Who Benefits from the Data Economy?

Klaus M. Miller
HEC Paris

European Quant Marketing Workshop
January 16, 2025

2025

2025

About Me

- **Since Sept. 2021**
  - Assistant Professor of Marketing, HEC Paris
  - Chairholder Hi!PARIS Research Center for the Study of Data Analytics and Artificial Intelligence

- **Visiting Scholar**
  - Wharton School, University of Pennsylvania
  - Stanford University

- **Research on the Economics of Digitization**
  - Online Advertising
  - Privacy

Who Benefits from the Data Economy?

A Perspective on the Economic Value of User Tracking for Publishers using Augmented Inverse Probability Weighting (AIPW)

joint work with Rene Laub and Bernd Skiera

Supported by the European Research Council and the NET-Institute New York

Description of Problem

Description of User Tracking

- **User tracking …**
  - Collects information about a user over time
    - Requires identifier

- **Examples of online identifiers that enable user tracking**

E-Mail: xyz@gmail.com

Device IDs: abcd123456789

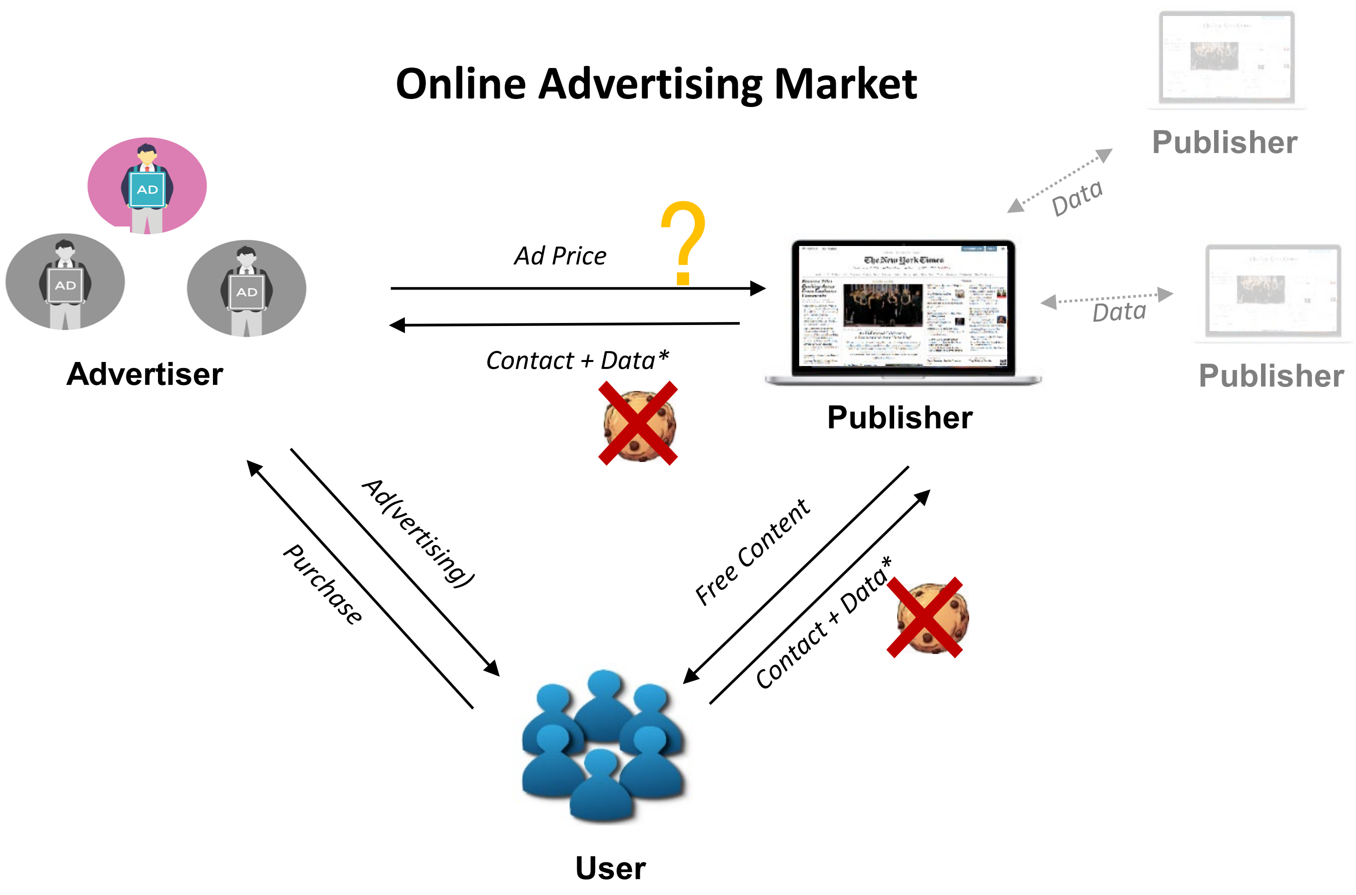
Phone Number: +49 123 456789

Fingerprints: "Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:109.0)

Cookies: Cfhhcnohhuknhuns.nytimes.com

Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Relevance of User Tracking in Online Advertising Market



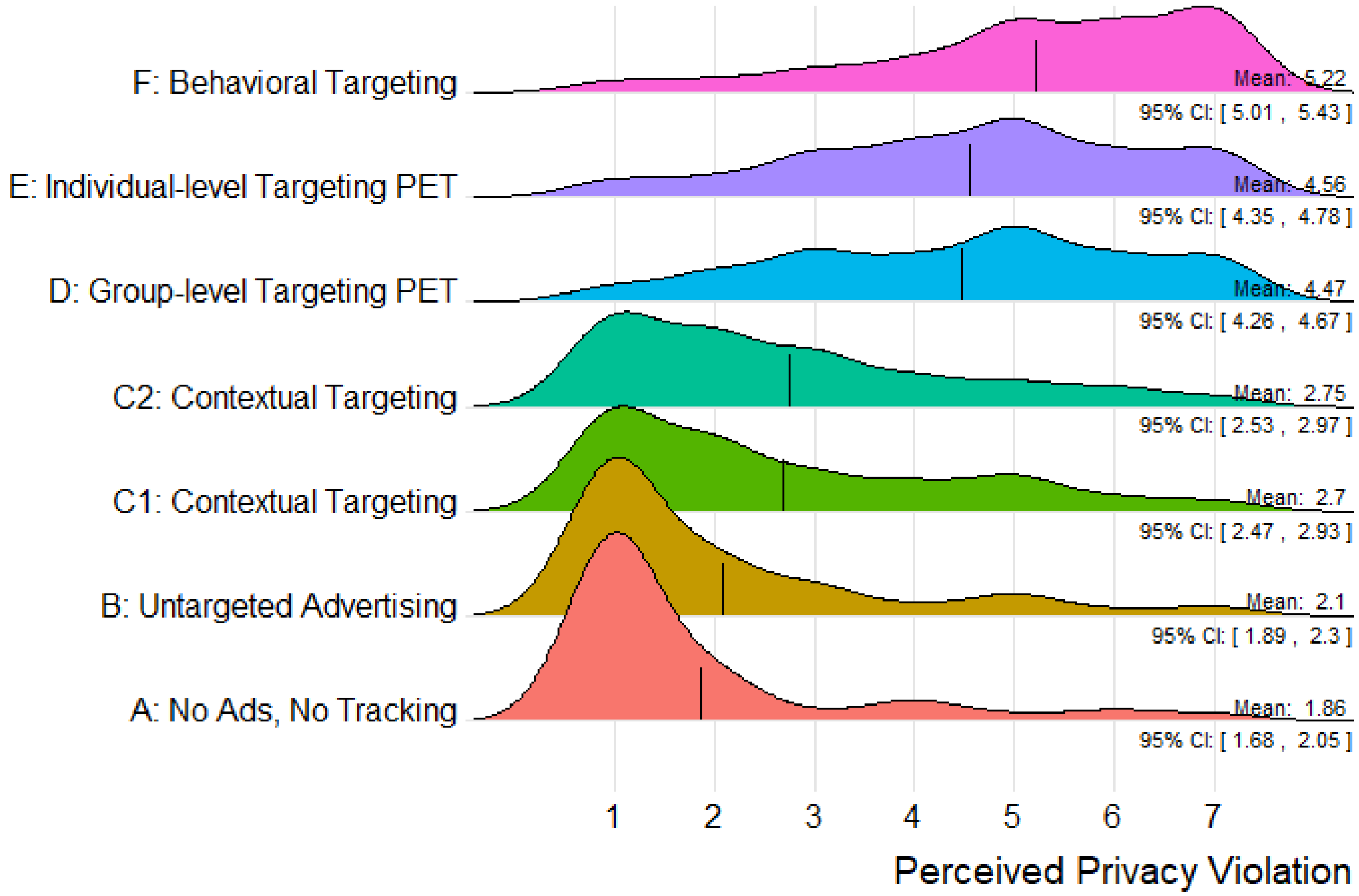
Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Two Main Usages of Tracked Data of Publishers

