Data Purchase and Third Degree Price Discrimination in Duopoly (**Preliminary Version**)

Constantatos Christos [1] Martis Apostolos Ioannis [1]

[1]Department of Economics, University of Macedonia

23rd Conference on Research on Economic Theory and Econometrics
Tinos, 7-11 July 2025

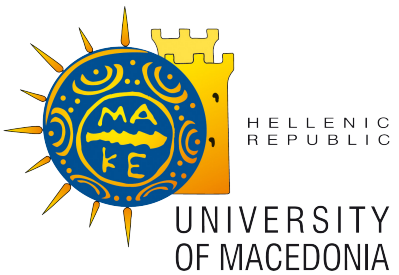


Table of Contents

Table of Contents

Motivation

- Digital markets increasingly rely on consumer data for advertising, personalization, and competitive advantage.
- Mikians et al. (2013) found 10–30% price variation for identical products, suggesting possible discrimination.
- Hannak et al. (2014) showed sector-wide discriminatory pricing using Amazon's Mechanical Turk (AMT) data.
- Woollaston (2014) documented real-world cases: Amazon (2010) and Staples (2012) priced differently by location.
- Understanding firms' incentives for data use is vital for competition policy and regulation.

Research Question

Purpose of the paper

How do firms' strategic data purchasing decisions affect pricing strategies, firm profits, consumer surplus, and total welfare under the existence/ non-existence of the possibility of consumers to hide their data (privacy).

Setting

- We employ a differentiated duopoly model.
- Demand side:
  - Products are imperfect substitutes
  - Consumers have a Bowley-type utility function with the corresponding linear demand for the two products.
  - There exist a continuum of consumer types, uniformly distributed according to their willingness-to-pay for the product
- Production side:
  - Two downstream firms engage in price competition.
  - An independent data broker collects consumer data and sells them to downstream firms at a fixed price.
  - Data access allows a firm to practice third-degree price discrimination based on each consumer's product valuation.

Table of Contents

Literature Review

- Acquisti and Varian (2005): Behavior-based price discrimination, based on past purchases, may enhance consumer welfare vs. uniform pricing.
- Taylor and Wagman (2014): Welfare effects of data usage vary by modeling framework.
- Montes, Sand-Zantman, and Valletti (2019): Shared data maximizes consumer surplus; data acquisition by a single firm (exclusivity) maximizes profits but reduces total welfare. Upstream firm sells data exclusively.
- Bounie, Dubus, and Waelbroeck (2021): Data brokers sell high-WTP consumers' data, typically to one firm.

Table of Contents

Demand Side

- Utility function:

$$U_i = k_i\ (q_{1i} + q_{2i}) - \frac{1}{2}\ (q_{1i}^2 + 2\ \theta q_{1i} q_{2i} + q_{2i}^2)$$

- The parameter $k_i \sim U[\underline{k}, 1]$ is idiosyncratic. It represents the intensity of each individual consumer's willingness-to-pay.

- $\underline{k}(\theta) = \dfrac{2-\theta^2}{6-\theta^2} > 0$ is an exogenous lower bound that ensures positive quantities.

- Individual Demand Function:

$$D_{ji}(p_{ji},\ p_{-ji},\ k_i) = \frac{k_i\ (1-\theta) - p_{ji}(k_i) + \theta\ p_{-ji}(k_i)}{(1-\theta^2)}$$

- Total Demand Function (horizontal summation)

Production Side

- Two symmetric single-product firms selling differentiated products. The marginal production cost is equal to zero.
- A data broker owns a data set that he can sell to one or both firms. The data set allows a firm to distinguish consumer types accurately.
- Three distinct scenarios presented:
  - Bertrand Competition (No data acquisition) - NN
  - Data Exclusivity - DN
  - Data Acquisition - DD

Production - Profits

- Bertrand Competition

$$\Pi_j^N = p_j \int_{k_i \in [\underline{k},1]} D_{ji}(p_j,\ p_{-j},\ k_i)\ f(k_i)\ dk_i$$

