Buying Data from Consumers

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

Yizhou Jin
UC Berkeley

Shosh Vasserman
Stanford

NBER SI Digitization & IO
July 2019

Motivation

IT + Privacy Standards $\rightarrow$ 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



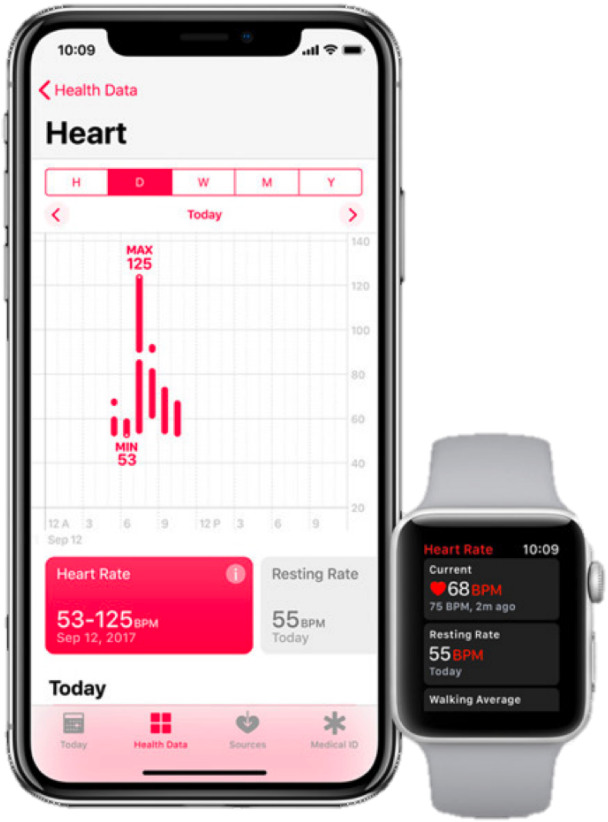




Direct Transactions of Consumer Data in General

back

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

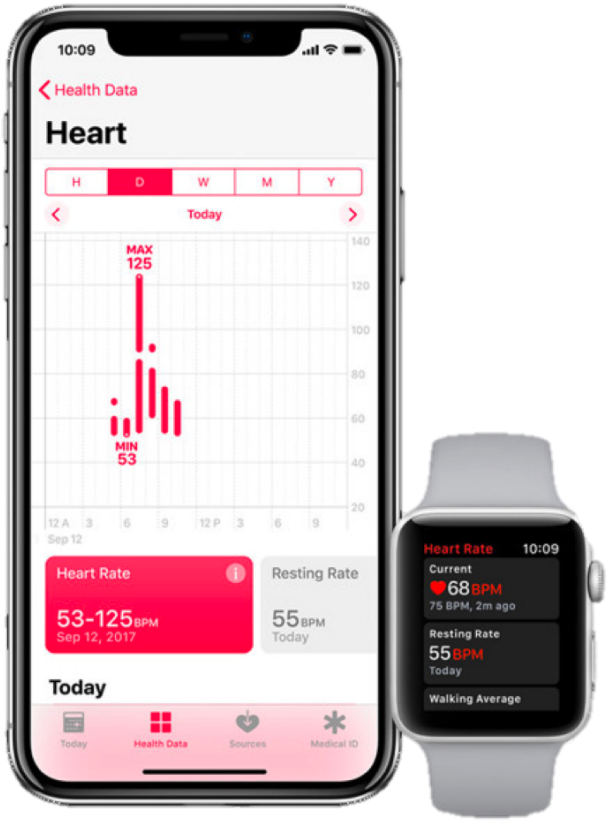


Vitality - John Hancock Life Insurance

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 Other

Ways to Improve Score:

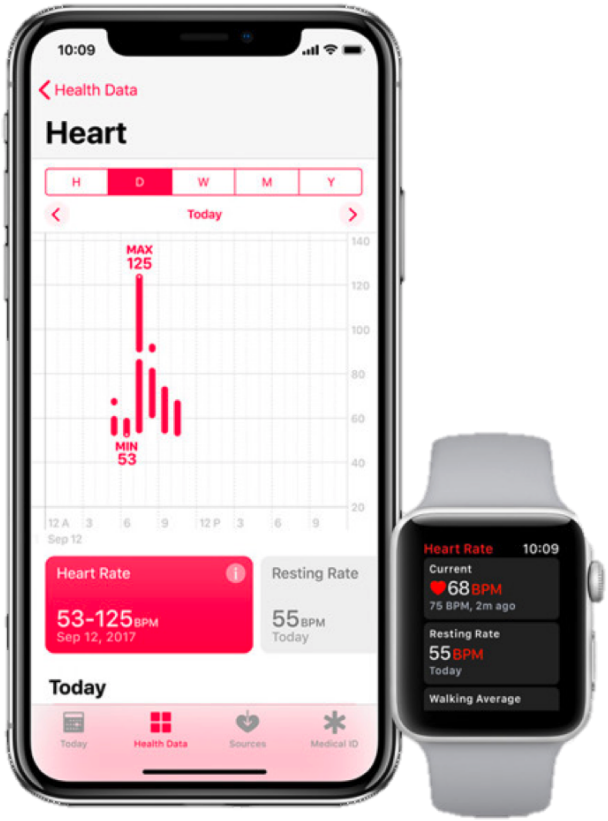
- Receive Income through the app
- Pay Utility Bills through the app
- Connect with friends on the app

Alibaba - Proprietary Credit Scores

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 Other

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

Motivation

IT + Privacy Standards $\rightarrow$ Direct transactions of consumer data

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

$\Longrightarrow$ Economic impact wrt. total surplus and its division?

- More information can improve market efficiency
- Consumers self select into providing data
- Proprietary data can facilitate higher markups

Motivation

IT + Privacy Standards $\rightarrow$ Direct transactions of consumer data

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

$\Longrightarrow$ Economic impact wrt. total surplus and its division?

- $\Longrightarrow$ How useful is monitoring technology?
- $\Longrightarrow$ How much information is revealed in equilibrium?

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. What 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. What 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. What much information is revealed in equilibrium?
   - ▷ Demand: estimate structural parameters to capture complex correlations among consumers' **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. What much information is revealed in equilibrium?
   - ▷ Demand: estimate structural parameters to capture complex correlations among consumers' 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. What much information is revealed in equilibrium?
   - ▷ Demand: estimate structural parameters to capture complex correlations among consumers' 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.

$\Longrightarrow$ No monitoring counterfactual

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

This Project: Preview of Findings

ag

1. Monitoring influences profit and welfare in two ways:
   - ▷ Incentive: drivers become 30% **safer** when monitored.
   - ▷ Allocative: outcome signals **unobserved risk differences** across drivers.

2. Information equilibrium is shaped by demand and supply forces:
   - ▷ Demand: safer drivers are more likely to opt in, but the average driver faces large friction against monitoring.
   - ▷ Supply: pricing of monitoring depends on product market competition and whether data collected are proprietary

$\Longrightarrow$ Introducing monitoring raises both profit and welfare, but mostly because drivers are safer.

$\Longrightarrow$ Information sharing harms both profit and welfare: pricing adjustments lead to fewer drivers monitored in equilibrium.

