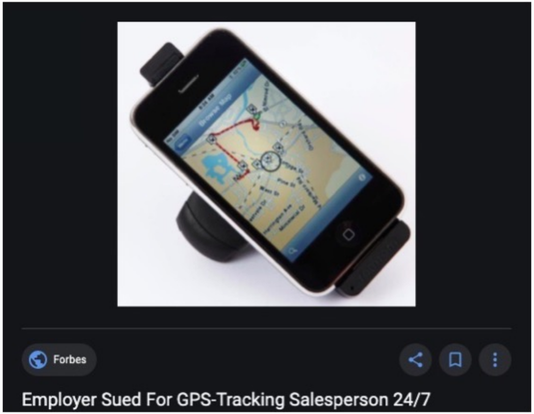
Valuing Intrinsic and Instrumental Preferences for Privacy

Tesary Lin

Boston University

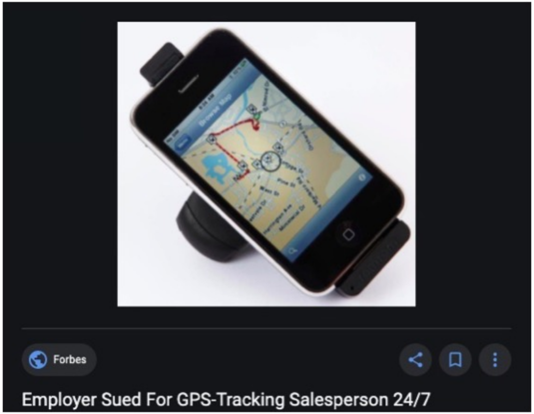
July 17, 2020

What Are Intrinsic & Instrumental Preferences for Privacy?



Intrinsic: taste, right

What Are Intrinsic & Instrumental Preferences for Privacy?



Intrinsic: taste, right

Instrumental: expected economic outcome from revealing one's type

Why Empirically Separate the Two?

1. Intrinsic & instrumental preferences induce different selection patterns

Instrumental preference only:

- Consumers who do not share $\Longrightarrow$ "low type"

  e.g. risky drivers can be more concerned about revealing their private info

Why Empirically Separate the Two?

1. Intrinsic & instrumental preferences induce different selection patterns

Heterogeneous intrinsic + *instrumental preference:*

- Consumers who do not share $\not\Rightarrow$ "low type"

  e.g. if safer drivers intrinsically dislike sharing location info more

Why Empirically Separate the Two?

1. Intrinsic & instrumental preferences induce different selection patterns

Heterogeneous intrinsic + instrumental preference:

- Consumers who do not share $\not\Rightarrow$ "low type"

  e.g. if safer drivers intrinsically dislike sharing location info more

2. Intrinsic—utility primitive; instrumental—endogenous

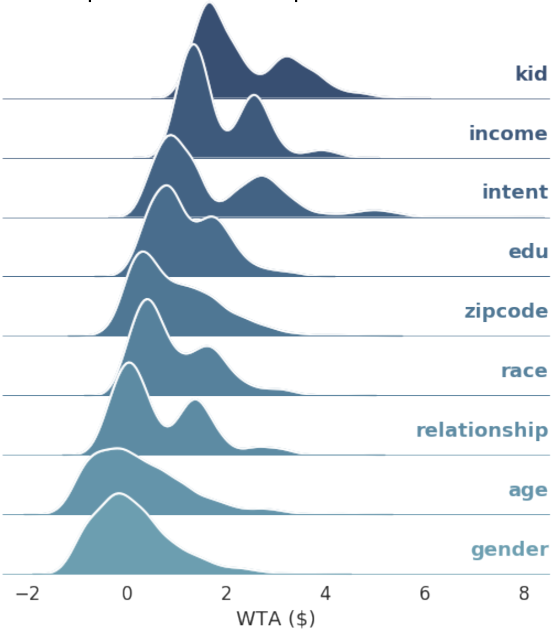
- Instrumental preferences respond to changes in firm's data collection & usage practices, e.g. due to new regulation

