Buying Data from Consumers

The Impact of Monitoring Programs in U.S. Auto Insurance

Yizhou Jin — UC Berkeley
Shosh Vasserman — Stanford

NBER SI IT Digitization & IO
July 2019

Motivation

IT + Privacy Standards → Direct transactions of consumer data

- Firms directly incentivize consumers to reveal information **voluntarily**
- **Own** collected data as proprietary

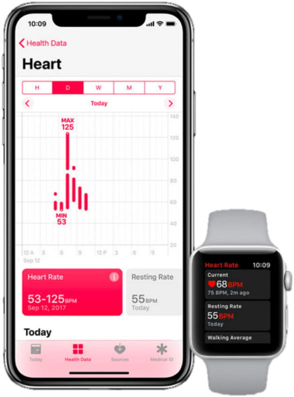
Monitoring in Auto Insurance

A simple device that reveals “how people drive.” more examples

Direct Transactions of Consumer Data in General

back

Prevalent in insurance and lending. Empirical evidence on its economic impact is limited.



Vitality - John Hancock Life Insurance

Services:

Borrowing Investing Rental Others

Ways to Improve Score:

- Receive Income through the app
- Pay Utility Bills through the app
- Connect with friends on the app

Alibaba - Proprietary Credit Scores

Uber Visa Card

Earn $100 after spending $500 on purchases in the first 90 days.

No annual fee + rebates on:

- Dining 4%
- Travels 3%
- Online purchases 2%

Uber – Credit Card

This Project: Research Question and Context

What is the **profit and welfare impact** of introducing a **monitoring program** in U.S. auto insurance?

- Acquire proprietary panel datasets from a major U.S. auto insurer
  - A **monitoring program** is introduced during our research window
  - Matched to competitors' price menus based on regulatory filings

This Project: Empirical Strategy

1. How useful is monitoring?

2. How much information is revealed in equilibrium?

This Project: Empirical Strategy

1. How useful is monitoring?
   - Provide reduced-form evidence and quantify monitoring's ability to both **incentivize safer driving** and **improve risk rating**.

2. How much information is revealed in equilibrium?

This Project: Empirical Strategy

1. How useful is monitoring?
   - Provide reduced-form evidence and quantify monitoring's ability to both incentivize safer driving and improve risk rating.

2. How much information is revealed in equilibrium?
   - **Demand**: estimate structural parameters to capture correlations of **monitoring opt-in choice**, **insurance choices**, **cost** to insure.
   - Supply: firm's information set is endogenous to prices: propose two-period two-product model to characterize pricing in counterfactual equilibria.

This Project: Empirical Strategy

1. How useful is monitoring?
   - Provide reduced-form evidence and quantify monitoring's ability to both incentivize safer driving and improve risk rating.

2. How much information is revealed in equilibrium?
   - Demand: estimate structural parameters to capture correlations of monitoring opt-in choice, insurance choices, cost to insure.
   - **Supply**: firm's **information set is endogenous to prices**: propose two-period two-product model to characterize **pricing** in counterfactual equilibria.

This Project: Empirical Strategy

1. How useful is monitoring?

   Evaluate the degree to which the IT can address information problems

2. How much information is revealed in equilibrium?

   Stricter privacy standards mean that the firm must “buy” data from consumers.

   Use IO tools to characterize the equilibrium price and quantity of information, and how it interacts with product market primitives.

This Project: Empirical Strategy

1. How useful is monitoring?

   Evaluate the degree to which the IT can address information problems

2. How much information is revealed in equilibrium?

   Stricter privacy standards mean that the firm must "buy" data from consumers.

   Use IO tools to characterize the equilibrium price and quantity of information, and how it interacts with product market primitives.