This Project: Preview of Findings

ag

1. Monitoring influences profit and welfare in two ways:
   - ▷ Incentive: drivers become 30% safer when monitored.
   - ▷ Allocative: outcome signals unobserved risk differences across drivers.

2. Information equilibrium is shaped by demand and supply forces:
   - ▷ Demand: **safer** drivers are more likely to opt in, but the average driver faces large **friction** against monitoring.
   - ▷ Supply: pricing of monitoring depends on **product market competition** and whether data collected are **proprietary**

$\Longrightarrow$ Introducing monitoring raises both profit and welfare, but mostly because drivers are safer.

$\Longrightarrow$ Information sharing harms both profit and welfare: pricing adjustments lead to fewer drivers monitored in equilibrium.

Roadmap

Background and Data

Demand and Estimation

Pricing and Equilibrium

Roadmap

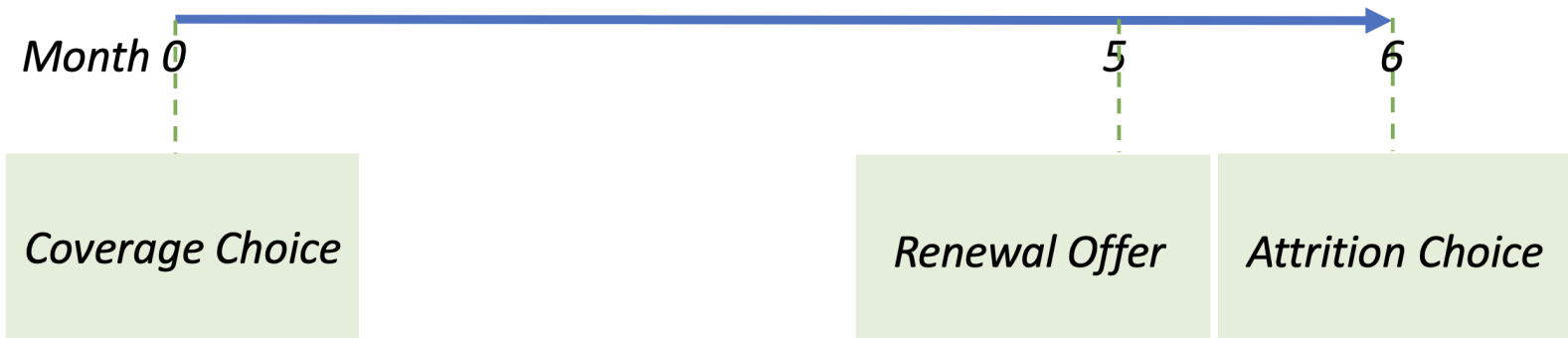
Background and Data

1. Timing and Setting
2. Data and Reduced-Form Evidence

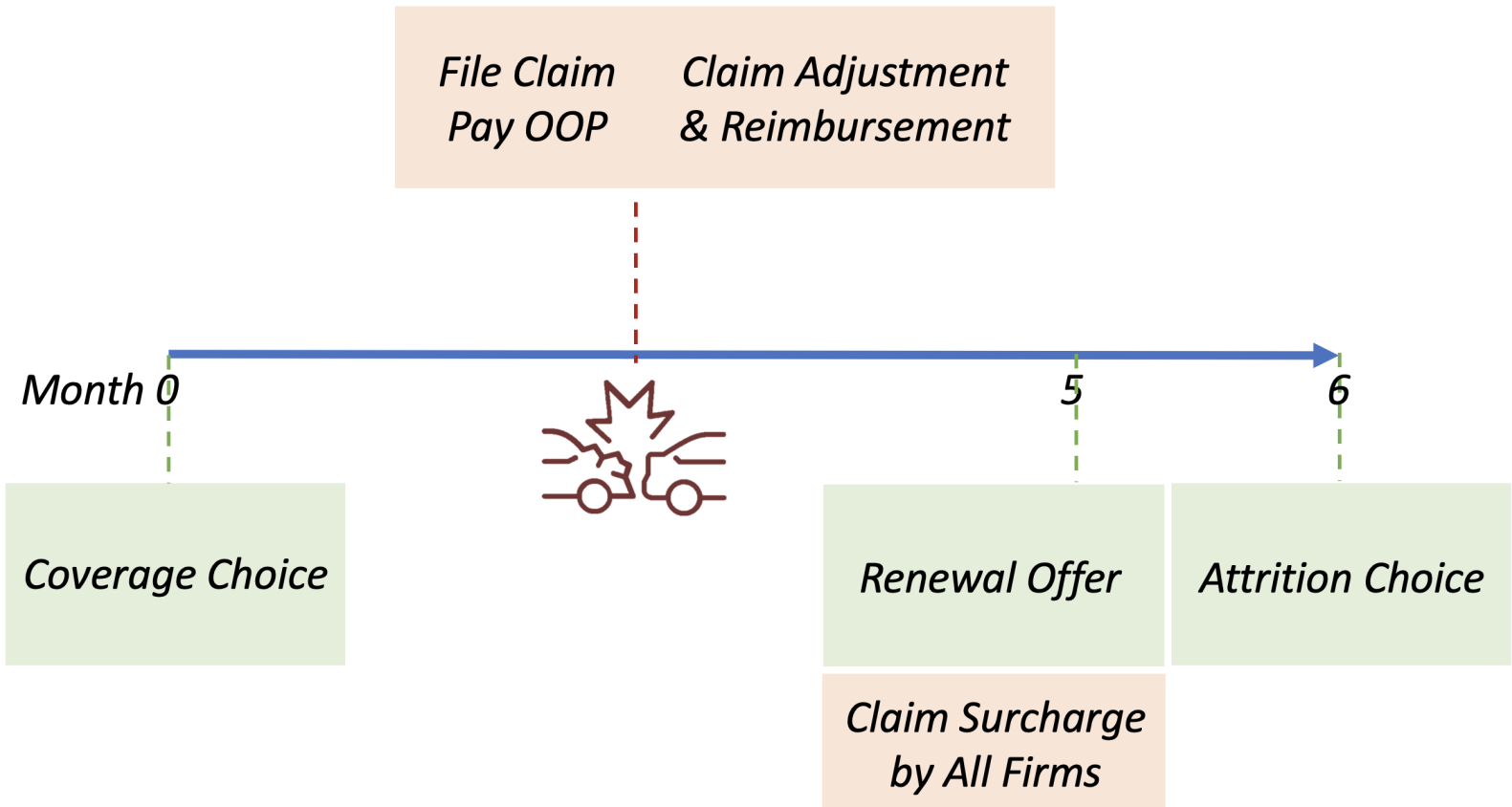
Demand and Estimation

Pricing and Equilibrium

Auto Insurance

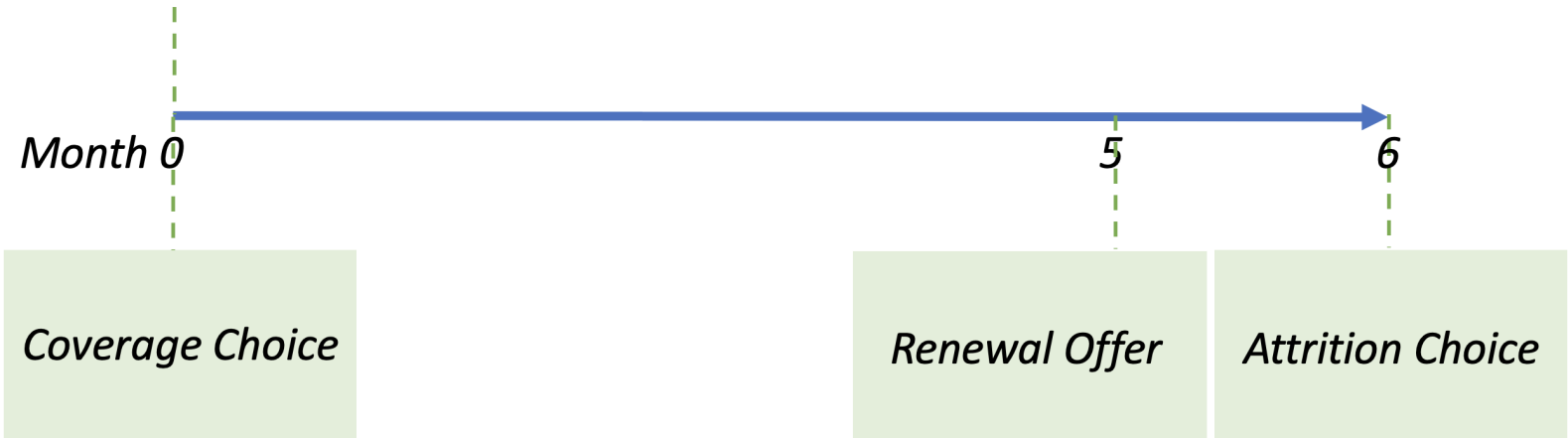


Auto Insurance



Auto Insurance - Data obs sum

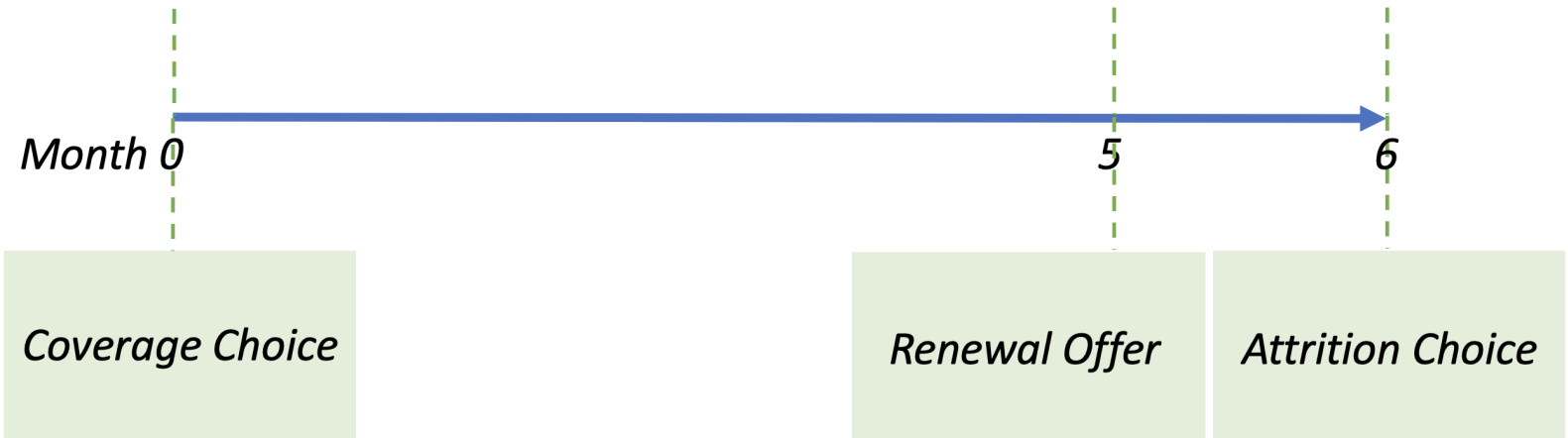
- ***Observable characteristics:*** 1-driver-1-vehicle, 22 states, 2012-16
- ***Quotes:*** liability limits ($30-500K, discrete choice)
- ***Competitor quotes:*** top 5 competitor per state
- ***Coverage choice:*** avg. $74K, and 48% in mandatory min
- ***Premium paid:*** avg. $380/period



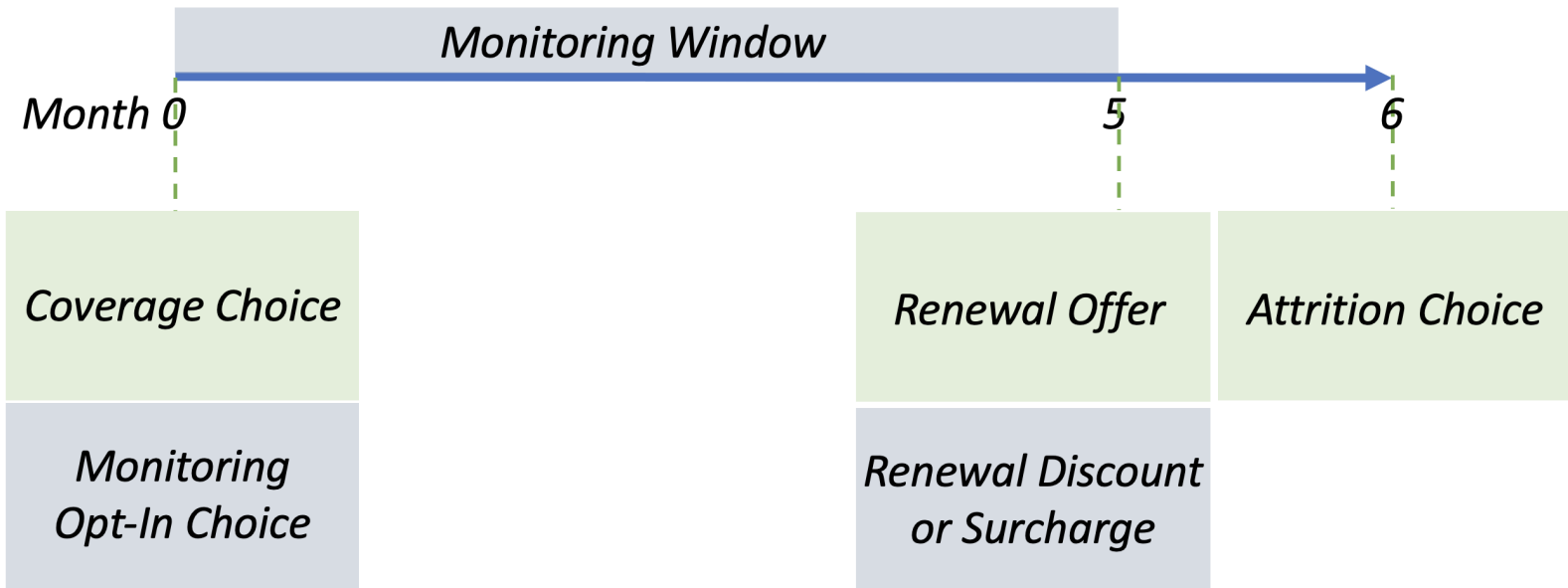
Auto Insurance - Data obs sum

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

- ***Claim realization**:* avg. 1 per 10 yrs
  $6K/claim
- ***Renewal quote change***
- ***Attrition choice**:* 15% first period