- **Better content, e.g.,**
  - Content personalization
    - Recommendations
    - User interface

- **Better advertising, e.g.,**
  - **Ad targeting**
    - Behavioral targeting
    - Retargeting
  - **Ad measurement**
    - Ad recency and frequency
    - Clicks
    - Conversions

Consumers' Perceived Privacy Violation



Source: Jerath and Miller 2024

Initiatives to Restrict User Tracking

- **Regulators, e.g.,**
  - Europe: General Data Protection Regulation (GDPR)
  - US: California Consumer Privacy Act (CCPA)
  - China: Personal Information Protection Law (PIPL)

- **Firms, e.g.,**
  - Apple's App Tracking Transparency (ATT)
  - Mozilla Firefox: Enhanced Tracking Protection (ETP)
  - Apple Safari: Intelligent Tracking Prevention (ITP)
  - Brave, Tor: Privacy-Focused Browsers

- **Activities from users, in particular, consumer protection agencies, e.g.,**
  - NOYB
  - Ad- and tracking blockers

Aim of Project

- **Research aim:**
  - Determine (economic) value of user tracking (for publishers)

		Value of Data for Firms	
		Low	High
Value of Data for User	Low	?!	Rather allow tracking
	High	Rather restrict tracking	?!

- **Research Questions:**
  - **RQ1: Average value of user tracking?**
  - **Differences of value of user tracking across**
    - RQ2: Users
    - RQ3: Publishers

Knowledge on the Value of User Tracking

Overview on Prior Research

- **Prior research on the value of user tracking**
  - **Focus of prior research**
    - Advertisers and ad intermediaries
  - **Little work on**
    - Users
    - **Publishers (our focus)**

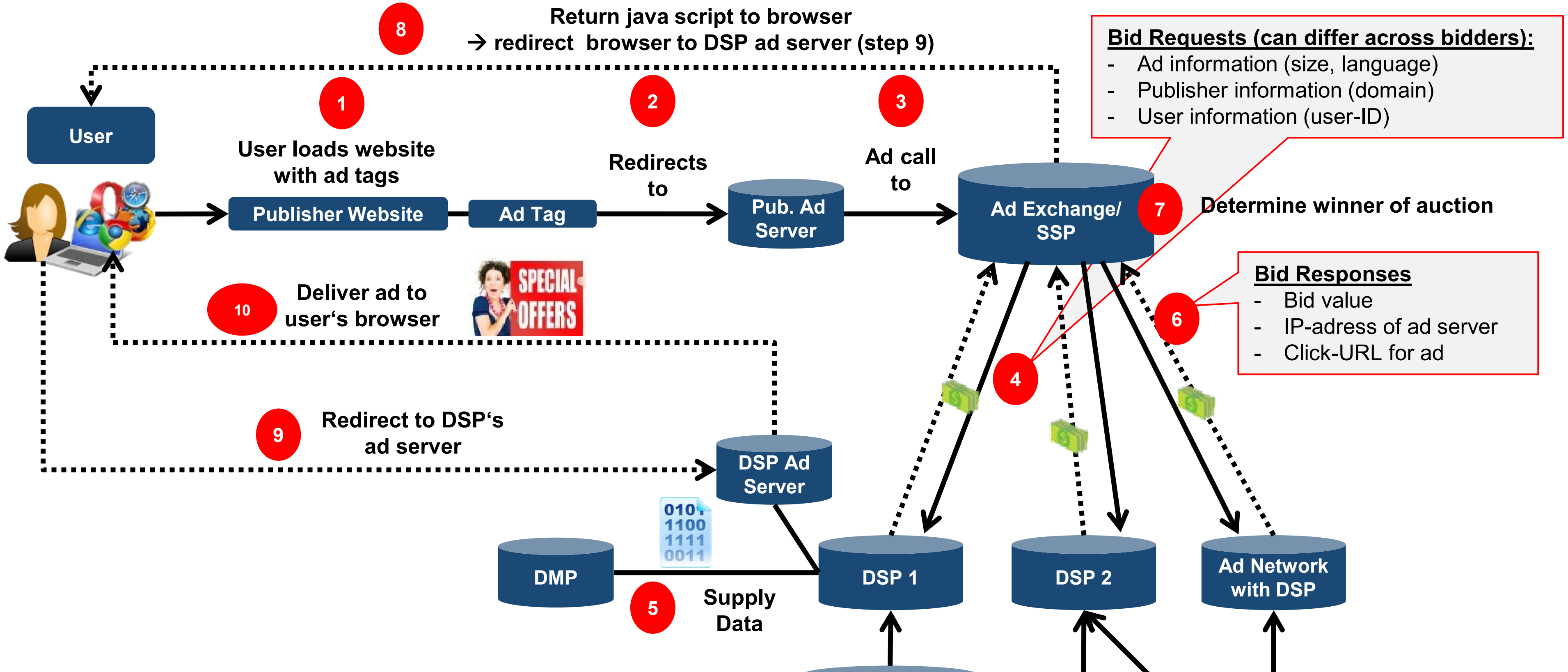
Empirical Academic Studies on the Value of User Tracking for Publishers