$\Longrightarrow$ No monitoring counterfactual

$\Longrightarrow$ Counterfactual equilibria: optimal pricing + data sharing

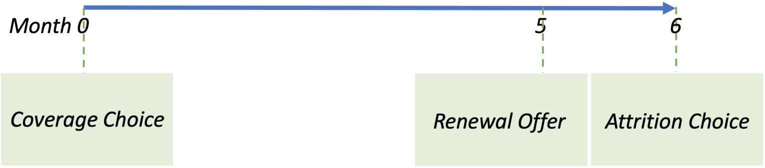
Roadmap

Background and Data

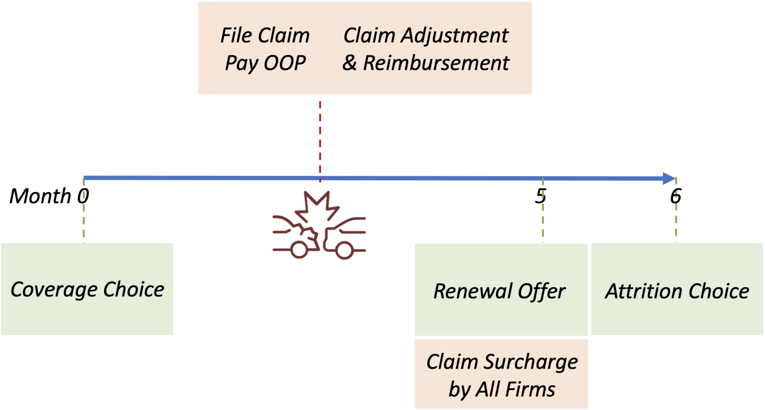
Demand and Estimation

Pricing and Equilibrium

Auto Insurance

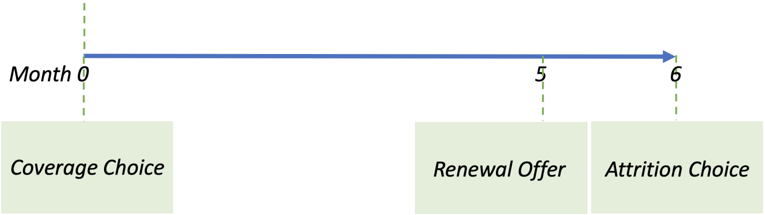


Auto Insurance



Auto Insurance - Data obs sum

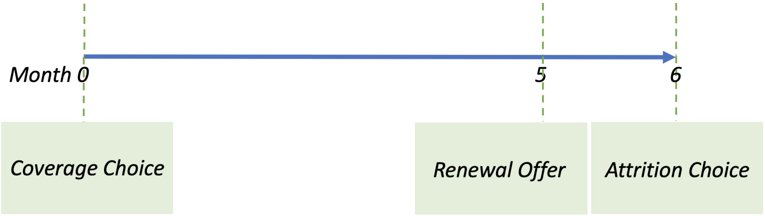
- ***Observable characteristics:*** 1-driver-1-vehicle, 22 states, 2012-16
- ***Quotes:*** liability limits ($30-500K, discrete choice)
- ***Competitor quotes:*** top 5 competitor per state
- ***Coverage choice:*** avg. $74K, and 48% in mandatory min
- ***Premium paid:*** avg. $380/period



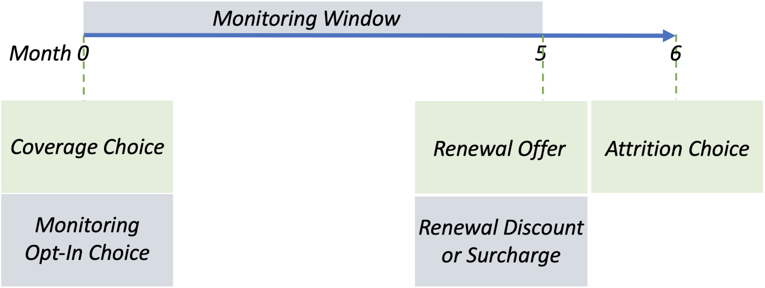
Auto Insurance - Data obs sum

- *Observable characteristics*
- *Quotes*
- *Competitor quotes*
- *Coverage choice*
- *Premium paid*