This Paper

1. Use an experiment to separately measure intrinsic & instrumental preferences
   - Revealed preference, in dollar terms; heterogeneity across demographics
   - Structurally estimate *intrinsic preference* & *belief on instrumental outcome* as primitives
2. Demonstrate the empirical selection pattern driven by the coexistence of the two preference types
3. Evaluate methods for firms & researchers to address privacy-induced selection

Result 1: Intrinsic Preferences are Highly Heterogeneous

WTA distribution across personal data requested (Y-axis) and consumers (X-axis)



Result 2: Instrumental Preference Matches Actual Outcome

- Consumer belief on the instrumental outcome determines the magnitude of instrumental preference
  - E.g. if risky drivers are unaware that firm uses driving data to customize premium, then instrumental preference = 0
- Estimation result shows that consumer beliefs are consistent with actual payoff qualitatively & quantitatively

Result 2: Instrumental Preference Matches Actual Outcome

- Consumer belief on the instrumental outcome determines the magnitude of instrumental preference
  - E.g. if risky drivers are unaware that firm uses driving data to customize premium, then instrumental preference = 0
- Estimation result shows that consumer beliefs are consistent with actual payoff qualitatively & quantitatively
  - Actual payment: w = 2 vs. Consumer belief estimates:

Model	1. No Heterogeneity		2. Heterogeneous c		3. Heterogeneous c & δ		4. Heterogeneous c & β	
	mean	95% CI	mean	95% CI	mean	95% CI	mean	95% CI
w_{income}	2.00	[0.15, 3.87]	2.12	[0.11, 3.99]	2.02	[0.14, 3.92]	1.90	[0.04, 3.88]
w_{intent}	2.63	[1.07, 3.88]	1.94	[0.38, 3.76]	1.97	[0.29, 3.77]	1.90	[0.35, 3.70]

Result 3: Intrinsic & Instrumental Jointly Determines Selection Pattern

Classical prediction: *low types are more willing to hide*

Result shows two opposite cases (for different personal data requested)

1. Classical prediction rejected
   - Reason: high types have higher intrinsic preferences; magnitude dominates instrumental
2. Classical prediction confirmed
   - Reason: intrinsic preference heterogeneity independent of consumer type

Result 3: Intrinsic & Instrumental Jointly Determines Selection Pattern

Classical prediction: *low types are more willing to hide*

Result shows two opposite cases (for different personal data requested)

1. Classical prediction rejected
   - Reason: high types have higher intrinsic preferences; magnitude dominates instrumental
2. Classical prediction confirmed
   - Reason: intrinsic preference heterogeneity independent of consumer type

Takeaway: need to measure heterogeneity & correlation between the two preference components to understand the empirical selection pattern

Implication

Separating intrinsic & instrumental preferences for privacy can help us

1. Measure privacy preferences by understanding how much they respond **endogenously** to ways that the firm uses data
2. Improve methods to collect & analyze consumer data by understanding its **selection** pattern

Experiment Design

Stage 1: Collect Full Data

1

2

Participants complete survey sent by UChicago

- Smartwatch preference questions (camouflage)
- Personal questions $\longrightarrow$ contents of personal data
  - Gender, age, edu, income, relationship, number of children, zipcode, race, digital product preference

Stage 2: Collect Privacy Choices & Shared Data

Participants receive data sharing request

Treatments (factorial):

- Compensation (price for data):
  Gift card value: {\$0; \$10, \$20, \$50} × 1%
- Instrumental Incentive:
  {on, off}

Treatment Variation to Identify Model Primitives

U(protect data) - U(share data) = intrinsic utility + instrumental utility − compensation

Compensation (price for data):

- Same for each datapoint **regardless of what the firm learns about you**
- Translate privacy preferences to dollar terms

Instrumental Incentive:

- Payoff that **depends on what the firm learns about you based on data shared**
- Separate instrumental utility from intrinsic