- Data Acquisition

$$\Pi_j^D = \int_{k_i \in [\underline{k},1]} p_{ji}\ D_{ji}(p_{ji},\ p_{-ji},\ k_i)\ f(k_i)\ dk_i$$

- Data Exclusivity: Firm j follows Π_j^D with $p_{-ji} = p_{-j}$ and Firm $-j$ follows Π_j^N

Table of Contents

Timing

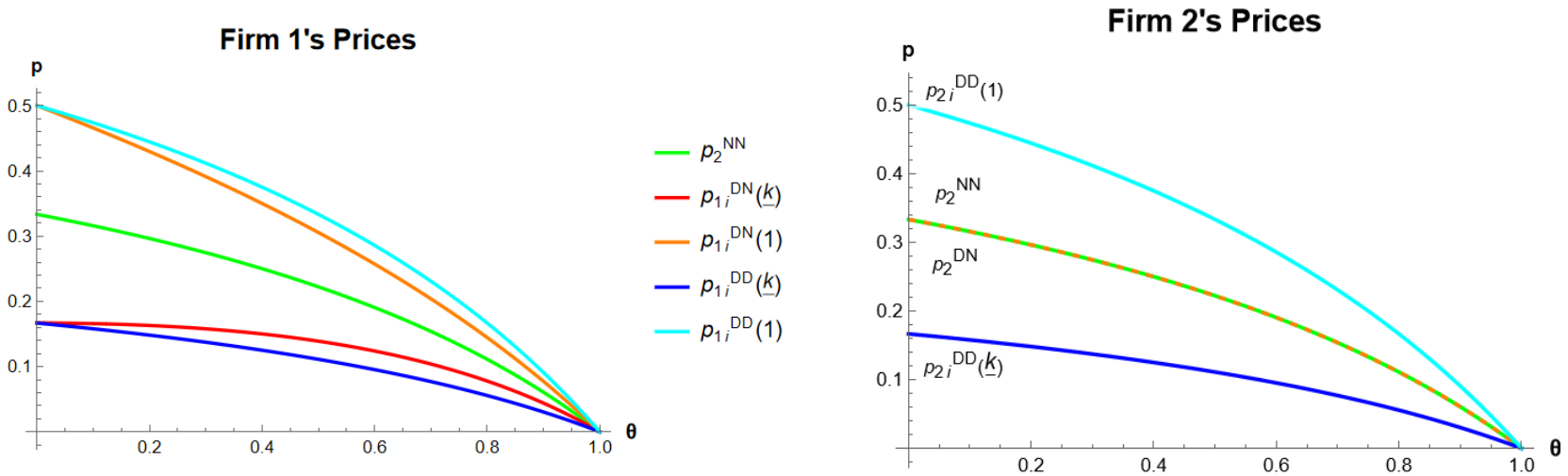
1. The data broker offers access to the database at a fixed cost (to one or both firms).
2. Firms decide whether or not to purchase the data.
3. Firms set their prices.