- ***Claim realization**:* avg. 1 per 10 yrs
  $6K/claim
- ***Renewal quote change***
- ***Attrition choice***

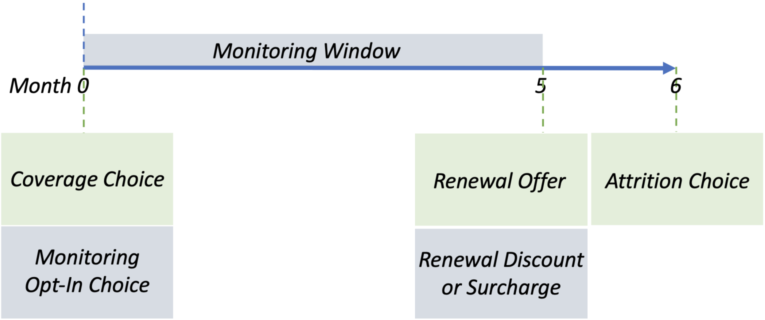


Monitoring Program



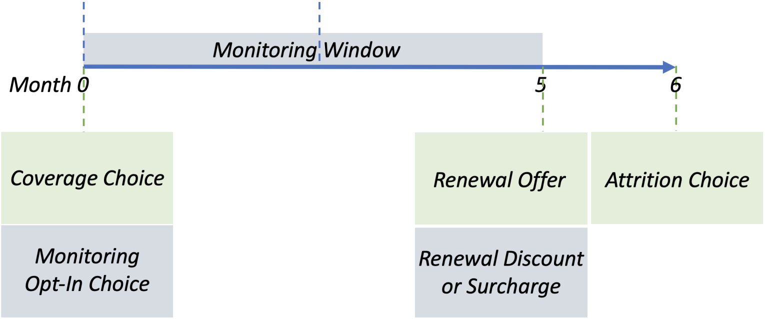
Monitoring Program

- ***Monitored behavior**:* mileage, hard brakes, speed, late night driving
- ***Duration**:* First period only (before renewal offer)
- ***Opt-in discount**:* First period only
- ***Renewal discount range**:* Lasts forever after first period



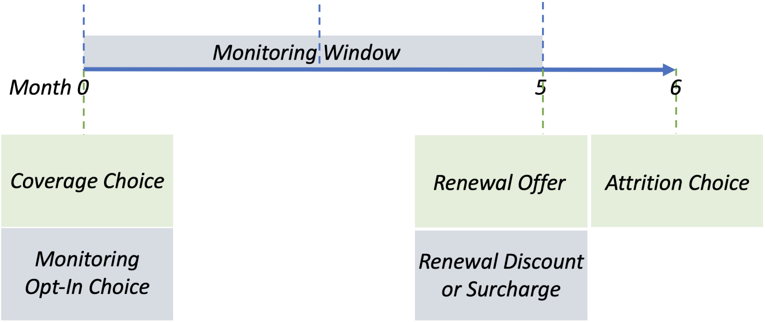
Monitoring Program

- *Monitored behavior*
- *Duration*
- *Opt-in discount*
- *Renewal discount range*
- ***Real-time feedback***



Monitoring Program

- *Monitored behavior*
- *Duration*
- *Opt-in discount*
- *Renewal discount range*
- *Real-time feedback*
- ***Score & discount**: proprietary data (verified with filing)*



Monitoring is useful in two ways

Result #1.1 Monitoring changes consumer behavior - drivers become 30% safer when they are monitored

Incentive Effect: drivers can exert effort to send a better signal of their type (Fama 1980, Holmstrom 1999, Villas-Boas and Fudenburg 2005).

- Within-driver comparison: opt-in drivers become riskier after the monitored (first) period; no such effect for drivers that did not opt in.