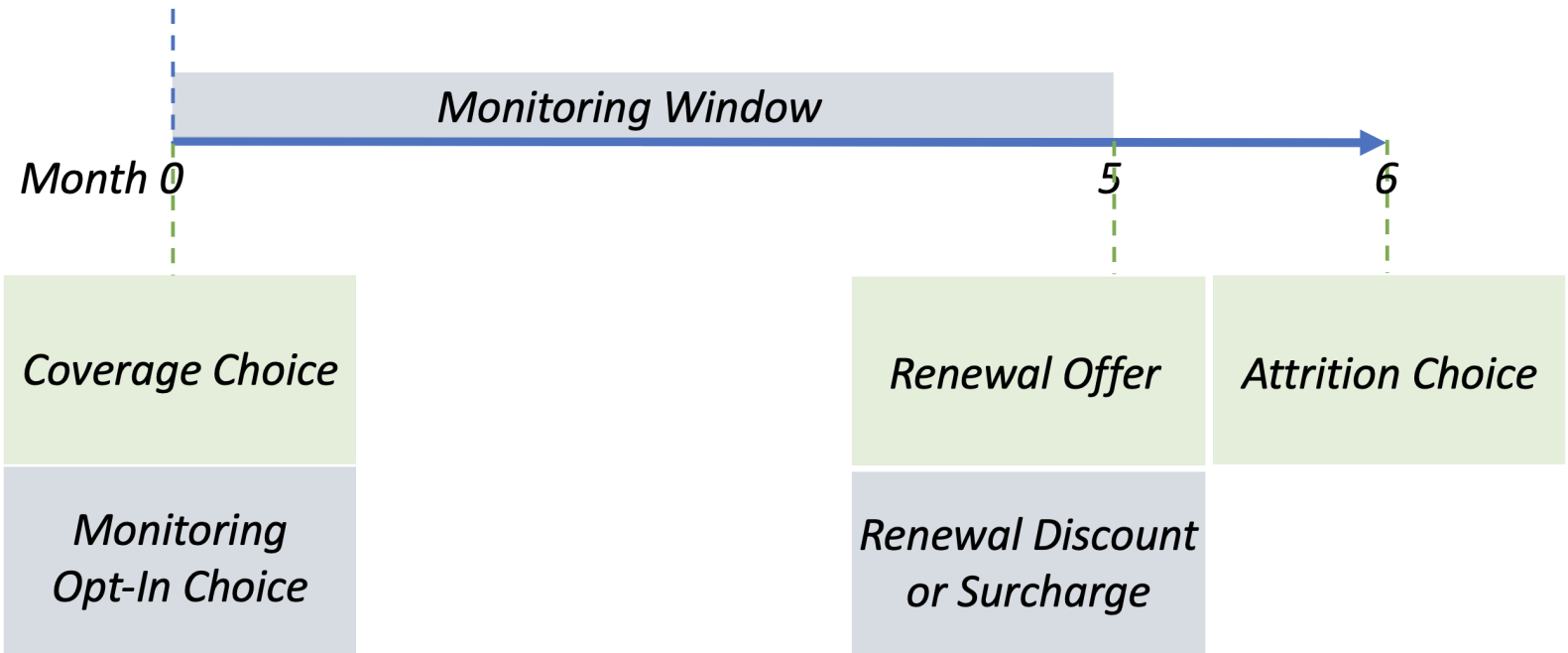


Monitoring Program



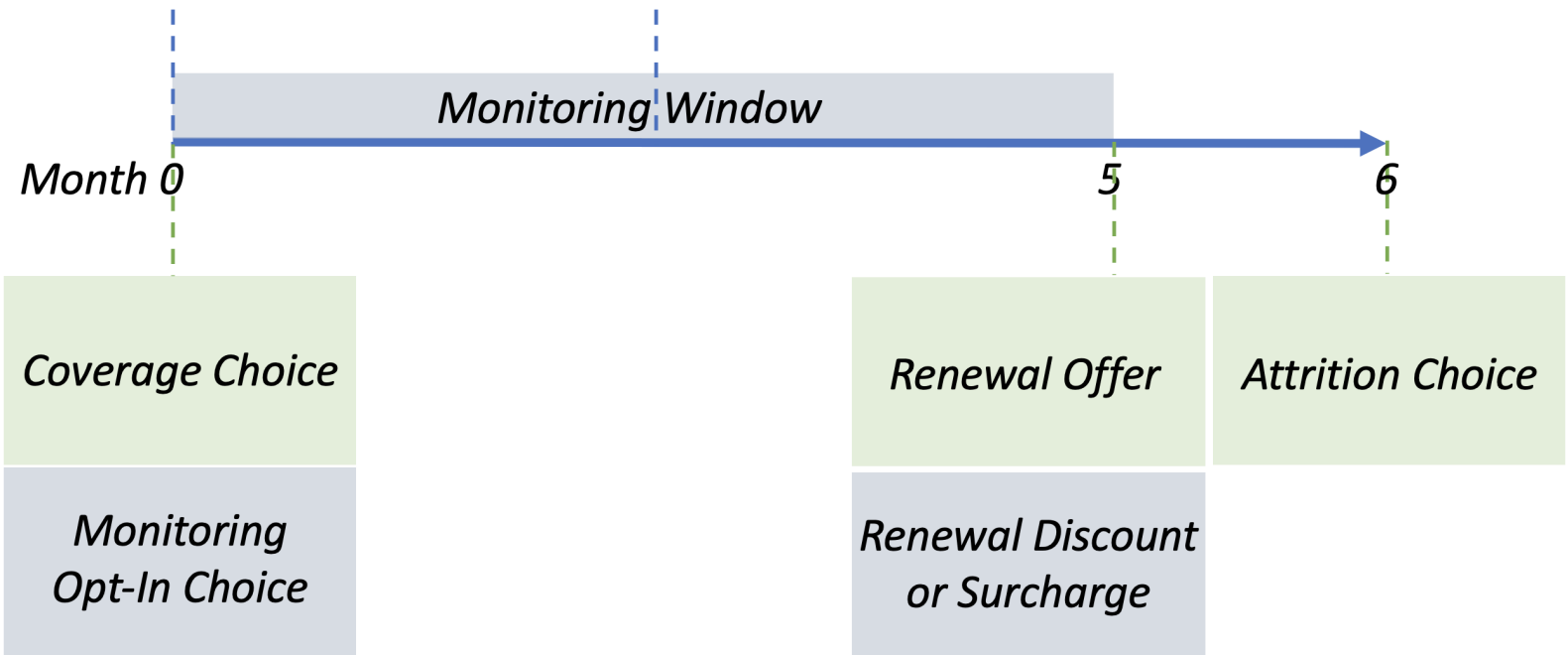
Monitoring Program

- ***Punishable behavior:*** mileage, hard brakes, speed, late night driving
- ***Duration:*** First period only (before renewal offer)
- ***Opt-in discount:*** 0—10%, only for first period
- ***Renewal discount range:*** -10—30%, lasts forever after first period



