Online Tracking and Publishers Revenues: An Empirical Analysis

Work in progress

Veronica Marotta (University of Minnesota)
Vibhanshu Abhishek (UC Irvine)
Alessandro Acquisti (Carnegie Mellon University)

Acknowledgements: We acknowledge support from a grant by the Sloan Foundation. For other grants not related to this work, see https://www.heinz.cmu.edu/~acquisti/grants.htm

PRIVACYCON

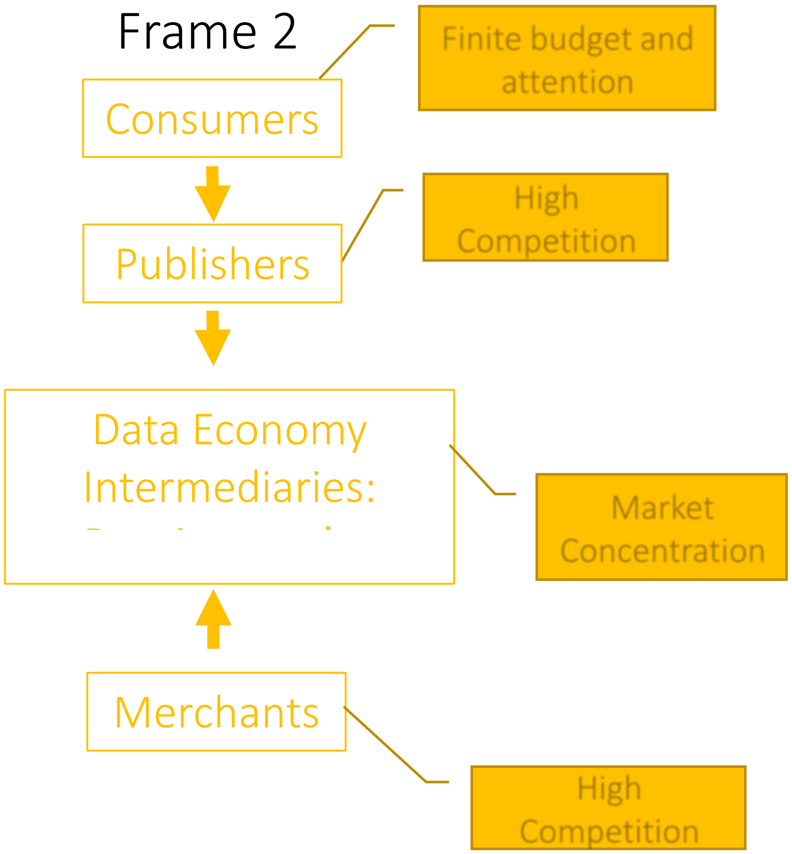
- To the extent that economic surplus is being generated by increasing (and increasingly sophisticated) consumer tracking, how is that surplus allocated?

Online advertising:

Frame 1

Online advertising:

Frame 2



The Online Advertising Market Puzzle

- Advertising revenues in US reached $88 billion in 2017 *(IAB, 2017)*
  - Growth rate of about 21.4%, relative to 2016

- However, revenues for about 40% of publishers – the final seller of ads – seem stagnant or shrinking *(Econsultancy, 2015)*
- Following GDPR enactment, NYT focused on contextual and geographical targeting and did not experience ad revenues drop (Jean-Christophe Demarta, SVP for global advertising at New York Times International, quoted by Digiday 2019b)
- A Digiday 2019 poll of publisher executives found that for 45% of respondents, behavioral ad targeting "has not produced any notable benefit, while 23% of publisher executives said behavioral targeting has actually caused their ad revenues to decline" (Digiday, 2019a)

Research Goals

- Provide insights on the relationship between advertisers ability to behaviorally target ads and publishers' revenues

- We leverage a unique dataset to investigate increase in publisher's revenues, after accounting for other factors, when the ads they sell can, or cannot, be behaviorally targeted via cookies to users
  - We focus on programmatic, open-auctions
  - We exploit the fact that if the user's cookie is not available, audience-based targeting is not implemented (other types of targeting can still be possible)