Result #1.2 Monitoring outcome still signals unobserved risk differences across drivers after monitoring.

Allocative Effect: better risk-rating can mitigate adverse selection and raise quantity (Akerlof 1970, Einav, Finklestein, and Cullen 2010).

- Receiving a score 1 sd above the mean $\rightarrow$ 29% higher claim count in subsequent (unmonitored) period, conditional on observables

Roadmap

Background and Data

Demand and Estimation

Pricing and Equilibrium

Structural Model - Overview

- **Cost** model - claim count C
- Monitoring **technology** - monitoring score s
- **Choice** model $d = \{f, y, m\}$

  Product choices - firm f and coverage y

  Information choice - monitoring opt-in m

Structural Model - Overview

- **Cost** model - claim count C: $\lambda(\sigma_\lambda, \theta)$
  - Consumers have latent risk types λ with unobserved heterogeneity σ_λ
  - Consumers can change λ by θ when monitored

Structural Model - Overview

why · cost mon · demand · id

- **Cost** model - claim count C: $\lambda(\sigma_\lambda, \theta)$
  - Consumers have latent risk types λ with unobserved heterogeneity σ_λ
  - Consumers can change λ by θ when monitored
- Monitoring **technology** - monitoring score s: σ_s
  - s is an informative signal of λ with precision σ_s

Structural Model - Overview

why | cost mon | demand | id

- **Cost** model - claim count C: $\lambda(\sigma_\lambda, \theta)$

  Consumers have latent risk types λ with unobserved heterogeneity σ_λ

  Consumers can change λ by θ when monitored

- Monitoring **technology** - monitoring score s: σ_s

  s is an informative signal of λ with precision σ_s

- **Choice** model $d = \{f, y, m\}$

  Product choices - firm f and coverage y: λ, γ, η

  - risk $\leftarrow \lambda$
  - preference: risk aversion (γ) and inertia for switching firms (η)

Structural Model - Overview

- **Cost** model - claim count C: $\lambda(\sigma_\lambda, \theta)$

  Consumers have latent risk types λ with unobserved heterogeneity σ_λ

  Consumers can change λ by θ when monitored

- Monitoring **technology** - monitoring score s: σ_s

  s is an informative signal of λ with precision σ_s

- **Choice** model $d = \{f, y, m\}$

  Product choices - firm f and coverage y: λ, γ, η

  risk $\leftarrow \lambda$

  preference: risk aversion (γ) and inertia for switching firms (η)

  Information choice - monitoring opt-in m

  - financial risk and rewards
    - risk reduction $\leftarrow \lambda(\theta)$
    - renewal discount and reclassification risk $\leftarrow \lambda(\sigma_\lambda)$, γ, σ_s
  - unobserved disutility from being monitored (ξ)

Structural Model - Overview

- **Cost** model - claim count C: $\lambda(\sigma_\lambda, \theta)$

    Consumers have latent risk types λ with unobserved heterogeneity σ_λ

    Consumers can change λ by θ when monitored

- Monitoring **technology** - monitoring score s: σ_s

    s is an informative signal of λ with precision σ_s

- **Choice** model $d = \{f, y, m\}$

  Product choices - firm f and coverage y: λ, γ, η

    risk $\leftarrow \lambda$

    preference: risk aversion (γ) and inertia for switching firms (η)