Treatment Group Payment Scheme

Your winning probability is determined both by the baseline probability and by the adjustment terms. The baseline winning probability is calculated as follows:

$$\text{Baseline probability of winning} = \text{Number of boxes checked} \times 1\%$$

This baseline probability is then adjusted to encourage response sharing from the customer group that Odde intends to serve, as shown in the following chart:

Income	< $50,000	$50,000 – $75,000	> $75,000
Adjustment	**-2%**	**Unchanged**	**+2%**
Plan to purchase any digital device in the next 3 months	Somewhat or extremely unlikely	Neither likely nor unlikely	Somewhat or extremely likely
Adjustment	**-2%**	**Unchanged**	**+2%**

Data

Sample Source and Size

Participant: Qualtrics Panel

- Typical source when firms estimate demand before product launch
- Lower bounds of population-level intrinsic preferences; alleviate the gap by
    - Stratified sampling using US census demographics
    - Characterize heterogeneity using observables

Sample Size:

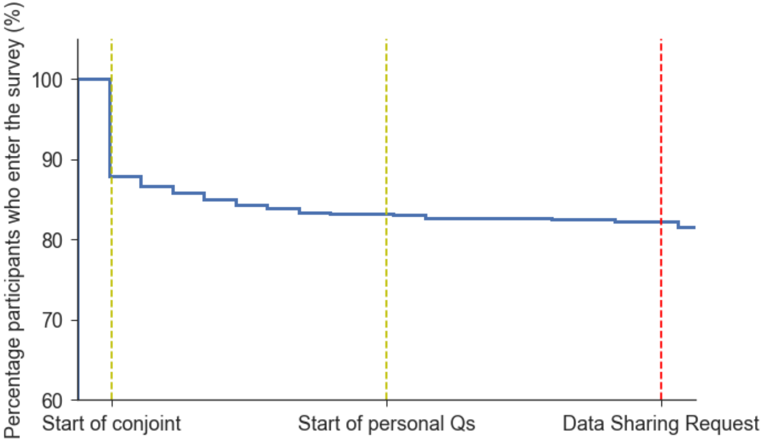
- 4,142 enter the survey; 2,583 qualified complete responses

Demographics of Final Sample

	Variables	Experiment Sample	2018 Census
	Female	65.31%	50.80%
	Married	47.39%	51.16%
	Have young kids	24.78%	–
	Mean age	47.60 (16.89)	45.9 (–)
Education	High school degree or less	47.00%	39.93%
	College degree	40.65%	48.67%
	Master's degree or higher	11.39%	11.40%
Race	White	71.27%	76.60%
	Black	15.37%	13.40%
Annual Household Income	$25,000 or less	21.99%	20.23%
	$25,000 to $50,000	29.54%	21.55%
	$50,000 to $100,000	30.12%	28.97%
	$100,000 or more	13.55%	29.25%
No. Observations		2,583	–

- More female, fewer with college degree, fewer in high-income bucket

Attrition Pattern



- Most attrition occurs at the start; not induced by concern about personal Qs

Model

Conceptual Model

Firm wants: personal data $\longrightarrow$ consumer's "type" $\longrightarrow$ targeted payoff

may offer (type-invariant) compensation to encourage data sharing

Consumer chooses protect vs. share data: protect iff

$$\text{U (protect) - U (share)} > 0 \iff \text{intrinsic utility}$$
$$+ \underbrace{\text{payoff by hiding his type - payoff from disclosing his type}}_{\text{instrumental utility}}$$
$$- \text{compensation} > 0$$

Estimation Model & Identification

U(not share k) − U(share k)