Bertrand Competition/ Data Exclusivity/ Data Acquisition

	Bertrand Competition	Data Exclusivity		Data Acquisition
Firm j's Profits	$\frac{(1-\theta)(\underline{k}+1)^2}{4(\theta-2)^2(\theta+1)}$	$\frac{(1-\theta)\left((\theta^2-4\theta)(1-\underline{k})^2+16\left(\underline{k}^2+\underline{k}+1\right)\right)}{48(2-\theta)^2(\theta+1)}$	$\frac{(1-\theta)(\underline{k}+1)^2}{4(2-\theta)^2(\theta+1)}$	$\frac{(1-\theta)\left(\underline{k}^2+\underline{k}+1\right)}{3(2-\theta)^2(\theta+1)}$
Aggregate Profits	$\frac{(1-\theta)(\underline{k}+1)^2}{2(\theta-2)^2(\theta+1)}$	$\frac{(1-\theta)\left((\theta^2-4\theta)(1-\underline{k})^2+4\underline{k}(7\underline{k}+10)+28\right)}{48(2-\theta)^2(\theta+1)}$		$\frac{2(1-\theta)\left(\underline{k}^2+\underline{k}+1\right)}{3(2-\theta)^2(\theta+1)}$
Consumer Surplus	$\frac{(\theta-4)\theta(\underline{k}-1)^2+7\underline{k}^2-2\underline{k}+7}{12(\theta-2)^2(\theta+1)}$	$\frac{(3\theta^3-7\theta^2-8\theta)(1-\underline{k})^2+44\underline{k}^2+8\underline{k}+44}{96(2-\theta)^2(\theta+1)}$		$\frac{\underline{k}^2+\underline{k}+1}{3(2-\theta)^2(\theta+1)}$
Welfare	$\frac{(\theta^2-10\theta+13)\underline{k}^2+(-2\theta^2-4\theta+10)\underline{k}+\theta^2-10\theta+13}{12(\theta-2)^2(\theta+1)}$	$\frac{(\theta^3+3\theta^2)(\underline{k}-1)^2-24\theta(\underline{k}(3\underline{k}+2)+3)+4(\underline{k}(25\underline{k}+22)+25)}{96(2-\theta)^2(\theta+1)}$		$\frac{(3-2\theta)\left(\underline{k}^2+\underline{k}+1\right)}{3(2-\theta)^2(\theta+1)}$

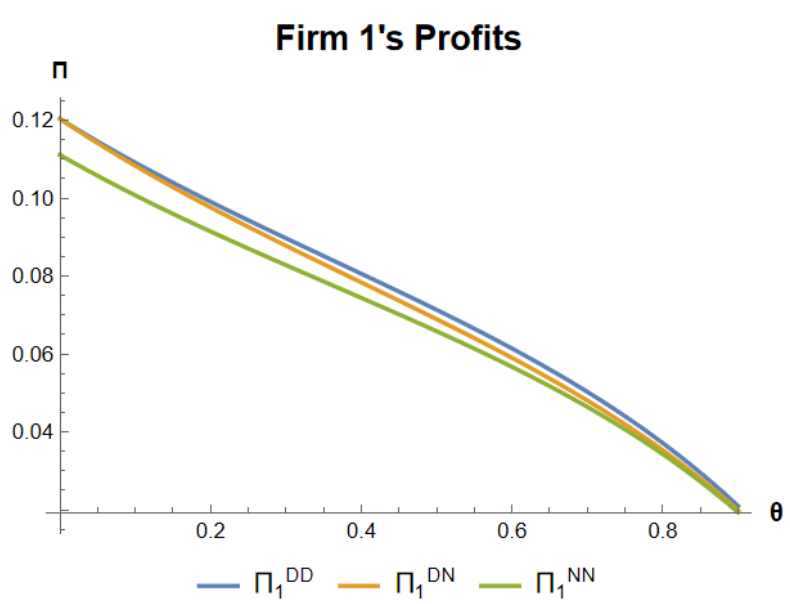
Table: Equilibrium values

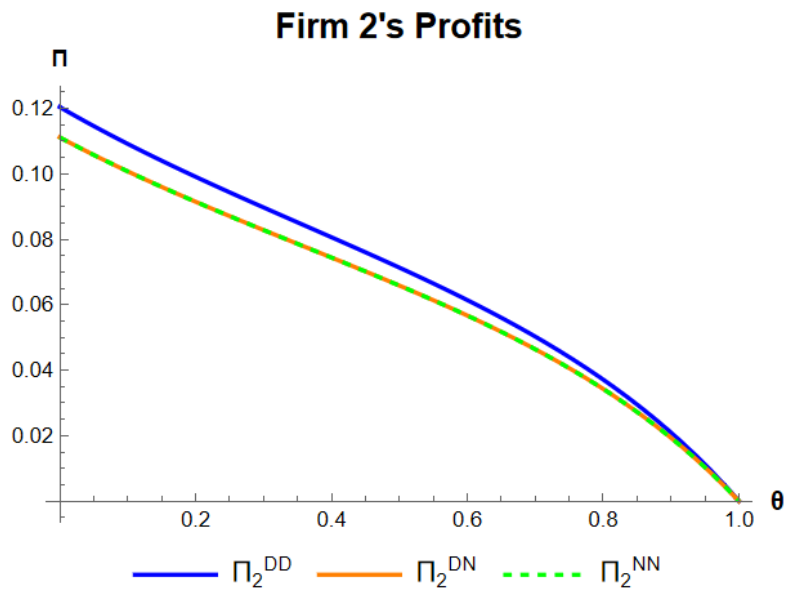
Competition analysis



- A data-holding firm prices more aggressively in low-k, less so in high-k markets.
- When both firms acquire data, competition intensifies or relaxes more significantly within their respective market segments.

