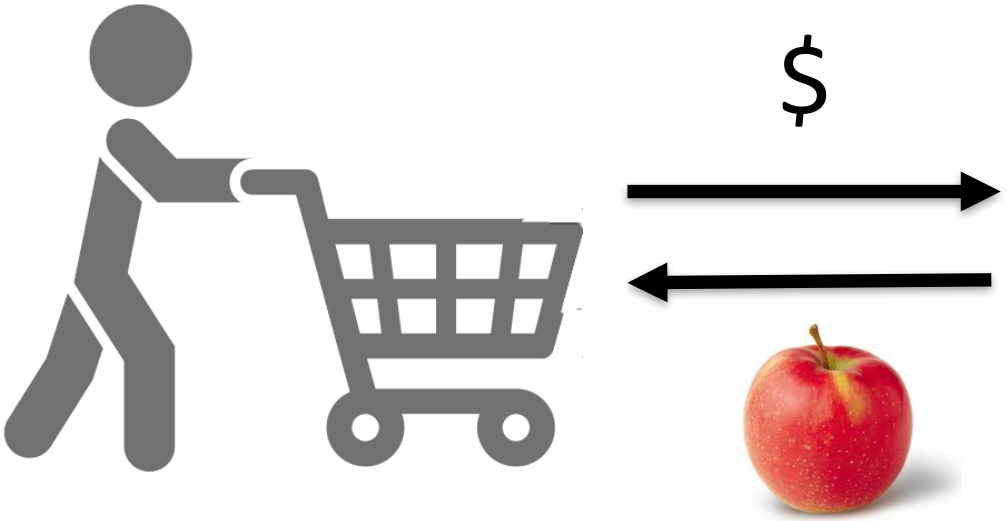
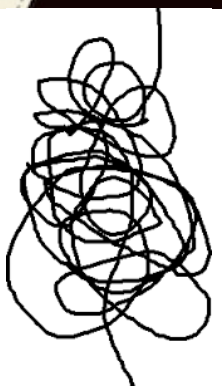
Discussion of Catherine Tucker "Artificial Intelligence and Privacy"

Ginger Zhe Jin
University of Maryland & NBER

Sept. 13, 2017

What is the problem?





Data persist
Data repurposed
Data spillover

- Intended use of AI:
  - Use your search history, emails, shopping history, social media usage to predict your "type"
  - Assume there is a stable "type" to learn about you

- "Unintended" use of AI:
  - Consumers do not know their types either
    ➔ discover or develop their types by following cues from AI
  - "Shadow" company creates fake accounts and buys social media ads
    ➔reach targeted groups via AI
  - Techy firms with AI in mind store enormous data
    ➔ become a target of hackers
  - Bad players such as Robocalls start to use AI
    ➔ Large volume is concerning even if they are not as smart as legitimate users of AI

Quality of AI algorithm

- Biased prediction if AI algorithm is imperfect
  - Magnitude and source of bias
  - Is it more biased than human's rule of thumb?
  - Can we correct the bias with more and better data?

- Things can also go wrong if AI algorithm is perfect
  - Perfect 1st degree price discrimination
    - Can competition address this problem?
  - Robust collusion among AI-driven competitors
    - Require humans have absolute confidence in AI

Should consumers give up privacy for the benefits of AI?

- Approach 1: Give consumers notice and choice
  - Hard to predict and evaluate all strings attached to a data flow in a focal transaction
  - Consumer preference for privacy is still developing
  - Little recourse to retract data or limit data use

How far can notice and choice go?

Should consumers give up privacy for the benefits of AI?

- Approach 2: Push firms to be transparent
  - About how they collect, store and use data
  - How to ensure authentic, complete and timely disclosure?
  - Then what? (who will use these information and how?)

Should consumers give up privacy for the benefits of AI?

- Approach 3: Direct regulation
  - E.g. minimum quality standard in data collection and data security
  - E.g. require opt-in or opt-out by type of information
  - How to set such a regulation, especially when there are horizontal preferences?
    - Is the Fair Credit Reporting Act a good example to follow?
  - How to ensure the regulation keeps up with technology?

How do privacy and data security assimilate or differ from safety regulations?

- Similarity
  - Consumers facing a severe, persistent information disadvantage
  - Firms may have incentives to hide, twist and obfuscate information transmission
  - A long chain between cause and consequence
  - Argument for policy actions before disaster happens
- Difference
  - Everything else equal, more safety is better for everyone
  - But preference on privacy can be horizontal and context-dependent
  - More uncertainty in the pros and cons of future AI for privacy?
- Integration: Self-driving cars

Detailed comments on Miller and Tucker (2017)

- Why compare to Census?
  - Facebook users are a selected group of population
  - Some Facebook users are more active than others
  - Advertisers aim to target active Facebook users, not the overall Census-represented population
- Facebook algorithm
  - Does Facebook know the “true” ethnicity of users and use it to validate AI algorithm?
  - How does Facebook react to this result? Is it a “bias” from their point of view?
- Counter-intelligence of Facebook users
  - How many take measures to hide their true ethnicity?
  - How does this affect the AI algorithm?