Related Works

- Advertising effectiveness:
  - Purchase Probabilities, Click-Through rates (Manchanda et al., 2006; Sahni, 2015; Farahat and Bailey, 2012; Bleier and Eisenbeiss, 2015; Lewis and Reley, 2014)
  - Page visits and online searches (Ghose and Todri-Adamopoulos, 2016; Johnson et al., 2017; Fong, 2016)
- Publishers' incentives and impact of targeting on revenues (Chen and Stallaert, 2014; Ghosh et al., 2015; Levin and Milgrom, 2010; Hummel and McAfee, 2016)
  - Theoretical predictions are mixed
- Empirical works on publishers' side are lacking

How Targeting May Affect Publishers' Revenue

- Advertisers willingness to pay increases if they can target audiences (Chen and Stallert, 2014; Board, 2009)
  - Ad prices increases, publisher's revenue increases

- When targeting audiences, advertisers reach narrow markets with reduced competition (Levin and Milgrom, 2010; Hummel and McAfee, 2016)
  - Ad prices decreases, publisher's revenue decreases

Data

- 2 million advertising transactions, over 60 different websites, 5,000 different advertisers, including:
  - Date and Time
  - Ad's features (size, type, etc..)
  - Webpage where ad was shown
  - Advertiser's name, industry, size
  - User's geo-location, device features, demographics
  - **User Cookie ID**
  - **Publisher's revenue**

Empirical Approach

- Observational data: a group of ads transactions has cookies associated and a group of transactions does not
    - Presence of cookies is associated with ability to behaviorally target (note, again: even in absence of cookies, other forms of targeting are possible – e.g. contextual targeting)
- Publisher's revenue is the outcome of a deterministic, programmatic process based on a given set of information
- Whether or not a user's cookie is available is outside the control of the publisher
    - Raw mean revenues are **higher with cookie is present**: average CPM $1.18 vs. $0.74
    - **However**: to isolate specific impact of cookie, we need to account for user's selection, and control for other factors

Empirical Approach

- Augmented Inverse Probability Weighting (Robins et al., 1994)

1. Estimate the Probability Model: Probability that user has a cookie associated

$$Prob_i(Cookie) = F(\beta_1 Demographics_i + \beta_2 Device_i + \beta 3 Location_i + \beta_4 X_i)$$

Where:

- X: vector of any other included features
- F : Logit function

Empirical Approach

2. Estimate two outcome models, one for transactions with cookies, one for transactions without

$$Y_i(t) = \beta_0 + \alpha Ad\_feat_i + \theta Website\_feat_i + \gamma User\_feat_i$$
$$+ \delta Advertisers\_feat_i + \eta X_i + \epsilon_i, t = (0, 1)$$

Where:

- Y_i: Publisher Revenue for transaction i
- Ad Features: Vector of ad level features
- Website Features: Vector of website level features
- User Features: Vector of user level features
- Advertisers Features: Vector of advertisers' features
- X: Vector of any additional covariate

Empirical Approach

3. Compute weighted means of treatment-specific predicted outcomes
4. Compute average treatment effect

- $Prob(Cookie|X) = \hat{c}_i$
- $m_1 = E(Y|T = 1, X), m_0 = E(Y|T = 0, X)$

$$\Delta_{DR} = \frac{1}{n}\sum_i \frac{T_i Y_i - (T_i - \hat{c}_i)m_1}{\hat{c}_i} - \frac{1}{n}\sum_i \frac{(1 - T_i)Y_i + (T_i - \hat{c}_i)m_0}{(1 - \hat{c}_i)}$$

- **Double-robustness:** only needs either the probability model or outcome models to be correctly specified for the estimate to be consistent

Results

	AIPW			
	Coeff.(Cookie)	*Std. Errors*	*P>—z—*	*[95% Conf. Interval]*
Seller_Revenue	0.0857	0.0009	0.000	[0.0837 - 0.0876]
$E(SellerRevenue \mid cookie = 0)$	0.9341	0.0033	0.000	[0.9276 - 0.9406]
$E(SellerRevenue \mid cookie = 1)$	1.0198	0.0034	0.000	[1.0129 - 1.0266]

- After controlling for other factors, when tracking cookie is available, revenue does increases – approximately by 4%, relative to when cookie is not available

Limitations

- The result can be interpreted as the increase in value generated for publishers specifically by the presence of a cookie
  - It cannot be interpreted as the value generated by behavioral advertising in general
- Our data pertain to a sample of websites of one large media company
  - Results may not apply to the entire universe of websites
- We observe publisher's revenue, already net of any intermediation fees
  - We do not have information on the actual amount of the fees
- We cannot capture presence of more sophisticated forms of tracking (e.g. device fingerprinting)