  Information choice - monitoring opt-in m: $\lambda, \sigma_s, \gamma$, ξ

    financial risk and rewards

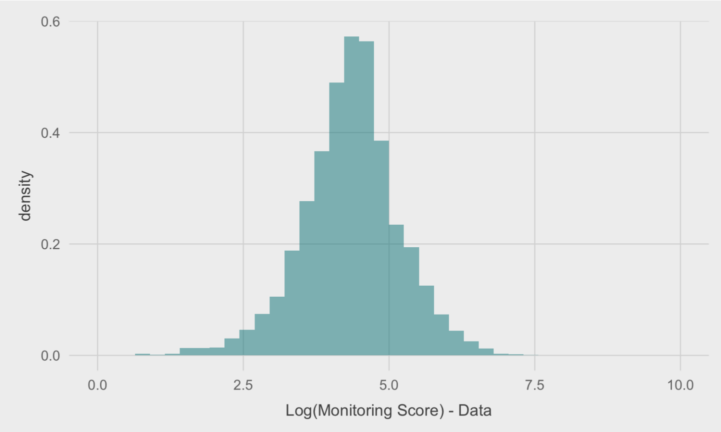
      risk reduction $\leftarrow \lambda(\theta)$

      renewal discount and reclassification risk $\leftarrow \lambda(\sigma_\lambda)$, γ, σ_s

    unobserved disutility from being monitored (ξ)

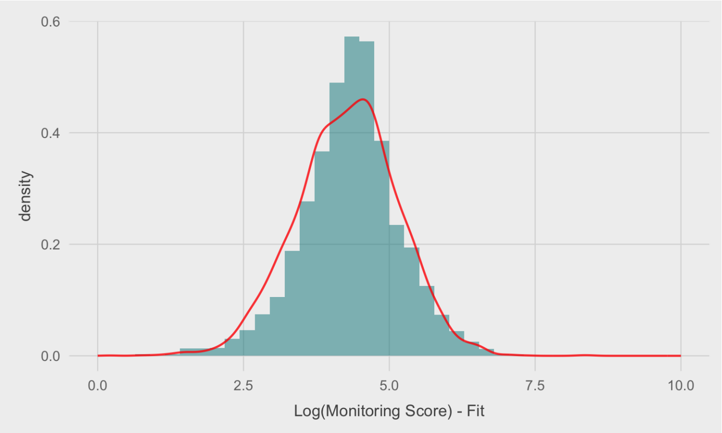
Estimation

Simulated MLE. Goal: fit monitoring share + selection pattern (*who* opts in). estimation



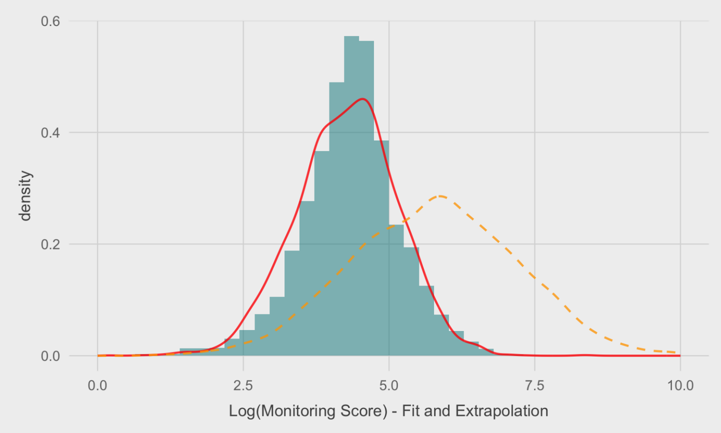
Fit

Simulated MLE. Goal: fit monitoring share + selection pattern (*who* opts in). estimation



Advantageous Selection into Monitoring...

Result #Demand.1 Safer drivers are more likely to opt in...



