RSAConference2019
San Francisco | March 4–8 | Moscone Center

SESSION ID: SEM-M01

How to Infer Actual Privacy Concern From Online Behavior

Laura Brandimarte

Assistant Professor of MIS
University of Arizona
@thefirstred

The Privacy Paradox

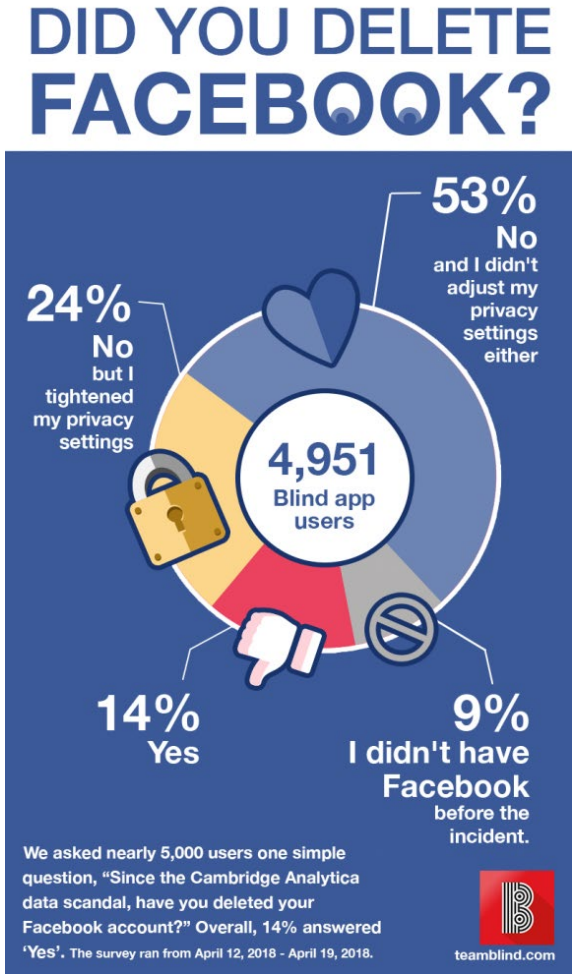
Conclusion

- People do not really care about their online privacy

Many flaws in that conclusion

- Asymmetric information
  - Average Internet user cannot have a clear picture of all the consequences of online behaviors
- Cognitive/behavioral/emotional biases
  - People are not utility-maximizing machines
- Risk perception
  - When there are benefits, risks are perceived as lower
- …
- Are we even using the right metrics?

Example



Only 14% deleted Facebook, so they don't really care about privacy

What if we asked the following question instead?

- Did people change their posting behavior on Facebook after the Cambridge Analytica scandal?