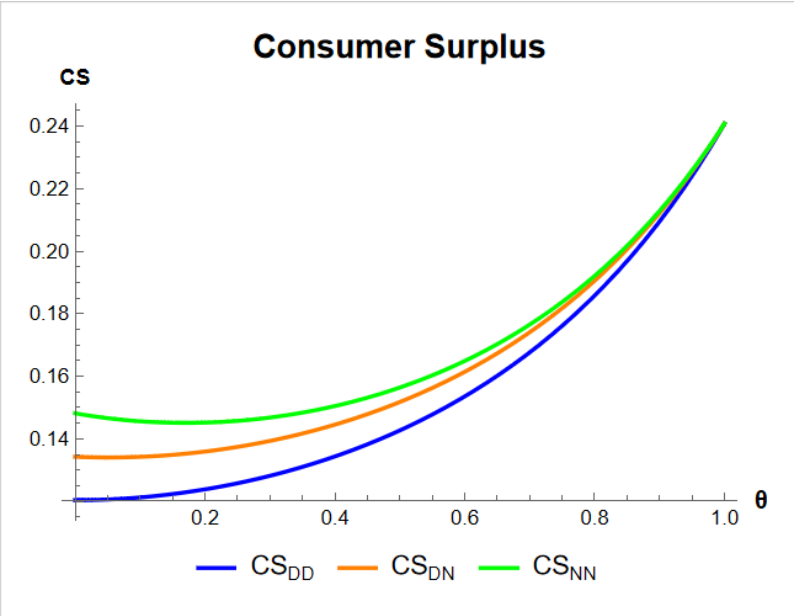
Individual Profit Comparison

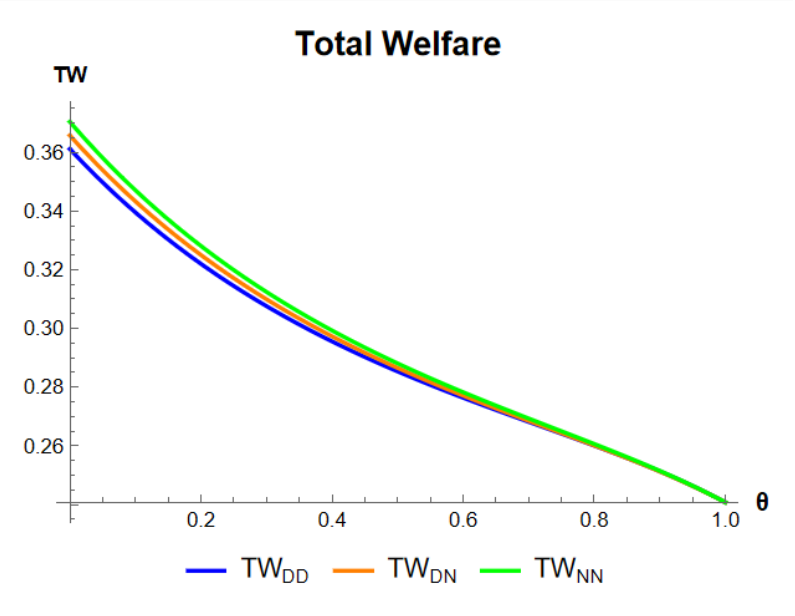




- Data acquisition relaxes competition.
- It intensifies competition in the low-k markets but relaxes it in the high-k ones, which matter more for profits.

Welfare Analysis





- While the average price remains the same, welfare is reduced because the consumers who value the good most, get lower quantity than without data.