...But Also Large Demand Friction Against Monitoring cost

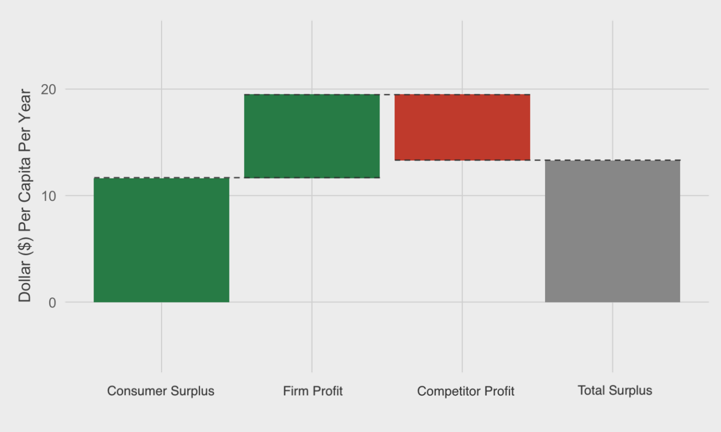
Result #Demand.2 ...but large friction exists so that most who can financial benefit do not opt in.

- $\hat{\xi}(x, \lambda)$ has mean \$93; higher for {younger, less educated, older cars, poorer prior insurance or traffic records}.
- $\hat{\xi}(x, \lambda)$ is increasing with λ: conditional on expected financial discounts, safer drivers are more likely to opt in $\rightarrow$ exacerbates advantageous selection into monitoring

risk aversion switch inertia cross validation

Welfare Calculation: Current - No Monitoring

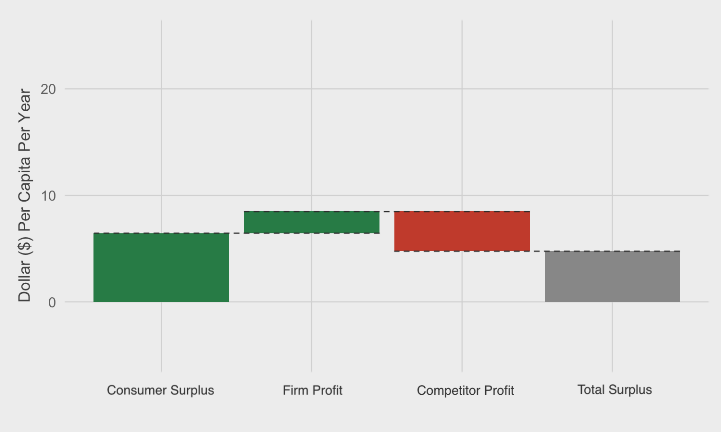
Introducing monitoring increases firm profit, consumer welfare, and total surplus. detail



- hold baseline (unmonitored) prices fixed event
- set resource cost of monitoring is $35 per capita

Welfare Calculation: Tease Out Allocative Effect

assume away incentive effect: drivers are no safer when monitored. detail



- ∼64% of the surplus gain comes from risk reduction (incentive effect)
- competitive cream-skimming with better risk information (vs. Rothschild and Stiglitz 1976): overall profit ↓ and quantity ↑

Roadmap

Background and Data

Demand and Estimation

Pricing and Equilibrium

Pricing and Counterfactual Equilibria detail prices

Pricing Model

- Firm profit
  - 2-period: pre- and post-information revelation
  - 2-product: insurance with and without monitoring

Pricing and Counterfactual Equilibria detail prices

Pricing Model

- Firm profit
  - 2-period: pre- and post-information revelation
  - 2-product: insurance with and without monitoring

- Firm actions: 3 types of price adjustments for monitoring
  - Parameterization corresponds to how monitoring changes the firm's *information set*

    $t = 0$ does not observe monitoring score

    $m = 0$: κ_0 surcharge unmonitored pool
    $m = 1$: κ_1 discount monitored pool

    $t = 1$ observes monitoring score iff $m = 1$

    $m = 1$: κ_s linear rent-sharing regime with monitored drivers

Pricing and Counterfactual Equilibria detail prices

Pricing Model

- Firm profit: 2-period-2-product
- Firm actions: 3 types of price adjustments for monitoring
  $t = 0, m = 0 : \kappa_0$ surcharge unmonitored pool
  $t = 0, m = 1 : \kappa_1$ discount monitored pool
  $t = 1, m = 1 : \kappa_s$ linear rent-sharing regime with monitored drivers

