Buying Data from Consumers

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

Yizhou Jin       Shosh Vasserman
UC Berkeley      Stanford

FTC Microeconomics Conference
November 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.”
Direct Transactions of Consumer Data in General

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

⇒ No monitoring counterfactual
⇒ Counterfactual equilibria: optimal pricing + data sharing
Roadmap

Background and Data

Demand and Estimation

Pricing and Equilibrium
Auto Insurance

Month 0

Coverage Choice

Renewal Offer

Attrition Choice
Auto Insurance

Month 0

Coverage Choice

File Claim Pay OOP

Claim Adjustment & Reimbursement

5

Renewal Offer

Attrition Choice

Claim Surcharge by All Firms

6
- **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

- 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

Month 0

Coverage Choice

Monitoring Opt-In Choice

Renewal Offer

Attrition Choice

Renewal Discount or Surcharge
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 Diagram]

- **Month 0**
- **Monitoring Window**: 5 months
- **Coverage Choice**
- **Monitoring Opt-In Choice**
- **Renewal Offer**
- **Attrition Choice**
- **Renewal Discount or Surcharge**
Monitoring Program

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

Monitoring Window

Month 0 5 6

Coverage Choice

Monitoring Opt-In Choice

Renewal Offer

Renewal Discount or Surcharge

Attrition Choice
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 $C: \lambda(\sigma, \theta)$
  - Consumers have latent risk types $\lambda$ with unobserved heterogeneity $\sigma$
  - Consumers can change $\lambda$ by $\theta$ when monitored
Structural Model - Overview

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

- Monitoring **technology** - monitoring score $s$: $\sigma_s$
  - $s$ is an informative signal of $\lambda$ with precision $\sigma_s$
Structural Model - Overview

- **Cost** model - claim count $C: \lambda(\sigma_\lambda, \theta)$
  
  Consumers have latent risk types $\lambda$ with unobserved heterogeneity $\sigma_\lambda$

  Consumers can change $\lambda$ by $\theta$ when monitored

- Monitoring **technology** - monitoring score $s: \sigma_s$
  
  $s$ is an informative signal of $\lambda$ with precision $\sigma_s$

- **Choice** model $d = \{f, y, m\}$
  
  Product choices - firm $f$ and coverage $y: \lambda, \gamma, \eta$

  ▶ risk $\leftarrow \lambda$

  ▶ preference: risk aversion ($\gamma$) and inertia for switching firms ($\eta$)
Structural Model - Overview

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

- Monitoring **technology** - monitoring score $s$: $\sigma_s$
  
  $s$ is an informative signal of $\lambda$ with precision $\sigma_s$

- **Choice** model $d = \{f, y, m\}$
  
  **Product choices** - firm $f$ and coverage $y$: $\lambda, \gamma, \eta$
  
  risk $\leftarrow \lambda$
  
  preference: risk aversion ($\gamma$) and inertia for switching firms ($\eta$)

  **Information choice** - monitoring opt-in $m$
  
  - financial risk and rewards
    
    - risk reduction $\leftarrow \lambda(\theta)$
    
    - renewal discount and reclassification risk $\leftarrow \lambda(\sigma_\lambda), \gamma, \sigma_s$
  
  - unobserved disutility from being monitored ($\xi$)
Structural Model - Overview

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

- **Monitoring technology** - monitoring score $s$: $\sigma_s$
  $s$ is an informative signal of $\lambda$ with precision $\sigma_s$

- **Choice** model $d = \{f, y, m\}$
  
  **Product choices** - firm $f$ and coverage $y$: $\lambda, \gamma, \eta$
  
  risk $\leftarrow \lambda$
  
  preference: risk aversion ($\gamma$) and inertia for switching firms ($\eta$)

  **Information choice** - monitoring opt-in $m$: $\lambda, \sigma_s, \gamma, \xi$
  
  financial risk and rewards
  
  risk reduction $\leftarrow \lambda(\theta)$
  
  renewal discount and reclassification risk $\leftarrow \lambda(\sigma_\lambda), \gamma, \sigma_s$
  
  unobserved disutility from being monitored ($\xi$)
Choice Model - Realized Utility

Utility for choice \( d \), conditional on realized \( C \) and \( s \).

- Consumer \( i \) in period \( t \) with observables \( x_{it} \)

\[
u_{idt}(C, s) = u_\gamma \left[ w_{it} - p_{idt} - 1_{d,t-1} \cdot \psi_{idt} - e(C, y_d) - p_{idt} \cdot R_{idt}(C, s) \right]
\]

- one-year (two-period) horizon + myopic (Handel 2013)
- risk aversion \( \gamma \) with quadratic \( u(\cdot) \) → risk premium (accident oop + renewal price).
Choice Model - Realized Utility

Utility for choice \( d \), conditional on realized \( C \) and \( s \).