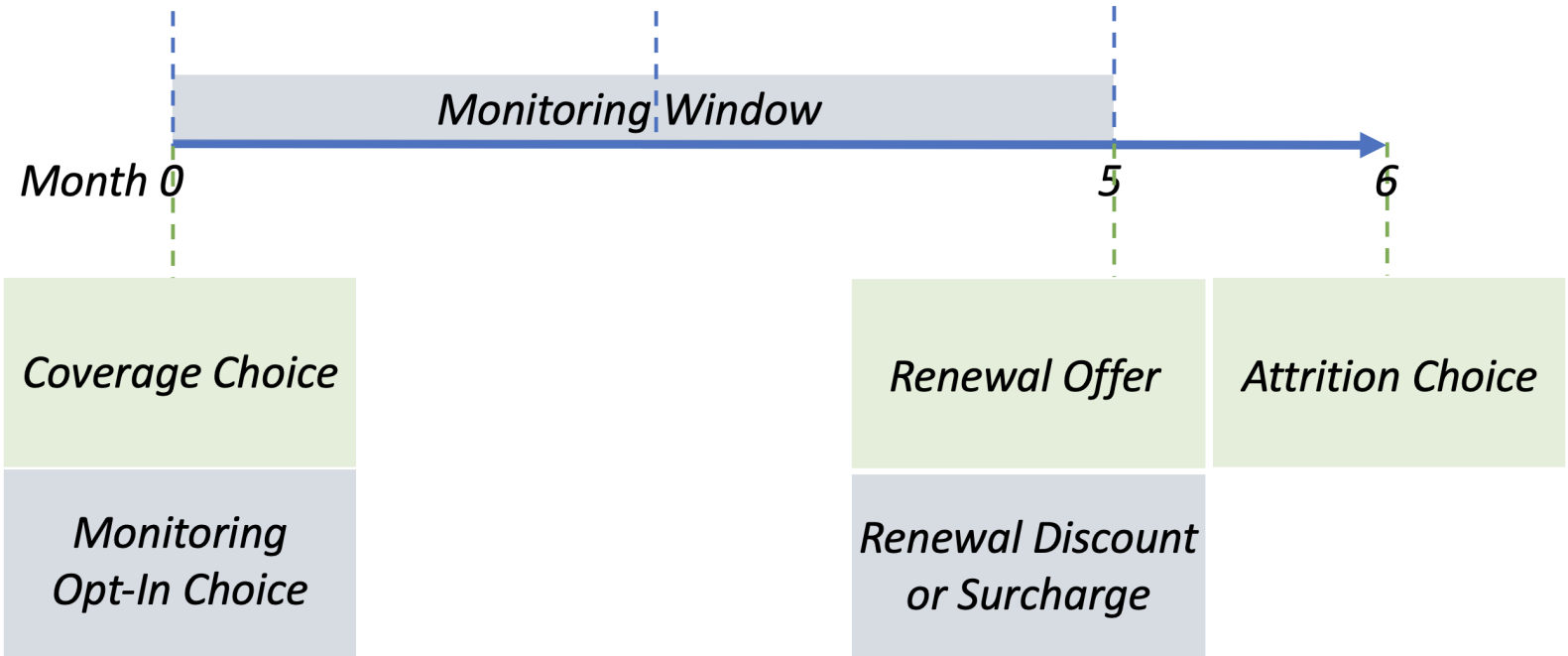
Monitoring Program

- *Punishable behavior*
- *Duration*
- *Opt-in discount*
- *Renewal discount range*
- ***Real-time feedback**:* beeping, web, texts, apps



Monitoring Program

- *Punishable 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).

- opt-in drivers are safer in the first period (monitored) than after, not opt-out drivers.

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

1. Key parameters and Identification
2. Results and No Monitoring Counterfactual

Pricing and Equilibrium

Structural Model - Overview why

- **Cost** model - claim count C details
  - Consumers have latent risk types (λ); unobserved heterogeneity (σ_λ)
  - Consumers can change λ by ($\theta_{\lambda,m}$) when monitored

Structural Model - Overview why

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

Structural Model - Overview why

- **Cost** model - claim count C details
  - Consumers have latent risk types (λ); unobserved heterogeneity (σ_λ)
  - Consumers can change λ by ($\theta_{\lambda,m}$) when monitored
- Monitoring **technology** - monitoring score s details
  - s is an informative signal of λ with precision (σ_s)
- **Choice** model d demand

  Product choices - firm f and coverage y
  - risk $\leftarrow \lambda$
  - preference: risk aversion (γ) and inertia for switching firms (η)

Structural Model - Overview why

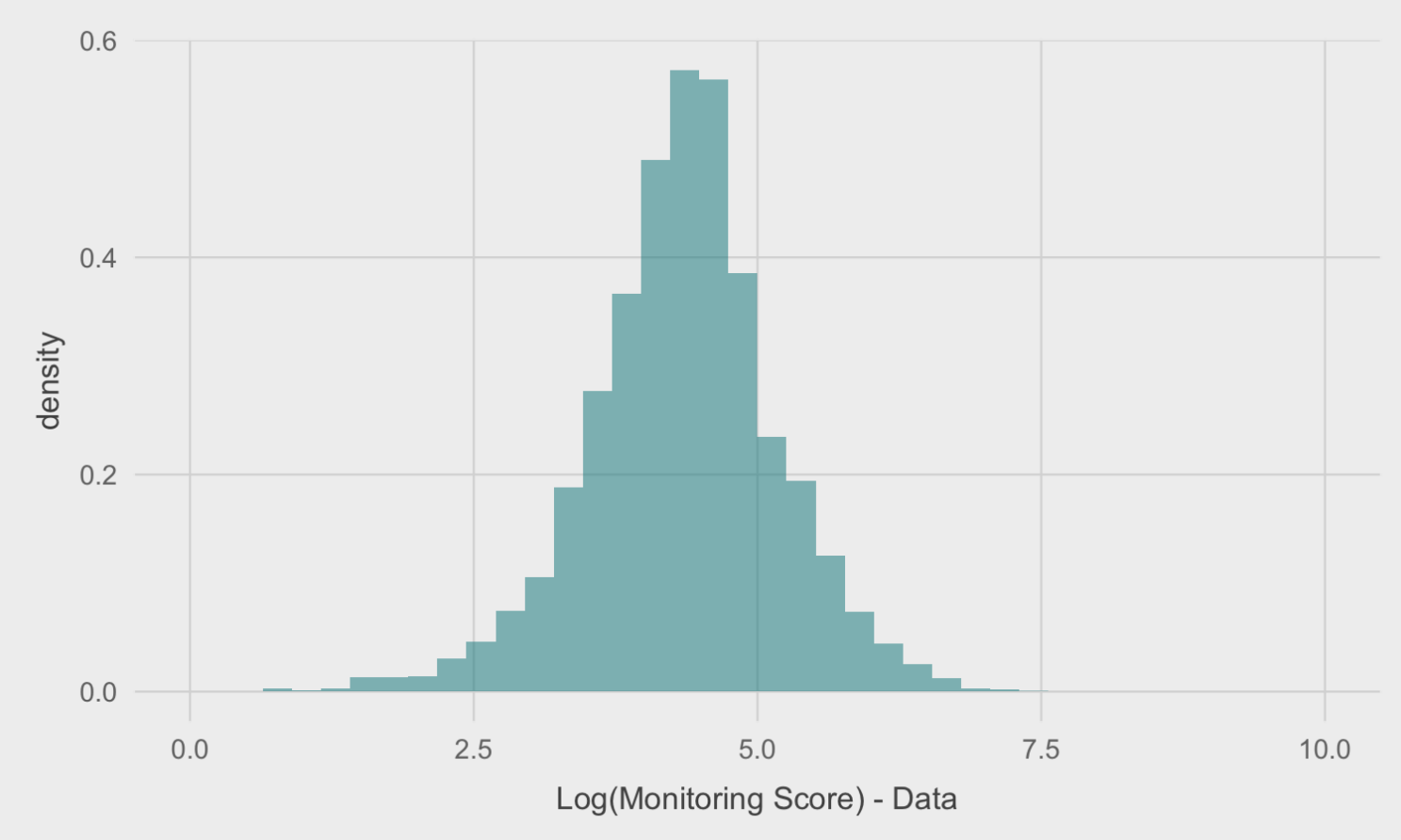
- **Cost** model - claim count C details
  - Consumers have latent risk types (λ); unobserved heterogeneity (σ_λ)
  - Consumers can change λ by ($\theta_{\lambda,m}$) when monitored
- Monitoring **technology** - monitoring score s details
  - s is an informative signal of λ with precision (σ_s)
- **Choice** model $d = \{f, y, m\}$ demand