Counterfactuals

- **Optimal pricing** of monitoring
  - marginal cost of monitoring is known
  - holding fixed competitor prices

Pricing and Counterfactual Equilibria detail prices

Pricing Model

- Firm profit: 2-period-2-product
- Firm actions: 3 types of price adjustments for monitoring
  $t = 0, m = 0 : \kappa_0$ surcharge unmonitored pool
  $t = 0, m = 1 : \kappa_1$ discount monitored pool
  $t = 1, m = 1 : \kappa_s$ linear rent-sharing regime with monitored drivers

Counterfactuals

- Optimal pricing of monitoring
  - marginal cost of monitoring is known
  - holding fixed competitor prices
- **Data sharing regulation** that eliminates propiretary data
  - assume competitors have symmetric belief and profit function
  - action: only set a single alternative rent-sharing scheme $\kappa_{s,-f^\star}$ to poach monitored drivers $(m = 1)$ at $t = 1$

Optimal Pricing

Result #Supply.1: Product market competition → firm can't coerce drivers into monitoring.

	Current Regime	Optimal Pricing
Surplus & division (/capita/year)		
Firm Profit		+14.7
Competitor Profit		−11.0
Consumer Welfare (in CE)		+4.7
Total Surplus		+8.4
Monitoring Market Share (%)	3.0%	4.4% ↑
Pricing: First Period (%)		
Unmonitored surcharge κ_0	0.0%	2.7% ↑
Opt-in discount κ_1	4.6%	22.1% ↑↑
Pricing: Second Period		
Rent-sharing κ_s	1x	0.80x ↓
Competitor rent-sharing $\kappa_{s,-f^*}$	-	-

- e.g. Post-GDPR, Google and Facebook can contingent service upon data consent, smaller firms/websites cannot (Schechner 2018).

Optimal Pricing

Result #Supply.2: Firm "buys" consumer data with upfront discount expecting ex-post rent.

	Current Regime	Optimal Pricing
Surplus & division (/capita/year)		
Firm Profit		+14.7
Competitor Profit		−11.0
Consumer Welfare (in CE)		+4.7
Total Surplus		+8.4
Monitoring Market Share (%)	3.0%	4.4% ↑
Pricing: First Period (%)		
Unmonitored surcharge κ_0	0.0%	2.7%
Opt-in discount κ_1	4.6%	22.1% ↑↑
Pricing: Second Period		
Rent-sharing κ_s	1x	0.80x ↓
Competitor rent-sharing $\kappa_{s,-f^\star}$	-	-

- Information ("privacy") choice is contextual (Nissenbaum 2009), and firms can greatly affect that context through pricing.

Counterfactual Equilibrium: Information Sharing

Data sharing undermines firm incentives to "buy" consumer data.

	Optimal Pricing	Data Sharing
Surplus & division (/capita/year)		
Firm Profit		-11.9
Competitor Profit		$+8.9$
Consumer Welfare (in CE)		-2.5
Total Surplus		-5.5
Monitoring Market Share (%)	4.4%	3.4% ↓
Pricing: First Period (%)		
Unmonitored surcharge κ_0	2.7%	1.6% ↓
Opt-in discount κ_1	22.1%	8.3% ↓
Pricing: Second Period		
Rent-sharing κ_s	0.80x	1.14x ↑
Competitor rent-sharing $\kappa_{s,-f^\star}$	-	1.81x

- hurts welfare as monitoring is "socially-valuable" (Posner 1979).
- real-world regulation: data portability + algorithm transparency

Takeaway

The **optimal privacy standards** should depend on

Social value of the data collected, and...

- Don't underestimate how data collection can change consumer behavior

Demand and supply primitives in the product market

- Customers self-select into revealing information
- Firms can compete on information through prices

Information structure is an equilibrium object. Regressing other equilibrium outcomes on the amount of information fall prey to the same critiques as the S-C-P approach