Welfare Analysis

For all $\theta \in (0,1)$ and $\frac{2-\theta^2}{6-\theta^2} \leq \underline{k} < 1$.

- **Total Profits:** $T\Pi_{DD} > T\Pi_{DN} > T\Pi_{NN}$
- **Consumer Surplus:** $TCS_{NN} > TCS_{DN} > TCS_{DD}$
- **Total Welfare:** $TW_{NN} > TW_{DN} > TW_{DD}$

Interpretation

Data Acquisition (DD) maximizes firm profits and total welfare, while no-data access (NN) maximizes consumer surplus. Data Exclusivity (DN) is intermediate across all dimensions.

Table of Contents

Privacy Decision I

- Assume that consumers can pay a fixed price, δ, in order to hide their data (both firms obtain consumer data).
- Two options for privacy:
  - Avoiding data collection via social media abstention and persistent consent refusal.
    - The consumer hides her data from all firms in the network.
  - Paying for their data to be erased before a specific transaction.
    - This is possible through applications such as VPN, or by simply using a physical store.
- The action of the type described above imply a different timing decision.

Privacy Decision II

- We add a second consumer group—"new" or "e-unsophisticated" customers.
- They lack online purchase history and cannot be identified through data.
- They share the same preferences and uniform distribution as existing users, but their market density is scaled by $h \geq 1$.

Privacy Decision and Time setting

	Model A	Model B
1st stage	Consumers pay δ to hide their data	Firms send price offers and set the price at their physical store
2nd stage	Firms decide their actual price	Consumers decide whether to accept the offer or purchase at the common price of the physical store

Table: Models A and B (Endogenous Privacy Decision)

- For both models, we assume that $\delta = 0$

Model A

- Consumers hide their data based on anticipated prices.
- Under this model, we distinguish two markets: Personalized and Pool.
- Firm's profits:

$$\Pi_j^A = \underbrace{\int_{\underline{k}}^{k_c} p_{ji}\ D_{ji}(p_{ji}, p_{-ji}, k_i) f(k_i)\, dk_i}_{\text{Personalized Market}} + p_j \underbrace{\left(\int_{k_c}^{1} D_{ji}(p_j, p_{-j}, k_i) f(k_i)\, dk_i + \int_{\underline{k}}^{1} D_{ji}(p_j, p_{-j}, k_i) f(k_i)\, dk_i \right)}_{\text{Pool Market}}$$

Model B I

- Consumers hide their data right before making a market-specific purchase.
- Both firms also supply their product through a physical store.
- Two sub-cases: Symmetric (equal) prices and Asymmetric prices ($p_j > p_{-j}$)
- Two markets under the symmetric model: Personalized and Pool. Firms' profits are the same as in Model A, but now $k_c = k_c(p_1,\ p_2)$.

Model B II

- Three markets under the asymmetric model: Personalized, Mixed, and Pool
  - Personalized Market: Tailored prices
  - Mixed Market: Low-priced firm offers uniform price; high-priced firm offers tailored prices
  - Pool Market: Uniform prices

Model B III

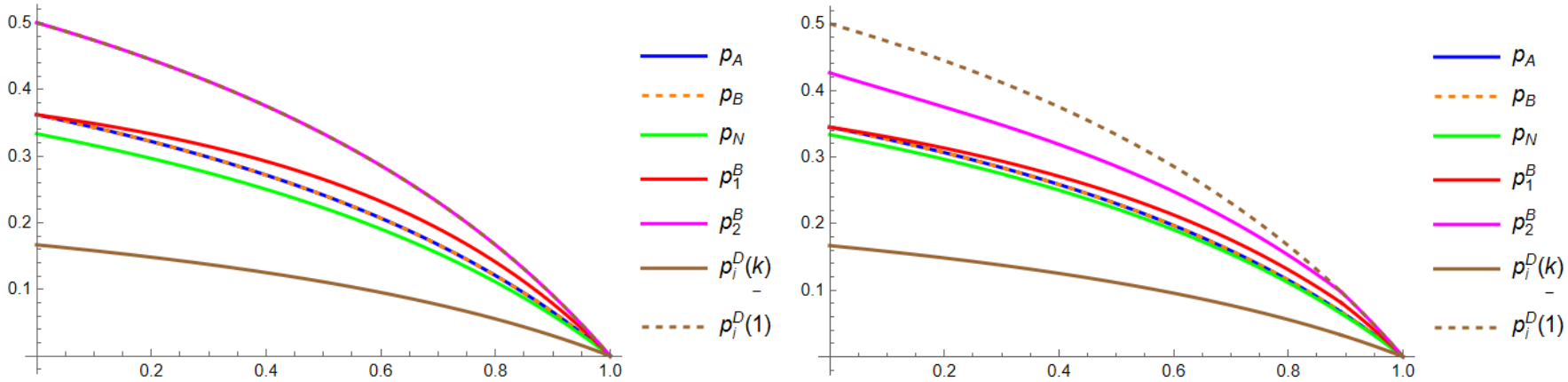
- Firm's profits:

$$\Pi_1^B = \underbrace{\int_{\underline{k}}^{k_{1c}} p_{1i}\, D_{1i}(p_{1i}, p_{2i}, k_i) f(k_i)\, dk_i}_{\text{Personalized Market}} + p_1 \underbrace{\int_{k_{1c}}^{k_{2c}} D_{1i}(p_1, p_{2i}, k_i) f(k_i)\, dk_i}_{\text{Mixed Market}} + p_1 \underbrace{\left(\int_{k_{2c}}^{1} D_{1i}(p_1, p_2, k_i) f(k_i)\, dk_i + \int_{\underline{k}}^{1} D_{1i}(p_1, p_2, k_i) f(k_i)\, dk_i \right)}_{\text{Pool Market}}$$

$$\Pi_2^B = \underbrace{\int_{\underline{k}}^{k_{1c}} p_{2i}\, D_{2i}(p_{1i}, p_{2i}, k_i) f(k_i)\, dk_i}_{\text{Personalized Market}} + \underbrace{\int_{k_{1c}}^{k_{2c}} p_{2i}\, D_{1i}(p_1, p_{2i}, k_i) f(k_i)\, dk_i}_{\text{Mixed Market}} + p_2 \underbrace{\left(\int_{k_{2c}}^{1} D_{2i}(p_1, p_2, k_i) f(k_i)\, dk_i + \int_{\underline{k}}^{1} D_{2i}(p_1, p_2, k_i) f(k_i)\, dk_i \right)}_{\text{Pool Market}}$$

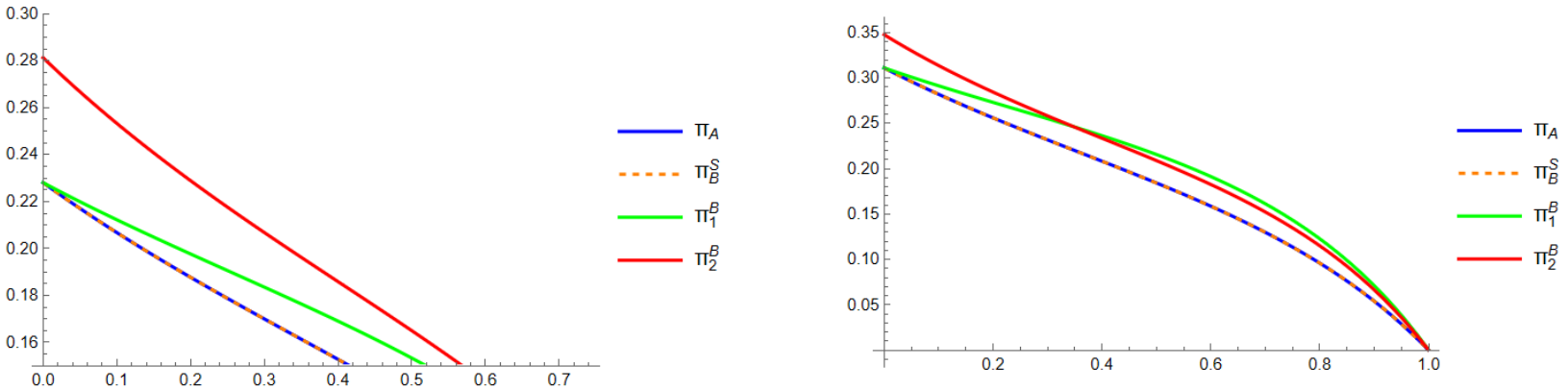
where $k_{jc} = k_{jc}(p_1,\ p_2),\ j = 1,\ 2$ are the critical consumers for firm 1 and firm 2 respectively.