- Consumer \( i \) in period \( t \) with observables \( x_{it} \)

\[
u_{idt}(C, s) = u_\gamma \left[ w_{it} - p_{idt} - 1_{d,t-1} \cdot \psi_{idt} - e(C, y_d) - p_{idt} \cdot R_{idt}(C, s) \right]
\]

- one-year (two-period) horizon + myopic (Handel 2013)
- risk aversion \( \gamma \) with quadratic \( u(\cdot) \) → risk premium (accident oop + renewal price).
Choice Model - Realized Utility

Utility for choice $d$, conditional on realized $C$ and $s$.

$$u_{idt}(C, s) = u_\gamma \left[ w_{it} - p_{idt} - 1_{d,t-1} \cdot \psi_{idt} - e(C, y_d) - p_{idt} \cdot R_{idt}(C, s) \right]$$

- one-year (two-period) horizon + myopic (Handel 2013)
- risk aversion $\gamma$ with quadratic $u(\cdot) \rightarrow$ risk premium (accident oop + renewal price).

$$\psi_{it}(d, d_i, t-1) = 1_{f_d, t-1} \cdot \eta(x_{it}; \theta_\eta) + 1_{m_d} \cdot 1_{t=0} \cdot \xi(x_{it}, \lambda; \theta_\xi)$$

- firm-switching inertia
- monitoring disutility
Choice Model - Realized Utility

Utility for choice $d$, conditional on realized $C$ and $s$.

\[ u_{idt}(C, s) = u_\gamma \left[ w_{it} - p_{idt} - \mathbf{1}_{d,t-1} \cdot \psi_{idt} - e(C, y_d) - p_{idt} \cdot R_{idt}(C, s) \right] \]

- one-year (two-period) horizon + myopic (Handel 2013)
- risk aversion $\gamma$ with quadratic $u(\cdot) \rightarrow$ risk premium (accident oop + renewal price).

\[ \psi_{it}(d, d_i, t-1) = \mathbf{1}_{f_{d,t-1}} \cdot \eta(x_{it}; \theta_\eta) + \mathbf{1}_{m_d} \cdot \mathbf{1}_{t=0} \cdot \xi(x_{it}, \lambda; \theta_\xi) \]

- firm-switching inertia $\eta \rightarrow$ imperfect competition without monitoring
- monitoring disutility $\xi \rightarrow$ unobserved costs/pref. or irrationality
Demand Identification

- Very rich price variation
  - Conditional on observables, prices vary across **zips and time**
  - Flexible controls for zipcode income, year trend and seasonality

- Contract space variation
  - **Monitoring eligibility** depends on state and time
  - **Mandatory minimum** changes in two states - use one for estimation and the other for cross-validation
Demand Identification

- Very rich price variation
  - Conditional on observables, prices vary across zips and time
  - Flexible controls for zipcode income, year trend and seasonality
- Contract space variation
  - Monitoring eligibility depends on state and time
  - Mandatory minimum changes in two states - use one for estimation and the other for cross-validation
- Risk preference ($\gamma$) pins down gradient of WTP across multiple coverages $y$ (Cohen Einav 2007)
- Switching inertia ($\eta$) governs attrition rate $f$
- Monitoring disutility ($\xi$) rationalizes the monitoring share not explained by financial risk and rewards of monitoring.
  - Vary based on $\lambda$ to fit the observed selection pattern: $\theta_{\xi,\lambda}$. 

\(\downarrow\)
Estimation

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

Simulated MLE. Goal: fit monitoring share + selection pattern (*who* opts in).
Advantageous Selection into Monitoring...

Result #Demand.1 Safer drivers are more likely to opt in...
...But Also Large Demand Friction Against Monitoring

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 $\lambda$: conditional on expected financial discounts, safer drivers are more likely to opt in → 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.

- hold baseline (unmonitored) prices fixed
- 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.

- ~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

Pricing Model

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

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

  \[ m = 0 : \kappa_0 \] surcharge unmonitored pool
  \[ m = 1 : \kappa_1 \] discount monitored pool

  \[ t = 1 \] observes monitoring score iff \[ m = 1 \]

  \[ m = 1 : \kappa_s \] linear rent-sharing regime with monitored drivers
Pricing and Counterfactual Equilibria

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

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 proprietary data
  - assume competitors have symmetric belief and profit function
  - action: only set a single alternative rent-sharing scheme \( \kappa_s, -f^* \) to poach monitored drivers \((m = 1)\) at \( t = 1 \)
Optimal Pricing