Study	Data	Average Value of User Tracking	Obser-vations	Observation Window	Observation Level	Geographical Focus	Type of Advertising	Type of Devices	Number of Publishers	Integration of Publisher	Heterogeneity of User Tracking across Characteristics of: Publishers	Heterogeneity of User Tracking across Characteristics of: Users
Marotta et al. 2019 (WP)	Observational data from one multisite publisher in 2016	8.0%	~ 2 Mio.	1 week	Individual ad impression	US	Display (Banner)	Desktop, Mobile, Tablet	1[a]	Horizontal	No	No
Wang et al. 2024 (JMR)	Observational data from one publisher in 2018	5.7%	~ 4 Mio.	10 weeks	Aggregate	US	Display (Native)	Desktop	1	Vertical	No	No
Sun et al. 2023 (MS)	Experimental data from one e-commerce platform in 2019	81%	~ 0.6 Mio.	7 hours	Individual user	China	Product Recommen-dations	Mobile	1	Vertical	No	No
Our Study	Observational data from two intermediaries in 2016 and 2023	EU: 18%-23% US: 47%	~ 42 Mio. (Study 1) ~ 31,890 (Study 2)	2 weeks (Study 1) 6 weeks (Study 2)	Individual ad impression, Aggregate	EU (Study 1) US & EU (Study 2)	Display (Banner, Video, Mobile)	Desktop, Mobile, Tablet	111[b] (Study 1) 10,526 (Study 2)	Independent	Ad Inventory Content Size	User ID Browsing History Ad Recency Ad Frequency

Notes: WP = Working Paper, JMR = Journal of Marketing Research, MS = Management Science. [a]One large publisher with 60 distinct websites. [b]Publishers with 84% reach in the respective market.

Description of Real-Time Bidding Data

Real-Time Bidding Auction Process



DSP = Demand Side Platform
DMP = Data Management Platform

Redirect:
If browser is redirected to URL of ad server, ad server gets access to its cookies, if placed, or possibility to place a cookie

Ad Tag

<script> src=„ib.adnxs.com/ttj?id=„123" type=text/javascript</script>

Bid Request Data in Real-Time Bidding

Data Categories		Common Variables	Example
User data	**Data generated by user tracking**	Identifier (ID)*	User ID “123-ABC-789”.
		Browsing history	The user visited www.sports.com three times already.
		Ad recency	The user saw the ad “ABC” two minutes ago.
		Ad frequency	The user saw the ad “ABC” four times already.
	Data not generated by user tracking	Device and software (e.g., operating system, browser) of user	The user is browsing the internet with a Samsung tablets, Android OS, and Firefox browser.
		Location of user	The user is in Paris, France.
		Date and time	The time of the user’s visit is 2 pm on a Monday.