  Product choices - firm f and coverage y
  - 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 $\leftarrow \theta_{\lambda,m}$
    - renewal discount and reclassification risk $\leftarrow \lambda$, γ, σ_s
  - unobserved disutility from being monitored (ξ) id

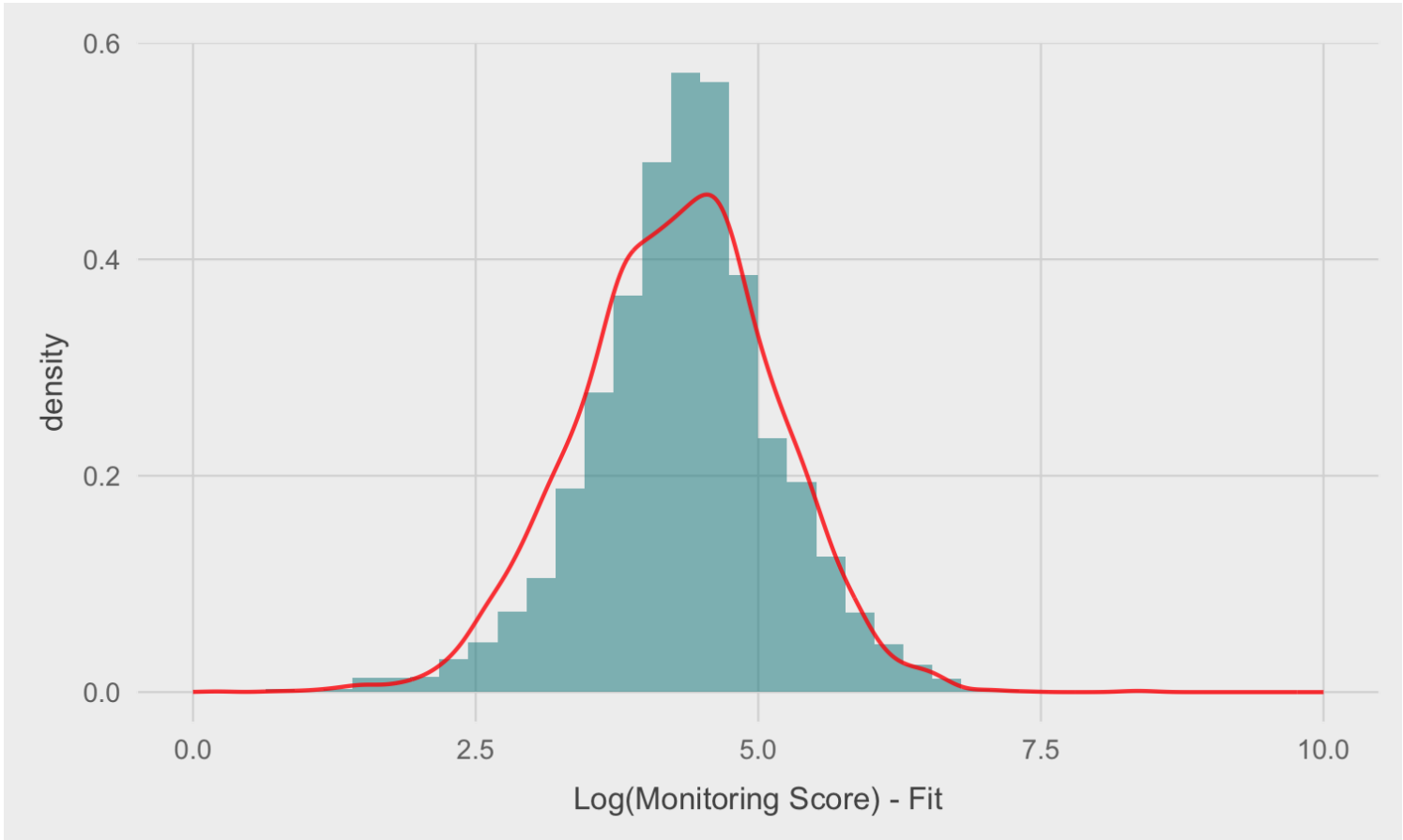
Estimation

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



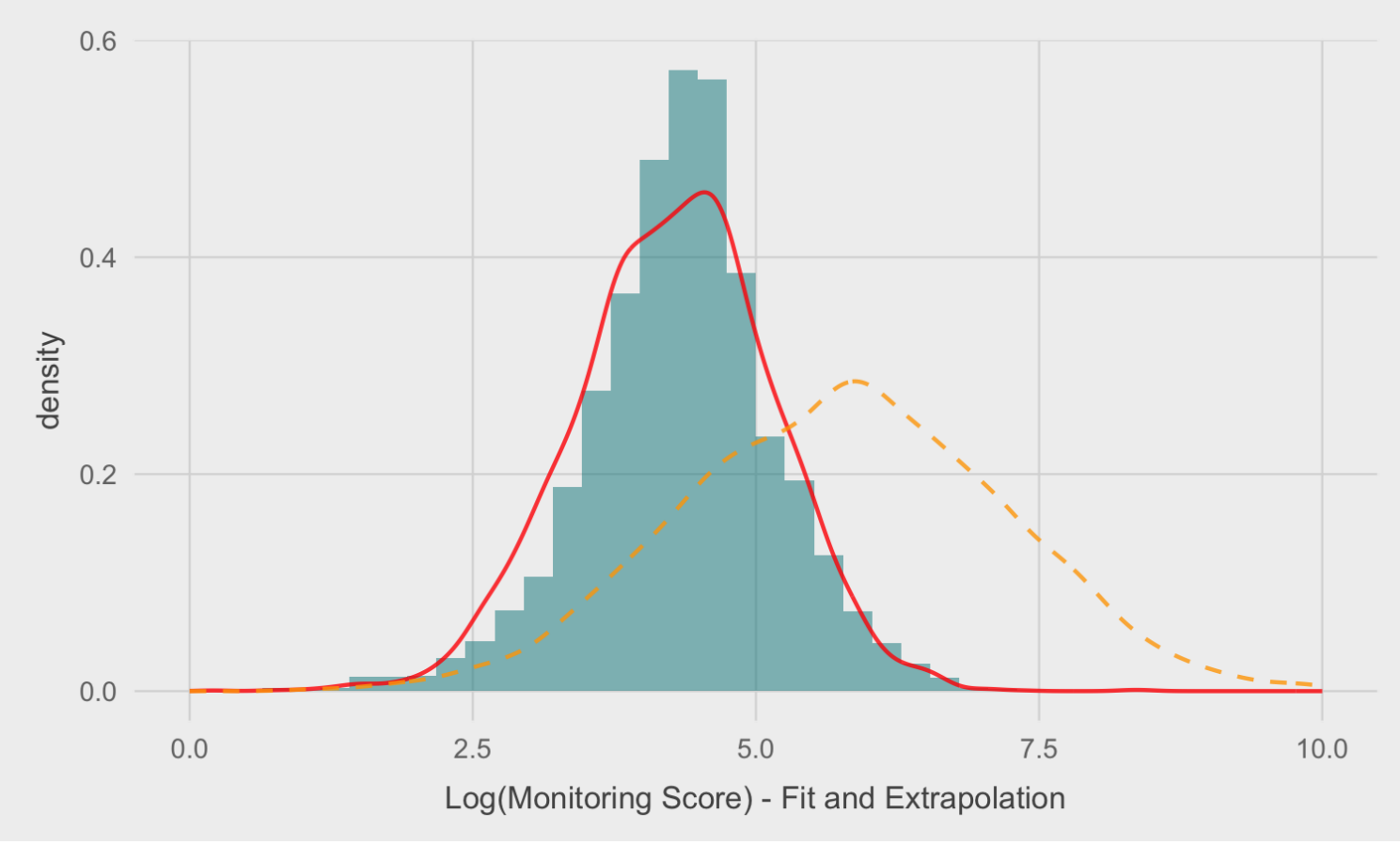
Fit

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

Fit and Cross Validation

estimation | fit | model detail | d-result

Goal: jointly several demand margins: coverage, attrition (firm), and monitoring

	Model Fit		Cross Validation
	Fit	Data	
Monitoring share (when eligible)	15.6%	15.3%	
Expected score	4.25	4.30	
Coverage share			
30K	12.5%	12.7%	
40K	8.2%	8.5%	
50K	49.8%	47.1%	
100K	15.4%	17.0%	
300K	11.9%	12.3%	
500K	2.3%	2.4%	