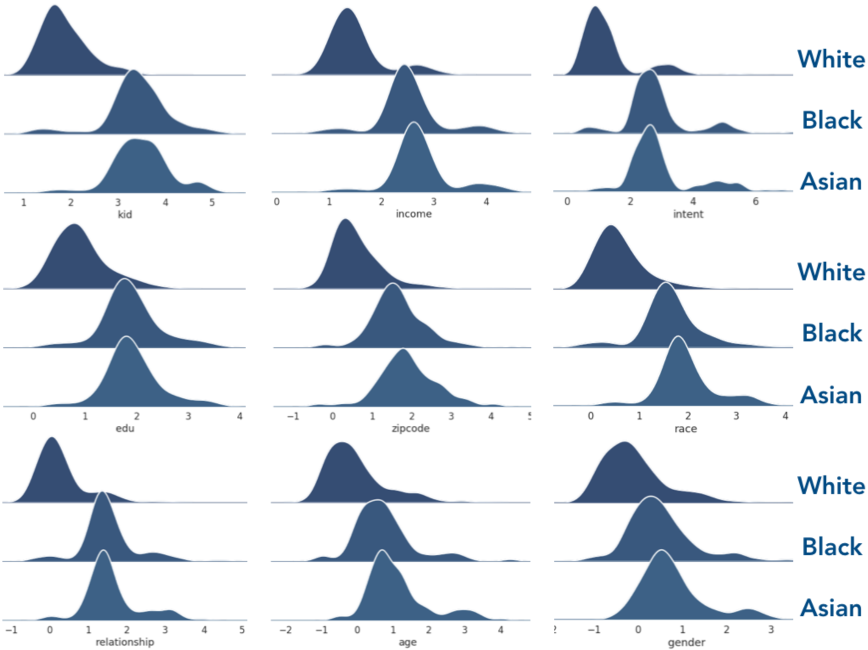
$$= \underbrace{c_k}_{\text{intrinsic}} + 1_{instr} \cdot 1_{k\in\{1,2\}} \cdot \underbrace{\beta \cdot p \cdot w_k \cdot \left(\widehat{E}[d_k|s_k = 0] - \widehat{E}[d_k|s_k = 1, d_{ik}]\right)}_{\text{instrumental}} - \underbrace{\beta \cdot p}_{\text{compensation}} + \epsilon_{ik}$$

Results converted to dollar space to address scale invariance problem

- c_k: utility intercept in the control group
- β: response to different amounts of compensation
- w_k: how different types react differently to instrumental incentives
- δ_{k0}, δ_{k1}: response to instrumental incentives that is common across types

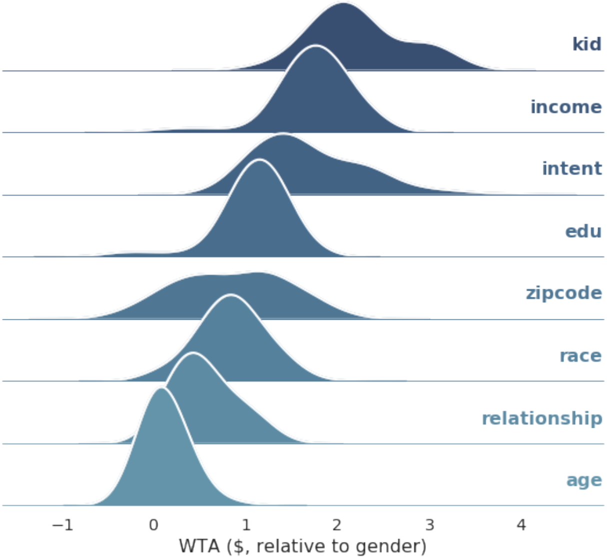
Intrinsic Preferences

Intrinsic Preference: Non-Whites Higher than Whites



Intrinsic Preference: Heterogeneous Even within Individual

Dollar Value of Intrinsic Preferences Relative to Individual Preference to Protect Gender Info



Experiment Replication

Replicate Experiment in the Field

- A treatment where consumers are given $$$ to get "ground truth"
- Separate intrinsic & instrumental in other treatments—Challenge: instrumental incentive hard to be removed
    - Vary intensity of instrumental incentive to measure consumer belief; project choices to 0 instrumental case, assuming belief stays constant