Theoretical Background of Empirical Studies

Value of User Tracking for Publishers

- **Effects with user tracking (compared to without user tracking)**
  - **Targeting effect**
    - Tracking data enables better targeting yielding
      - Higher willingness-to-pay
      - Higher ad prices
  - **Competition effect**
    - Tracking data enables targeting very narrow groups of users for which only few advertisers compete (“thin market”) yielding
      - Less competition
      - Lower ad prices

- **Need for empirical study**
  - Resulting overall effect of both contradicting effects unclear
  - Effect of publisher and user characteristics on effect unclear

Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Empirical Studies

Comparison of Setup of Both Empirical Studies

	Study 1	Study 2
Data Source	Ad exchange (2016)	Demand-side Platform (2023)
Number of Ad Impressions	41,767,963	218,394,708
Share of Trackable Ad Impressions	85%	Apple: 17%, Android: 91%
Observation Window	2 weeks (in April 2016)	6 weeks (Mid-September until end of October 2023)
Geographical Focus	EU	EU & US
Number of Publishers	111	10,526
Type of Advertising	Display	Display
Type of Devices	Desktop, Tablet, Mobile (Browser)	Mobile (In-App)
Average Value of User Tracking	Yes	Yes
Heterogeneity of Value of User Tracking across Characteristics of Publishers	Yes	Yes
Heterogeneity of Value of User Tracking Across Characteristics of Users	Yes	No

Setup of First Empirical Study

Description of Data

- **Price of 42 million ad impressions from large European Ad Exchange**
  - User tracking via third-party-cookies
    - 85% of ad impressions with cookie (~1.4 Mio. cookie IDs)
    - 15% of ad impressions without cookie

- **111 publishers**

- **Ad impression characteristics, e.g.,**
  - Ad position
  - Ad size

- **User characteristics, e.g.,**
  - Device of user
  - Internet browser of user
  - …

- **Publisher characteristics, e.g.,**
  - Topic area of content
  - …

Prices of Trackable and Untrackable Ad Impressions

(CPM) Price in US$	Mean	Standard Deviation	Median	Min	q95	q98	Max	N (= Ad Impressions)
Panel A: Raw Price Distribution								
With User Tracking	0.691	0.950	0.524	0.002	1.586	2.862	131.622	35,515,448
Without User Tracking	0.274	0.324	0.113	0.003	0.954	1.063	18.097	6,252,515
Relative Price Difference	-60.3%	-65.9%	-78.4%	-50.0%	-39.8%	-62.9%	-86.3%	-

Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Comparison of Trackable and Untrackable Ad Impressions

