. *American Economic Review*, 107(11):3363–3385, November 2017.

Giorgio Martini. Multidimensional Disclosure.

Paul R. Milgrom. Good News and Bad News: Representation Theorems and Applications. *The Bell Journal of Economics*, 12(2):380, 1981.

Stephen Morris and Hyun Song Shin. Social Value of Public Information. *American Economic Review*, 92(5):1521–1534, November 2002.

Kym Pram. Disclosure, welfare and adverse selection. *Journal of Economic Theory*, 197:105327, October 2021.

Anne-Katrin Roesler and Balázs Szentes. Buyer-Optimal Learning and Monopoly Pricing. *American Economic Review*, 107(7):2072–2080, July 2017.

Itai Sher and Rakesh Vohra. Price discrimination through communication: Price discrimination through communication. *Theoretical Economics*, 10(2):597–648, May 2015.

Vasiliki Skreta. Mechanism design for arbitrary type spaces. *Economics Letters*, 91(2):293–299, May 2006.

Vasiliki Skreta. Sequentially Optimal Mechanisms1. *The Review of Economic Studies*, 73(4):1085–1111, October 2006.

Alex Smolin. Disclosure and pricing of attributes. *The RAND Journal of Economics*, 54(4):570–597, December 2023.

Hyun Song Shin. Disclosures and Asset Returns. *Econometrica*, 71(1):105–133, January 2003.

Philipp Strack and Kai Hao Yang. Countering Price Discrimination with Buyer Information. *SSRN Electronic Journal*, 2024.

Curtis R. Taylor. Consumer Privacy and the Market for Customer Information. *The RAND Journal of Economics*, 35(4):631, 2004.

8. APPENDIX

8.1. Distribution conditional of x

Now consider the case where the consumer's type is informative about the distribution of v. Assume that the type is binary and that the willingness to pay is distributed according to $v \sim F(v|x)$. Let $x \in \{x_H, x_L\}$ and $P(x = x_H) = \lambda$. The consumer does not know v in the first stage but learns it in the second stage. The firm knows $F(v|x)$ and $F(v)$ but does not know x. In the first stage the firm offers a price r for disclosing x. In the second stage the firm offers three prices $\{p, p_H, p_L\}$ where p is the price for consumers who do not disclose in equilibrium.

8.1.1. Example

Consider the case where $f(v|x_H) = 2v$ and $f(v|x_L) = 2(1 - v)$. Thus, a high type chooses to disclose if

$$r + \int_{p_H}^{1} (v - p_H)2vdv \geq \int_{p}^{1} (v - p)2vdv$$

Here the left-hand side is the expected payoff from selling information to the firm; r for disclosing in the first stage and then the expected payoff from facing price p_H in the second stage. The right-hand side is the expected payoff in the product market facing the non-disclosure price, p.

Firm offers prices $p_H = 1/\sqrt{3}$ and $p_L = 1/3$. If firm sets $p \in [1/3, 1/\sqrt{3}]$ then low types disclose and it's optimal to set $p = 1/\sqrt{3}$. It's straightforward to check that there is a PBE where $r = 0, p = 1/\sqrt{3}$ (firm believes deviation is from high type) and both types disclose in equilibrium. Thus, you get the same unraveling result as in the fully revealing case where all types disclose at no cost for the firm.

8.2. Consumer knows v but can only share x

In perhaps the most natural case, the consumer knows v in the first stage but can only share x. The firm knows $F(v)$ and $F(v|x)$. In the second stage the firm offers prices $\{p, p(x)\}$ where p is the price for consumers who do not disclose in equilibrium, and $p(x)$ is charged of consumers disclosing their type, x.

Type (x, v) chooses to disclose if

$$r + \max(v - p(x), 0) \geq \max(v - p, 0)$$

The prices $p(x)$ will depend on the updated beliefs of the firm about what v's of type x choose to disclose in equilibrium.

8.2.1. Example

Case with two types and where $f(v|x_H) = 2v$ and $f(v|x_L) = 2(1 - v)$. High type chooses to disclose if

$$r + \max(v - p_H, 0) \geq \max(v - p, 0)$$