Prices under Privacy Cost



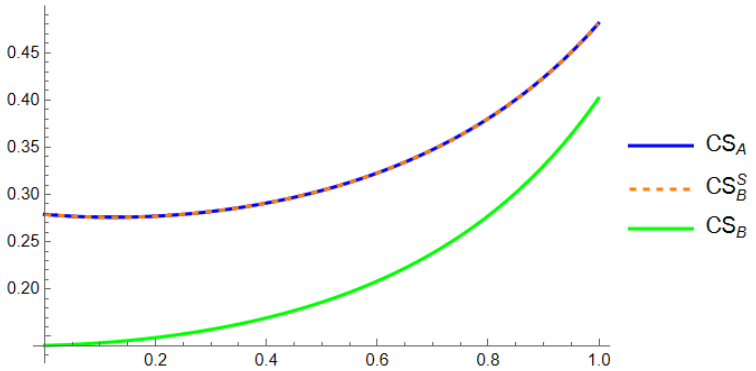
- Ex-ante and ex-post (symmetric prices) privacy decision generate the same price.
- Ex-post privacy decision with asymmetric prices relaxes competition.
- High-pricing firm specialize online for low/high h, but enters the offline market by lowering prices for intermediate h.

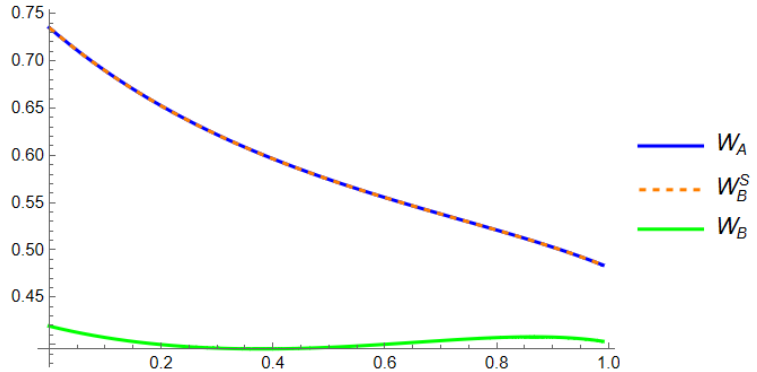
Profits under Privacy Cost – Comparison



- Relative profitability depends on h and θ.
- For $h > 1.53$:
  - At low θ, the high-price strategy wins: differentiation limits losses in the pool market.
  - At high θ, the low-price firm performs better.
- Neither firm prefers symmetric pricing.