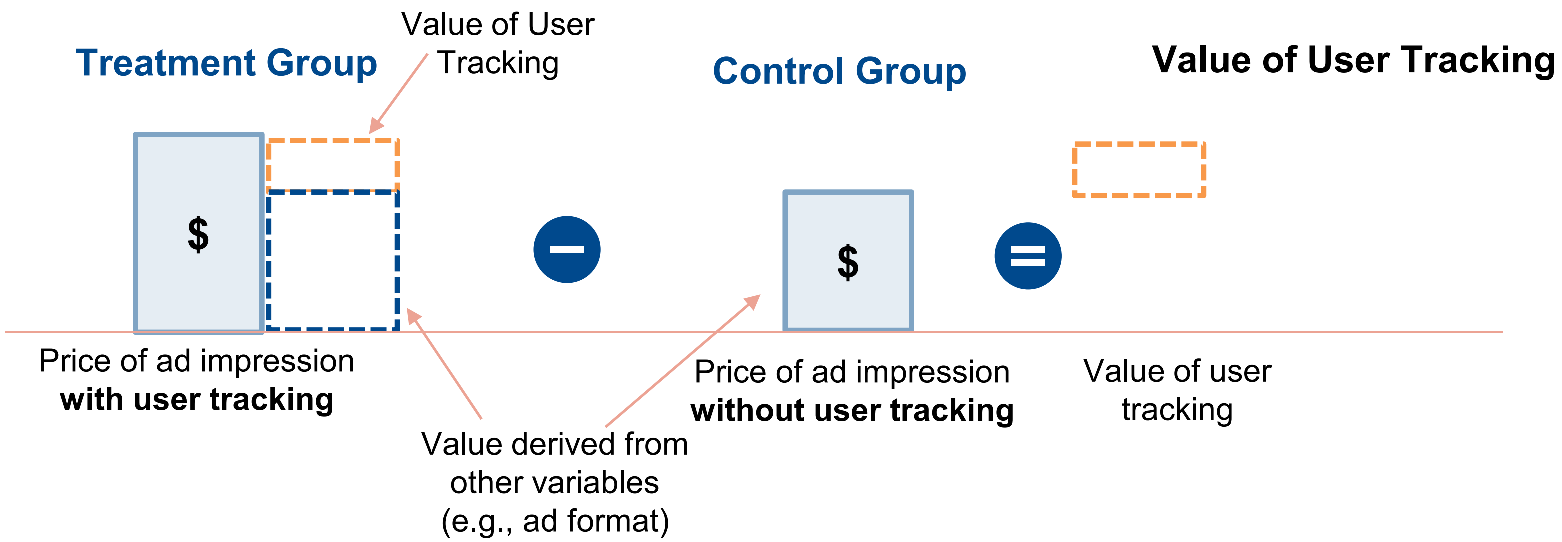
Variable		All Ad Impressions (41,767,963, 100.00%)		Ad Impressions with User Tracking (35,515,448, 85.03%)		Ad Impressions without User Tracking (6,252,515, 14.97%)	
		Number of Ad Impressions	Avg. Price (CPM)	Number of Ad Impressions (% of total)	Avg. Price (CPM)	Number of Ad Impressions. (% of total)	Avg. Price (CPM)
Device	Desktop	34,543,967	0.678 $	31,910,002 (92.38%)	0.705 $	2,633,965 (7.62%)	0.349 $
	Smartphone	2,679,746	0.180 $	1,179,826 (44.03%)	0.252 $	1,499,920 (55.97%)	0.123 $
	Tablet	1,990,033	0.682 $	1,097,181 (55.13%)	0.807 $	892,852 (44.87%)	0.528 $
	Unknown	2,554,217	0.385 $	1,328,439 (52.01%)	0.637 $	1,225,778 (47.99%)	0.111 $
Operating System	Android	5,116,678	0.292 $	2,601,298 (50.84%)	0.478 $	2,515,380 (49.16%)	0.099 $
	Apple Macintosh	1,606,277	0.548 $	662,930 (41.27%)	0.808 $	943,347 (58.73%)	0.365 $
	Apple iOS	1,543,166	0.604 $	500,762 (32.45%)	0.825 $	1,042,404 (67.55%)	0.498 $
	BlackBerry OS	13,417	0.336 $	13,218 (98.52%)	0.339 $	199 (1.48%)	0.191 $
	Linux	109,864	0.676 $	100,543 (91.52%)	0.704 $	9,321 (8.48%)	0.378 $
	Microsoft Windows	32,872,175	0.684 $	31,188,951 (94.88%)	0.703 $	1,683,224 (5.12%)	0.340 $
	Symbian OS	513	0.122 $	505 (98.44%)	0.123 $	8 (1.56%)	0.113 $
	Unknown	505,873	0.725 $	447,241 (88.41%)	0.769 $	58,632 (11.59%)	0.392 $
Browser	Android	782,883	0.386 $	774,737 (98.96%)	0.386 $	8,146 (1.04%)	0.455 $

Identification Strategy

Measuring the Value of User Tracking

- **Aim of Empirical Study:** Estimate price difference between ad impressions with and without user tracking

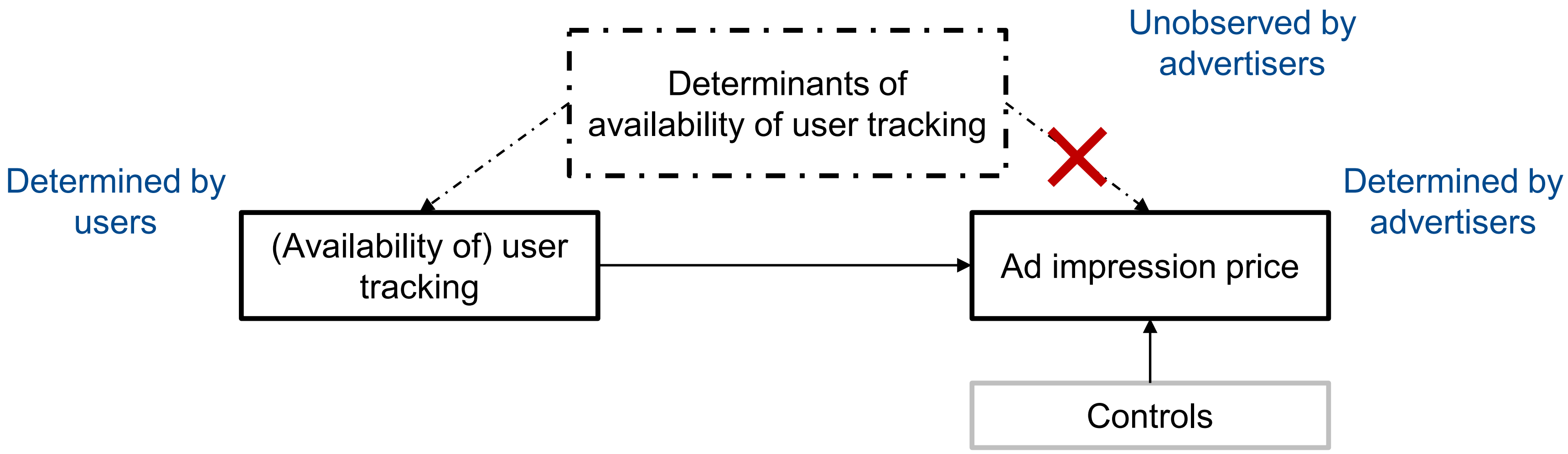
- User tracking via cookie (first study) or device id (second study)
- Treatment T = 1, if cookie/device id is available (0 otherwise)