**Result #Supply.1**: Product market competition $\rightarrow$ firm can’t coerce drivers into monitoring.

<table>
<thead>
<tr>
<th>Current Regime</th>
<th>Optimal Pricing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Surplus &amp; division (/capita/year)</strong></td>
<td></td>
</tr>
<tr>
<td>Firm Profit</td>
<td>+14.7</td>
</tr>
<tr>
<td>Competitor Profit</td>
<td>−11.0</td>
</tr>
<tr>
<td>Consumer Welfare (in CE)</td>
<td>+4.7</td>
</tr>
<tr>
<td>Total Surplus</td>
<td>+8.4</td>
</tr>
</tbody>
</table>

| Monitoring Market Share (%) | 3.0% | 4.4% ↑ |

| Pricing: First Period (%) | | |
|--------------------------|-----------------|
| Unmonitored surcharge $\kappa_0$ | 0.0% | 2.7% ↑ |
| Opt-in discount $\kappa_1$ | 4.6% | 22.1% ↑↑ |

| Pricing: Second Period | | |
|-----------------------|-----------------|
| Rent-sharing $\kappa_s$ | 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.

<table>
<thead>
<tr>
<th>Current Regime</th>
<th>Optimal Pricing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Surplus &amp; division (/capita/year)</strong></td>
<td></td>
</tr>
<tr>
<td>Firm Profit</td>
<td>+14.7</td>
</tr>
<tr>
<td>Competitor Profit</td>
<td>−11.0</td>
</tr>
<tr>
<td>Consumer Welfare (in CE)</td>
<td>+4.7</td>
</tr>
<tr>
<td>Total Surplus</td>
<td>+8.4</td>
</tr>
<tr>
<td><strong>Monitoring Market Share (%)</strong></td>
<td></td>
</tr>
<tr>
<td>3.0%</td>
<td>4.4% ↑</td>
</tr>
<tr>
<td><strong>Pricing: First Period (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Unmonitored surcharge $\kappa_0$</td>
<td>0.0%</td>
</tr>
<tr>
<td>Opt-in discount $\kappa_1$</td>
<td>4.6%</td>
</tr>
<tr>
<td>2.7%</td>
<td>22.1% ↑↑</td>
</tr>
<tr>
<td><strong>Pricing: Second Period</strong></td>
<td></td>
</tr>
<tr>
<td>Rent-sharing $\kappa_s$</td>
<td>1x</td>
</tr>
<tr>
<td>Competitor rent-sharing $\kappa_{s,f^*}$</td>
<td>-</td>
</tr>
<tr>
<td>0.80x ↓</td>
<td>-</td>
</tr>
</tbody>
</table>

- 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.

<table>
<thead>
<tr>
<th></th>
<th>Optimal Pricing</th>
<th>Data Sharing</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Surplus &amp; division (/capita/year)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm Profit</td>
<td>−11.9</td>
<td>−11.9</td>
</tr>
<tr>
<td>Competitor Profit</td>
<td>+8.9</td>
<td>+8.9</td>
</tr>
<tr>
<td>Consumer Welfare (in CE)</td>
<td>−2.5</td>
<td>−2.5</td>
</tr>
<tr>
<td>Total Surplus</td>
<td>−5.5</td>
<td>−5.5</td>
</tr>
</tbody>
</table>

**Monitoring Market Share (%)**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4%</td>
<td>3.4% ↓</td>
<td></td>
</tr>
</tbody>
</table>

**Pricing: First Period (%)**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Unmonitored surcharge $\kappa_0$</td>
<td>2.7%</td>
<td>1.6% ↓</td>
</tr>
<tr>
<td>Opt-in discount $\kappa_1$</td>
<td>22.1%</td>
<td>8.3% ↓</td>
</tr>
</tbody>
</table>

**Pricing: Second Period**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rent-sharing $\kappa_s$</td>
<td>0.80x</td>
<td>1.14x ↑</td>
</tr>
<tr>
<td>Competitor rent-sharing $\kappa_{s,-f^*}$</td>
<td>-</td>
<td>1.81x</td>
</tr>
</tbody>
</table>

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

Monitoring is welfare- and profit-improving.

1. Drivers respond to financial incentives and become a lot safer during monitoring

2. Strong advantageous selection but opt-in rate (amount of information) is low due to demand frictions and price competition

3. Insurer’s property right to monitoring data strongly influence their effort to elicit data through pricing
Takeaway

**Policy** Data regulation in insurance / broader privacy standards should depend on

- Social value of the data collected, and...
  - Incentive effects are important
- Demand and supply primitives in the product market
  - Consumers’ risk and privacy preferences + product market power
  \[\implies\] require disclosure of price/quantity effects (price filings/algorithm audits) vs. outright ban or full transparency

**Research** 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.