

Discussion of:
“Balancing User Privacy and Personalization”
By Korganbekova and Zuber

Discussion by Scott Shriver
University of Colorado
October 18, 2024

What this paper does

- What: Empirically measures value of historical consumer/firm clickstream (product search, purchase) data
 - To firms: impact on revenues/profits
 - To consumers: impact on welfare \equiv product utility - search costs
- How: Field experiment + structural model & counterfactuals
- Why: Informs firm/public policy w.r.t. consumer privacy
 - Economic trade-offs for using/not using personal information



What I like

- Very nice paper that addresses an important & timely issue
 - Relevant to firms and policy makers
- Field experiment data
 - Reduced endogeneity concerns (randomized, but not fully controlled)
 - Incredibly rich – complete clickstream observed, including scrolling
 - Product image data also captured and encoded
- Structural model
 - Calibrated using experimental variation
 - State-of-art modeling as Gaussian process

Field experiment – wayfair.com

- Outcomes: clicks, add-to-cart, purchases, revenues, profits, etc.
 - Dining chair product category, observed for 2 years (2020-2021)
- Treatment/control: session product rank page personalized/not using prior session data
 - Prior session data is individual-specific and time-varying (SUTVA?)
- Results:

TABLE 4: Effect of personalization on consumer and platform outcomes

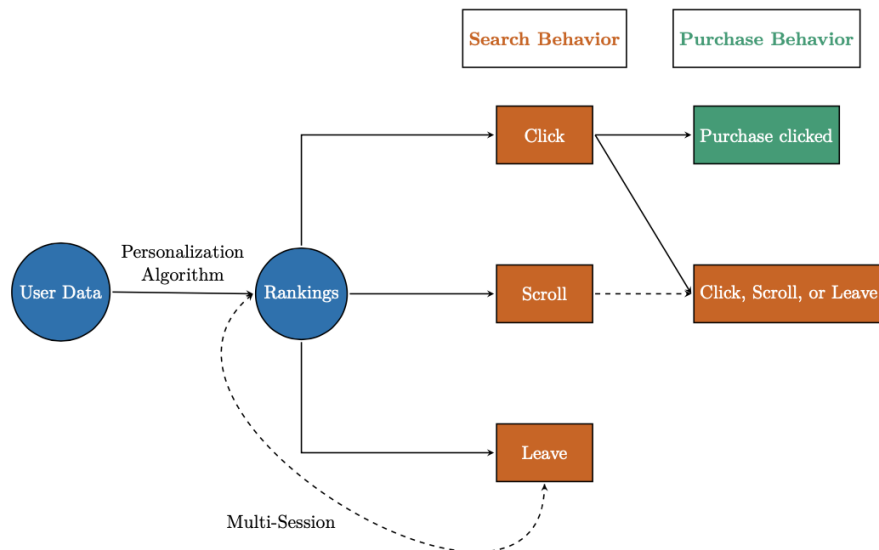
	<i>Logistic</i>				<i>OLS</i>		
	(1) Clicks	(2) Add-to-cart	(3) Basket page	(4) Converted	(5) Log(Revenue)	(6) Purchases	(7) Log(Profit)
Personalized	0.002 (0.012)	0.011** (0.005)	0.014*** (0.005)	0.014** (0.005)	0.021*** (0.008)	0.024*** (0.008)	0.015** (0.006)
Intercept	2.988*** (0.008)	0.246*** (0.004)	0.148*** (0.004)	-0.870*** (0.004)	1.947*** (0.005)	1.095*** (0.006)	– (0.004)
Observations	635,267	635,267	635,267	635,267	635,267	635,267	635,267

Key outcomes
increase ~1-2%,
except clicks

Structural model

- Modeled actions: search (click, scroll, leave), purchase
 - Heuristic (linear index, “near-optimal”) search assumed (no Bellman equation)
 - Product utility a Gaussian process, value uncertainty reduced by clicking (Bayesian updating)

FIGURE 13: Overview of the Model



Product utility

$$u_{ij} = m_i(X_j) + \xi_j + \varepsilon_{ij} \quad \xi_j \sim N(0, \sigma_\xi^2) \text{ and } \varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2).$$

$$m_i(X_j) \sim GP(\beta_i X_j, \kappa(X_j, X_{j-1}))$$

$$\beta_i \sim N(\beta, \Omega) \quad X_j = [\text{price}_j, \text{rating}_j, \#\text{ratings}_j, \text{image}_j]$$

Click, scroll costs

$$c_{ijt} = c_0 + \underbrace{\psi_{ijt}}_{\text{Type 1 EV}}$$

Welfare: different utility scales?