$\mathbb{E}[Y(1)]$ $\mathbb{E}[Y(0)]$ $ATE = \mathbb{E}[Y(1) - Y(0)]$

Selection into Treatment

- **Effect of interest:**



- **Outcomes of interest:** Ad impression price with and without user tracking

- Availability of user tracking is not-random, however
  - Underlying determinants (e.g., age) unobserved by advertiser
  - We observe same variables as advertisers (e.g., proxies for age)

-> selection on unobservables (self-selection) should not be a concern for estimation

Adjustment for Selection into Treatment

- **Regression Analysis (RA)**
- **Augmented Inverse Probability Weighting (AIPW)**
- Heckman Selection Model

Regression Analysis

Implementation of Procedure to Determine Results

- **Part 1: Estimation of Model (Regression Analysis)**

- **Part 2: Determination of Treatment Effects**
  - For each impression calculate treatment effect (TE) by determining difference of predicted price of
    - Trackable user
    - Untrackable user
  - Calculation of
    - Average treatment effect (ATE) (all observations)
    - Average treatment effect on the treated (ATET) (only trackable users)

Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Augmented Inverse Probability Weighting (AIPW)

Implementation of
Augmented Inverse Probability Weighting (AIPW)

- **Setup**
  - Estimate the **Average Treatment Effect (ATE)** of a binary treatment T_i on an outcome Y_i (ad price)
    - T = 1 if treated (trackable user)
    - T = 0 if untreated (untrackable user)

Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Implementation of Augmented Inverse Probability Weighting (AIPW)

- **Step 1:** Estimate the Treatment Model
- **Step 2:** Estimate the Outcome Models
- **Step 3**: Compute the AIPW Estimator

Step 1: Estimate the Treatment Model

Inverse Probability Weighting (IPW)

- Estimate treatment probability $\hat{\pi}_i$ for each ad impression *i* (propensity score)
  - Logistic regression
  - Boosted regression trees
- Weight observations:
  - Treatment group (T_i= 1): 1 / $\hat{\pi}_i$
  - Control group (T_i= 0): 1 / (1 - $\hat{\pi}_i$)

$$\widehat{ATE_{IPW}} = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{T_i Y_i}{\hat{\pi}_i} - \frac{(1-T_i)Y_i}{(1-\hat{\pi}_i)}\right\}$$

$$Y_i = \text{price of ad impression } i$$

Step 2: Estimate the Outcome Models

Regression Adjustment (RA)

- Estimate price of each ad impression i based on covariates X (e.g. ad size, ad position)
- Separate estimation per treatment group
  - Estimate counterfactual price using regression estimates from opposite group
  - Determine outcome regression models via linear regression or random forests

$$\mu_1(X_i) = \mathbb{E}[Y \mid T_i = 1, X_i], \quad \mu_0(X_i) = \mathbb{E}[Y \mid T_i = 0, X_i].$$

$$\widehat{ATE}_{RA} = \frac{1}{n}\sum_{i=1}^{n}\{\hat{\mathbb{E}}(Y_i|T_i = 1, X_i) - \hat{\mathbb{E}}(Y_i|T_i = 0, X_i)\}$$