Welfare Analysis under Privacy Cost – Comparison

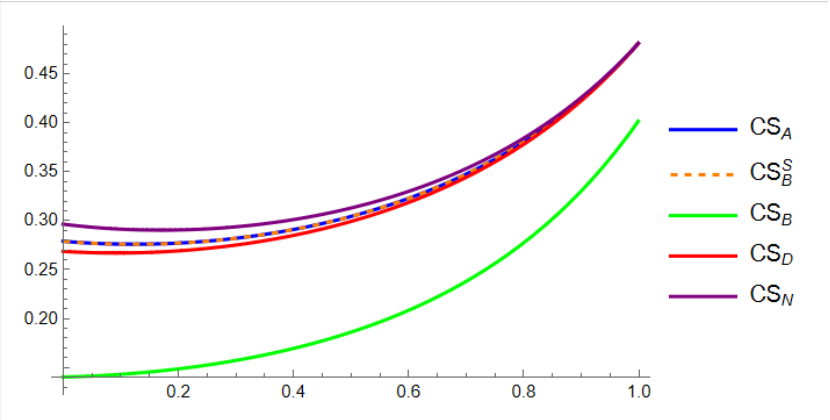


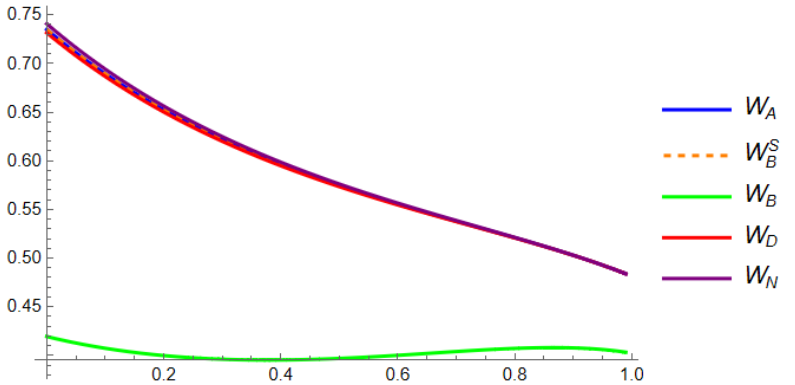


- While ex-post privacy benefits firms, it hurts consumers and society.
- In the asymmetric equilibrium, ex-post privacy allows firms to relax competition by "differentiating" their products.

Table of Contents

Welfare Analysis – Endogenous vs. Exogenous Privacy





- Privacy does not always improve consumer and social welfare.
- In the asymmetric case, old consumers who go private generate a negative externality to new consumers.
- Boosts surplus for high-WTP e-sophisticated consumers, but reduces it more for data-virgins.

Table of Contents

Concluding Remarks I

- Without endogenous privacy decision,
  - Data acquisition relaxes (intensifies) competition for the high-valuation (low-valuation) consumers
  - This leads to higher profits for both firms than under differentiated Bertrand and Data Exclusivity cases.

Concluding Remarks II

- With endogenous privacy:
  - Models A and B (symmetric) yield identical outcomes.
  - Asymmetric Model B softens competition, increasing profits.
  - Except at intermediate new-consumer densities, the high-price firm specializes on online sales, the low-price firm serves offline.
  - Welfare rises under symmetric Model B and Model A.
  - Welfare falls under asymmetric Model B

THANK YOU FOR YOUR ATTENTION

Any questions?

Table of Contents

References I

- Acquisti, Alessandro and Hal R Varian (2005). "Conditioning prices on purchase history". In: *Marketing Science* 24.3, pp. 367–381.
- Bounie, David, Antoine Dubus, and Patrick Waelbroeck (2021). "Selling strategic information in digital competitive markets". In: *The RAND Journal of Economics* 52.2, pp. 283–313.
- Hannak, Aniko et al. (2014). "Measuring price discrimination and steering on e-commerce web sites". In: *Proceedings of the 2014 conference on internet measurement conference*, pp. 305–318.
- Mikians, Jakub et al. (2013). "Crowd-assisted search for price discrimination in e-commerce: First results". In: *Proceedings of the ninth ACM conference on Emerging networking experiments and technologies*, pp. 1–6.

References II

- Montes, Rodrigo, Wilfried Sand-Zantman, and Tommaso Valletti (2019). "The Value of Personal Information in Online Markets with Endogenous Privacy". In: *Management Science* 65.3, pp. 1342–1362. DOI: 10.1287/mnsc.2017.2989. URL: https://pubsonline.informs.org/doi/10.1287/mnsc.2017.2989.
- Taylor, Curtis and Liad Wagman (2014). "Consumer privacy in oligopolistic markets: Winners, losers, and welfare". In: *International Journal of Industrial Organization* 34, pp. 80–84.

References III

Woollaston, Victoria (2014). *Internet shoppers beware: Websites change product prices depending on a customer's purchase history - and even where they live*. URL: https://www.dailymail.co.uk/sciencetech/article-2818613/Internet-shoppers-beware-Websites-change-product-prices-depending-customers-purchase-history-live.html.