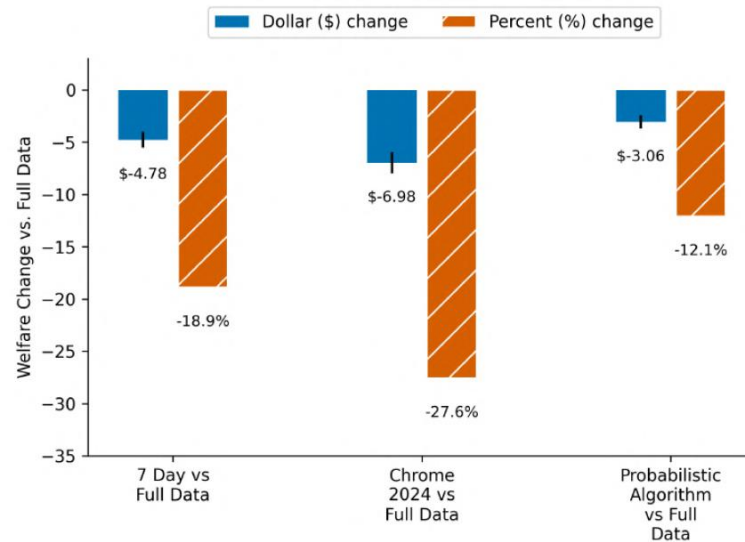
$$c_s(r_t) = c_s \cdot \log(r_t)$$

Deterministic? Functional form?

Counterfactuals

- Assess impact of browser-imposed privacy policies & potential work-arounds
 - Expiring first-party cookies after 7 days (Safari 2019)
 - Blocking third-party cookies (Chrome 2024)
 - Privacy restriction mitigation: Probabilistic Identity Recognition (vs. known)

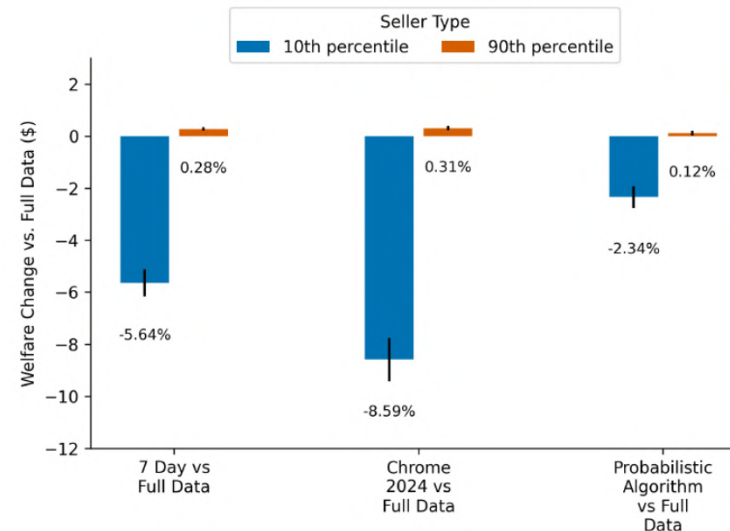
FIGURE 20: Counterfactual results: consumer welfare



Large (20-30%) welfare losses under policies 1 & 2

Moderate (~10%) welfare losses under policy 3

FIGURE 23: Counterfactual results: seller outcomes



Moderate (5-10%) profit losses for small products/sellers under policies 1 & 2

Small (<5%) profit losses under policy 3

Counterfactuals: Comments

- 20%-30% welfare losses under counterfactuals 1 & 2 seem large
 - Artifact of data? Modeling (scroll cost, utility scales)?
- Probabilistic recognition algorithm (counterfactual 3)
 - Interesting and promising, but no silver bullet
 - Highly predictive algorithm \Leftrightarrow No effective privacy
 - Acceptable probability threshold (u.b.) for identifiability?

Conclusion

- Extremely rich & novel data, applied to policy-relevant question
- Rigorous & novel methods
- Desirable improvements:
 - Effect robustness – model assumptions, functional forms, etc.
 - More streamlined narrative/exposition in paper
- Thank you for the opportunity to read such an interesting paper!