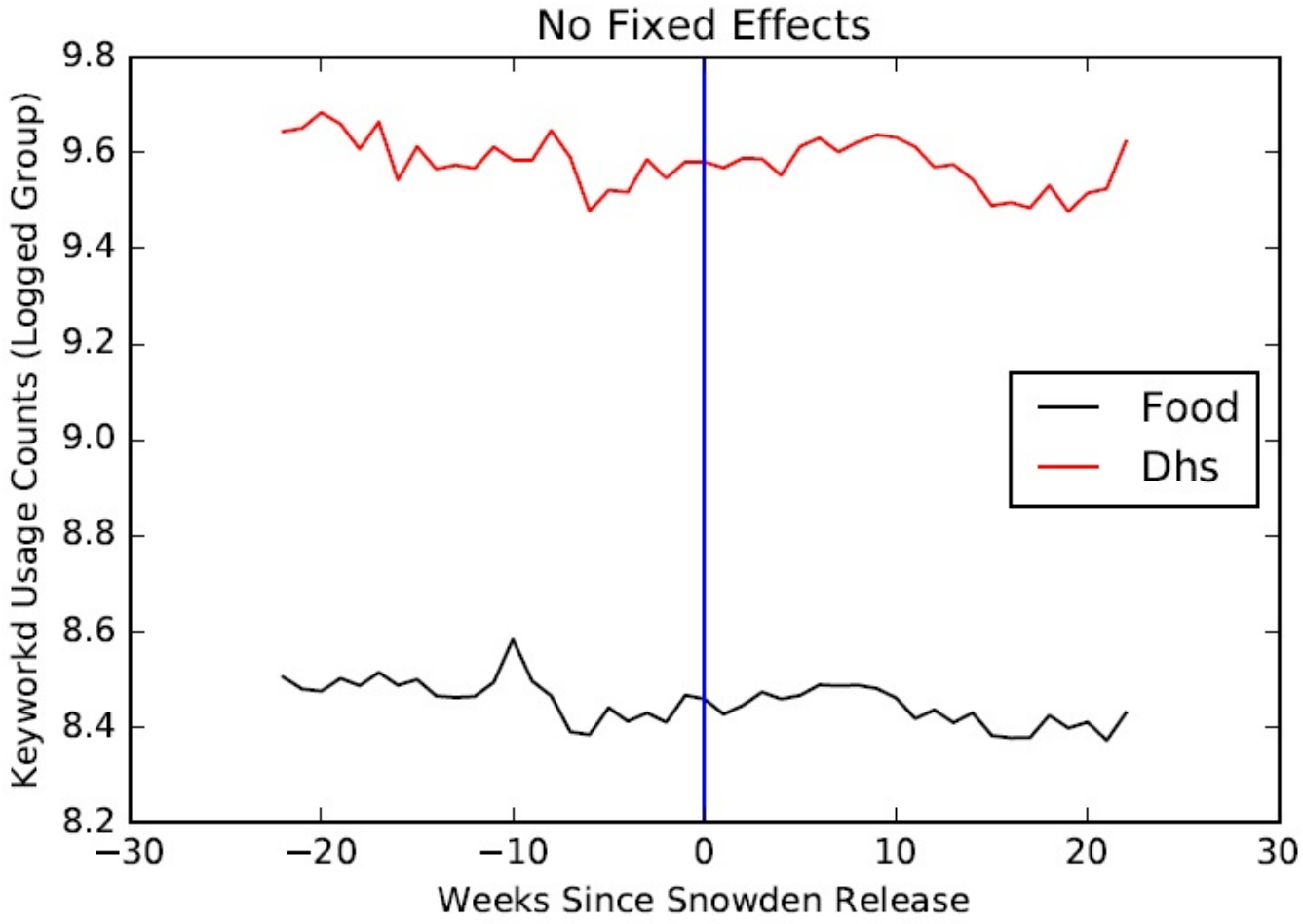
We asked a similar question for Twitter

- Did people change their use of sensitive* words on Twitter after the Snowden revelations about Government surveillance?

* that are monitored by the Department of Homeland Security or DHS (cybersecurity, terrorism, infrastructure security…)

Data

- 12 months (~80 TB) of data (18 bln Tweets)
- 22 weeks before and 29 weeks after June 6, 2013

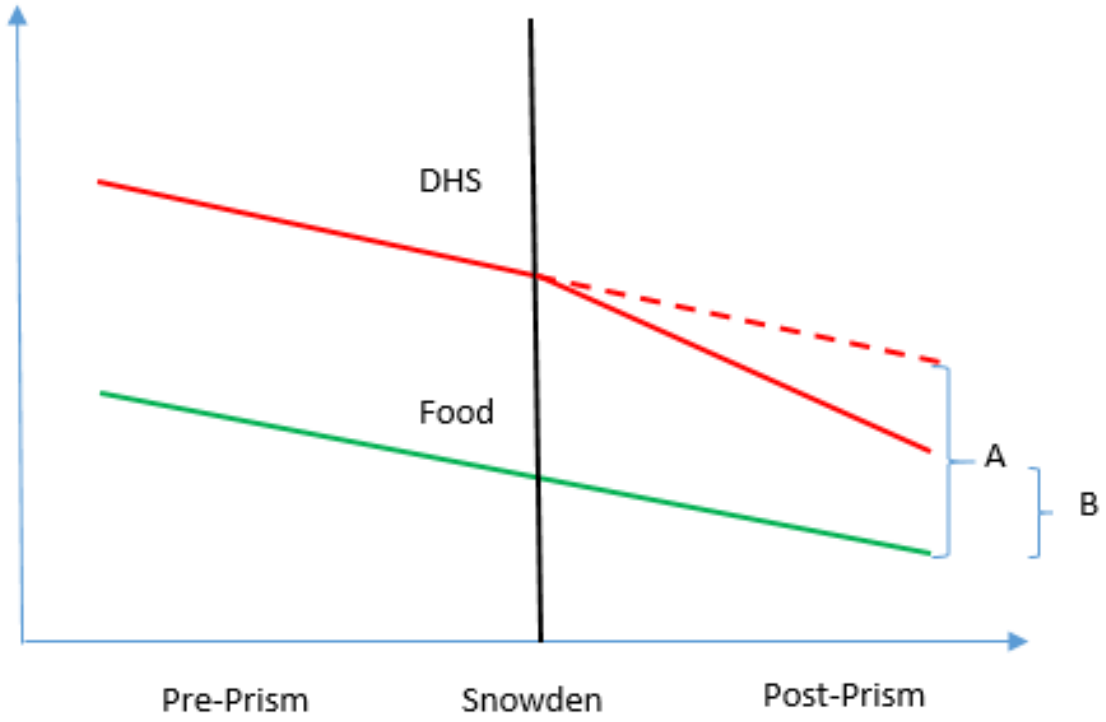


Methodology

- Combine machine learning with econometrics
  - ML algorithm finds anomalies in trends of word counts, and suggests how to solve problems caused by unmet assumptions of econometric model
  - Econometric analysis estimates size and significance of effect by comparing sensitive to non-sensitive (food-related) words, and establishes causality

Results

- Sensitive words decreased more than non-sensitive ones after Snowden revelations



Research questions – and answers

- Did awareness of Government surveillance programs affect the way people express themselves on Twitter?
  - Yes, self-censoring (chilling) effect of about 1%
- Where is the effect more pronounced?
  - Chilling appears to have a disproportionate effect on the US, “Blue” states

Conclusion

- Inferring privacy concern from online behavior is hard!
- There may be no privacy paradox at all depending on what metrics we use