Comparing Hypothetical and Realistic Privacy Valuations

Joshua Tan, **Mahmood Sharif**, Sruti Bhagavatula, Matthias Beckerle, Michelle L. Mazurek*, Lujo Bauer

> Carnegie Mellon

University

PRIVACYCON

* UNIVERSITY OF MARYLAND

Why measure privacy preferences?

- Privacy preferences = willingness/comfort sharing personal info
- Who benefits from understanding privacy preferences?
 - System designers
 - What data are users okay sharing?
 - How much value should users receive for sharing?
 - Policy makers

- How much "loss" do consumers incur through data breaches?
- What kind of data sharing (if any) should be disincentivized?

Measuring privacy preferences is challenging

- Contextual factors influence users' privacy preferences and behaviors
 - E.g., willingness to share PII depends on how it will be used
- Valuations of goods (estimations of worth) influenced by framing effects and cognitive biases
 - Hypothetical bias = overestimate value in hypothetical scenario
- Stated privacy attitudes often do not align with actual behavior (privacy paradox)



This talk: Can we predict privacy valuations?

- Privacy valuation = willingness to sell and selling price for personal info
- How do privacy valuations depend on combinations of factors?

Attribute type



PRIVACYCON

Receiving party



Scenario realism



• Does hypothetical bias explain the privacy paradox?

Methodology

- Online study with 434 Prolific participants
- Participants asked to assign selling prices to personal attributes
 - Could also choose to not sell
 - Selling scenario was information marketplace operated by CMU
 - Attributes in market are sold to buyers via an auction
 - Buyers have limited budgets and purchase lowest-priced offers first



Prices assigned to 7 attributes and 6 parties

For how much do you agree to sell your (attribute) to each one of the following parties?

• Age

- Email address
- Gender
- Relationship status
- Home address
- Occupation
- Phone number

C	ľ	70	0	С	e)	

```
Sell Do not sell
```

\$ ar	n	0	U	r	It	

- Ad networks
- Federal agencies
- Insurance companies
- Market research companies
- Political parties
- Research pools

We varied the realism of the scenario

More realistic



G Sign in with Google

Contact info sold for more \$



Selling price depends on who is buying

\$\$\$\$\$\$ **\$\$** Ad networks Market research **Research** pools companies **Federal agencies Political parties** COMPETITORS VENDORS CUSTOMER Insurance companies

Privacy paradox doesn't always hold

- *Hypothetical* values not generally different than *Realistic* values
 - Exceptions:

- Phone number (~\$9 vs. ~\$14)
- Home address (~\$8 vs. ~\$11)
- Calibration factor = Hypothetical / Real
 - Largest calibration factor in our study was 1.61
 - List and Gallet (2001): 4.44 for public goods, 8.41 for private goods
- No significant differences in likelihood of selling by scenario realism

Can we predict valuations?

• From scenario realism, attribute type, and receiving party

Dollar values?

- Not yet, individual users have very different baselines
- But, given baseline, accurate
 \$ prediction possible

PRIVACYCON

Attribute rankings?

- Yes, Same average rankings regardless of scenario realism
- Eliciting subset rankings further improves predictions

Takeaways

- Certain privacy preferences are possible to predict
- In contrast to other types of goods, privacy valuations not generally affected by hypothetical bias
- Attribute rankings stable regardless of scenario realism and receiving party
- Selling prices can be accurately predicted based on attribute type and receiving party, given baseline price for individual person

Joshua Tan, **Mahmood Sharif**, Sruti Bhagavatula, Matthias Beckerle, Michelle L. Mazurek*, Lujo Bauer