First renewal attrition	15.6%	15.2%	

- Drivers are not very risk averse: $\gamma = 9.7 \times 10^{-5***}$ bet | model

Fit and Cross Validation

estimation | fit | model detail | d-result

Goal: jointly several demand margins: coverage, attrition (firm), and monitoring

	Model Fit		Cross Validation
	Fit	**Data**	
Monitoring share (when eligible)	15.6%	15.3%	
Expected score	4.25	4.30	
Coverage share			
30K	12.5%	12.7%	
40K	8.2%	8.5%	
50K	49.8%	47.1%	
100K	15.4%	17.0%	
300K	11.9%	12.3%	
500K	2.3%	2.4%	
First renewal attrition	15.6%	15.2%	

- Drivers are not very risk averse: $\gamma = 9.7 \times 10^{-5***}$ bet | model
- large firm-switching inertia: mean $\hat{\eta}$ is \$283 (44% of premium)

Fit and Cross Validation

estimation | fit | model detail | d-result

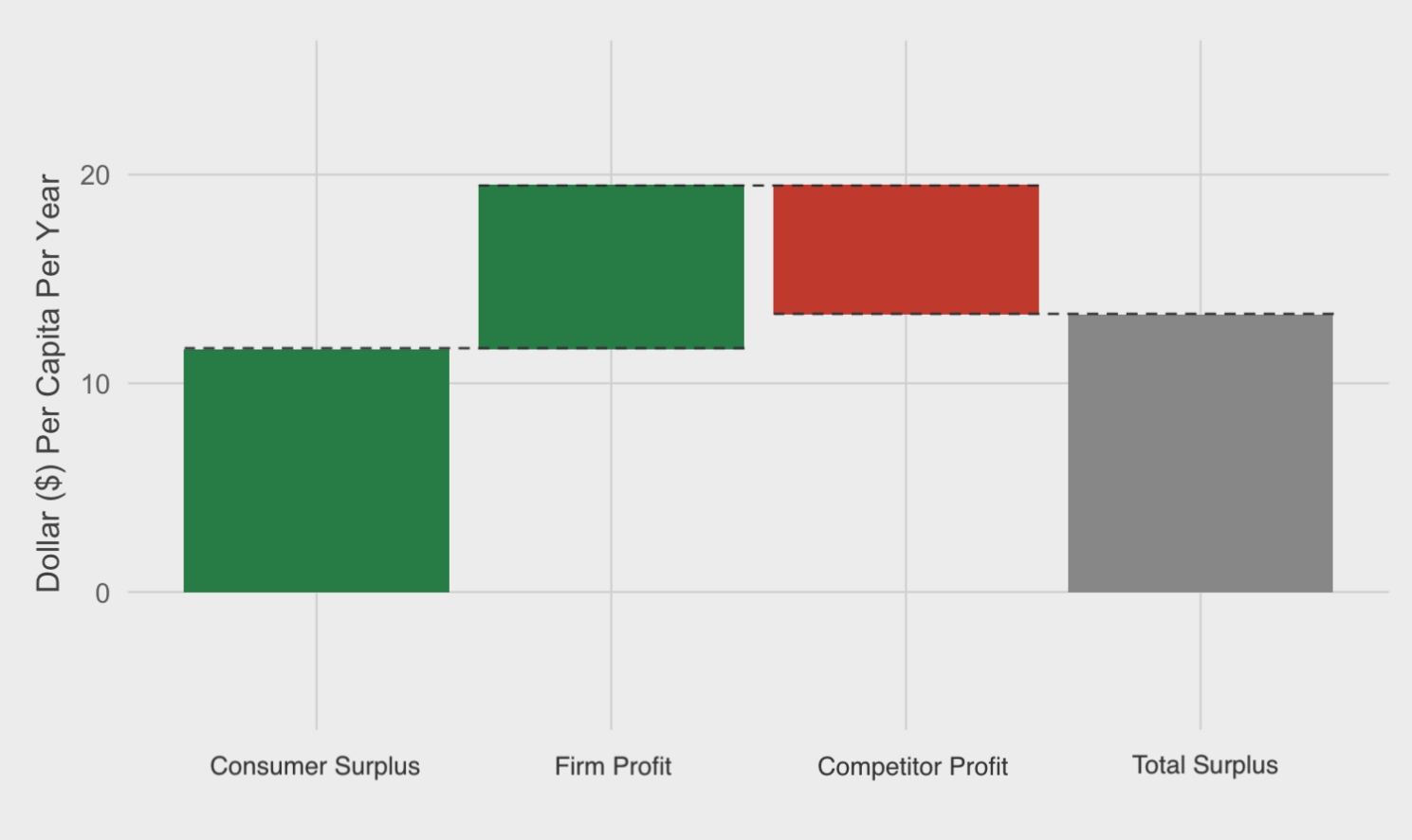
Goal: jointly several demand margins: coverage, attrition (firm), and monitoring

	Model Fit		Cross Validation	
	Fit	Data	Prediction	Hold-out Data
Monitoring share (when eligible)	15.6%	15.3%	17.9%	17.6%
Expected score	4.25	4.30	3.97	4.17
Coverage share				
30K	12.5%	12.7%	-	-
40K	8.2%	8.5%	7.6%	7.2%
50K	49.8%	47.1%	60.5%	58.1%
100K	15.4%	17.0%	17.5%	19.6%
300K	11.9%	12.3%	10.9%	12.8%
500K	2.3%	2.4%	3.6%	2.4%
First renewal attrition	15.6%	15.2%	15.4%	14.7%

- Drivers are not very risk averse: $\gamma = 9.7 \times 10^{-5***}$ bet | model
- large firm-switching inertia: mean $\hat{\eta}$ is \$283 (44% of premium)

Welfare Calculation: Current - No Monitoring

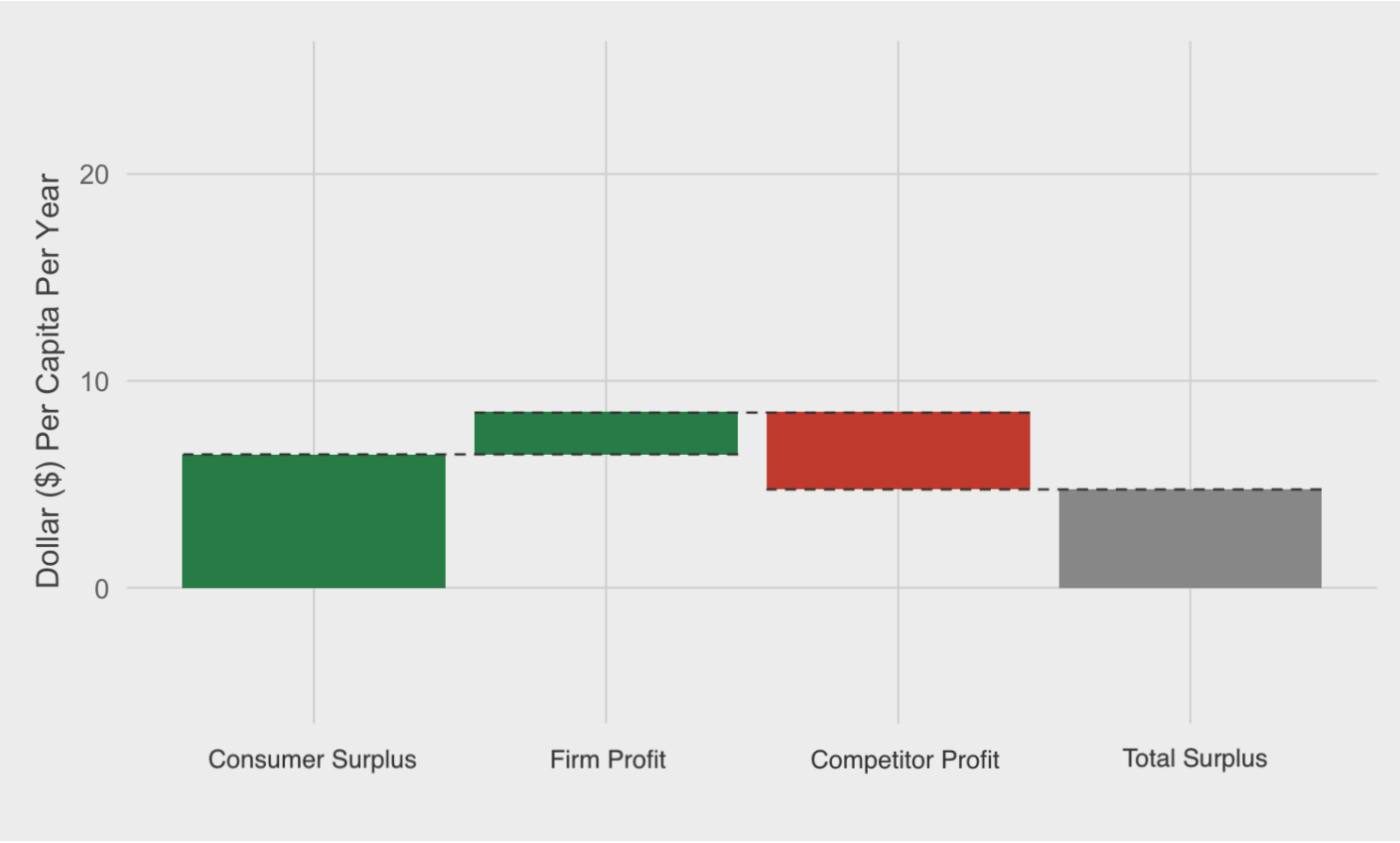
Introducing monitoring increases firm profit, consumer welfare, and total surplus. box Q



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

Welfare Calculation: Tease Out Allocative Effect

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



- ∼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

1. Pricing Model and Optimal Pricing
2. Equilibrium Counterfactual with Data Sharing

Pricing and Counterfactual Equilibria

1. Optimal Pricing

- Firm profit: two-period (pre-post information revelation) two-product (monitored vs. unmonitored pools) detail

- Three types of price *adjustments* corresponding to how monitoring changes the firm's *information set* detail
  $t = 0, m = 0 : \kappa_0$ 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

- Hold fixed competitor prices

Pricing and Counterfactual Equilibria

1. **Optimal Pricing**
   - ▷ Firm profit: two-period (pre-post information revelation) two-product (monitored vs. unmonitored pools) detail
   - ▷ Three types of price adjustments corresponding to how monitoring changes the firm's information set detail
     $t = 0, m = 0 : \kappa_0$ 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
   - ▷ Hold fixed competitor prices

2. **Data Sharing**
   Eliminates proprietary data but turns monitoring into a *public good*
   - ▷ Assume competitors have symmetric belief as the firm, and set a single alternative rent-sharing parameter for monitored drivers $\kappa_{s,-f^\star}$
   - ▷ Best response algorithm to calculate Nash Equilibrium result

Optional 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 $\kappa_{s,-f^\star}$	-	-

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

Optional 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