$

Step 3: Compute the AIPW Estimator

Double Robust Estimator

- Either treatment model or outcome model needs to be correctly specified
- Performs best* even under severe confounding (Glynn & Quinn 2010)

$$\widehat{ATE}_{AIPW} = \frac{1}{n}\sum_{i=1}^{n}\Big[\underbrace{\frac{T_i}{\hat{\pi}_i}\left(Y_i - \hat{\mu}_1(X_i)\right) + \hat{\mu}_1(X_i)}_{\text{Treated part}} - \underbrace{\frac{1-T_i}{1-\hat{\pi}_i}\left(Y_i - \hat{\mu}_0(X_i)\right) + \hat{\mu}_0(X_i)}_{\text{Untreated part}}\Big]$$

*Comparable or lower mean square error than propensity score matching, IPW or regression adjustment

Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Why Double Robust?

Double Robustness

- AIPW called "**double robust**" because it only requires **one** of the two models is correct:
  - Treatment model (propensity score model)
  - Outcome model (regressions)
- AIPW remains a consistent estimator of the true ATE

What if both are misspecified?

- Double-robust guarantee breaks down
- AIPW can be biased
- Model checking is critical
  - Theoretically (all variables included?)
  - Empirically (model fit?, robustness?)

Key Takeaways on AIPW

- **AIPW** stands for Augmented Inverse Probability Weighting
  - Augments simple IPW with plug-in outcome model
- **Double robustness** if correct specification of either
  - Propensity score model
  - Outcome models
- **If both are wrong**, no protection, and AIPW can be **biased**
  - Theoretical model specification
  - Model testing

Results of Empirical Study

Results for
Average Value of User Tracking (RQ1)

Regression Results

HEC PARIS

Value of User Tracking at the Ad Impression Level

Panel A: Linear Regression Model Estimation

Dependent Variable: Log(Price)	Model 1.1	Model 1.2	Model 1.3	Model 1.4
User Tracking (1/0)	0.939*** (0.000)	0.753*** (0.023)	0.539*** (0.067)	0.200*** (0.042)
Other User Data				
Device × Operating System × Browser		Yes	Yes	Yes
Time (Week, Weekday, Hour of Day)		Yes	Yes	Yes
Location of User (Continent, City)		Yes	Yes	Yes
Publisher Data				
Ad Position (Above Fold)			Yes	Yes
Ad Format			Yes	Yes
Publisher ID			Yes	Yes
Control				
Advertiser ID				Yes
Adj. R^2	0.119	0.224	0.415	0.654
N Ad Impressions	41,767,963	41,767,963	41,767,963	41,767,963

Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Value of User Tracking at the Ad Impression Level

Panel B: Price Predictions (Potential Outcomes) and Treatment Effects				
All Ad Impressions				
	Model 1.1	**Model 1.2**	**Model 1.3**	**Model 1.4**
E(Price| User Tracking =1)	0.689 \$	0.666 \$	0.671 \$	0.652 \$
E(Price| User Tracking =0)	0.269 \$	0.314 \$	0.392 \$	0.533 \$
ATE (in US\$)	**-0.420 \$**	**-0.352 \$**	**-0.279 \$**	**-0.119 \$**
ATE (%) / Relative Price Difference	**-60.9%**	**-52.9%**	**-41.6%**	**-18.3%**
N Ad Impressions	41,767,963	41,767,963	41,767,963	41,767,963

Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Value of User Tracking at the Ad Impression Level

Only Ad Impressions with User Tracking (Treatment Group)				
E(Price | User Tracking =1, Treatment =1)	0.689 $	0.675 $	0.696 $	0.703 $
E(Price | User Tracking =0, Treatment =1)	0.269 $	0.318 $	0.406 $	0.576 $
ATET (in US$)	**-0.420 $**	**-0.357 $**	**-0.290 $**	**-0.127 $**
ATET (%) / Relative Price Difference	**-60.9%**	**-52.9%**	**-41.7%**	**-18.1%**
N Ad Impressions	35,515,448	35,515,448	35,515,448	35,515,448

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.
Notes: ATE = Average Treatment Effect, ATET = Average Treatment Effect on the Treated. The ATE and ATET in percentages correspond to the relative price change from Equation (1) and correspond to a situation where we move from ad impressions with user tracking to ad impressions without user tracking. Thus, the estimated ATE and ATET are negative. Robust standard errors are in parentheses.

Augmented Inverse Probability Weighting (AIPW) Results

Value of User Tracking at the Ad Impression Level

- **Step 1: Treatment Model (Propensity Score Model)**

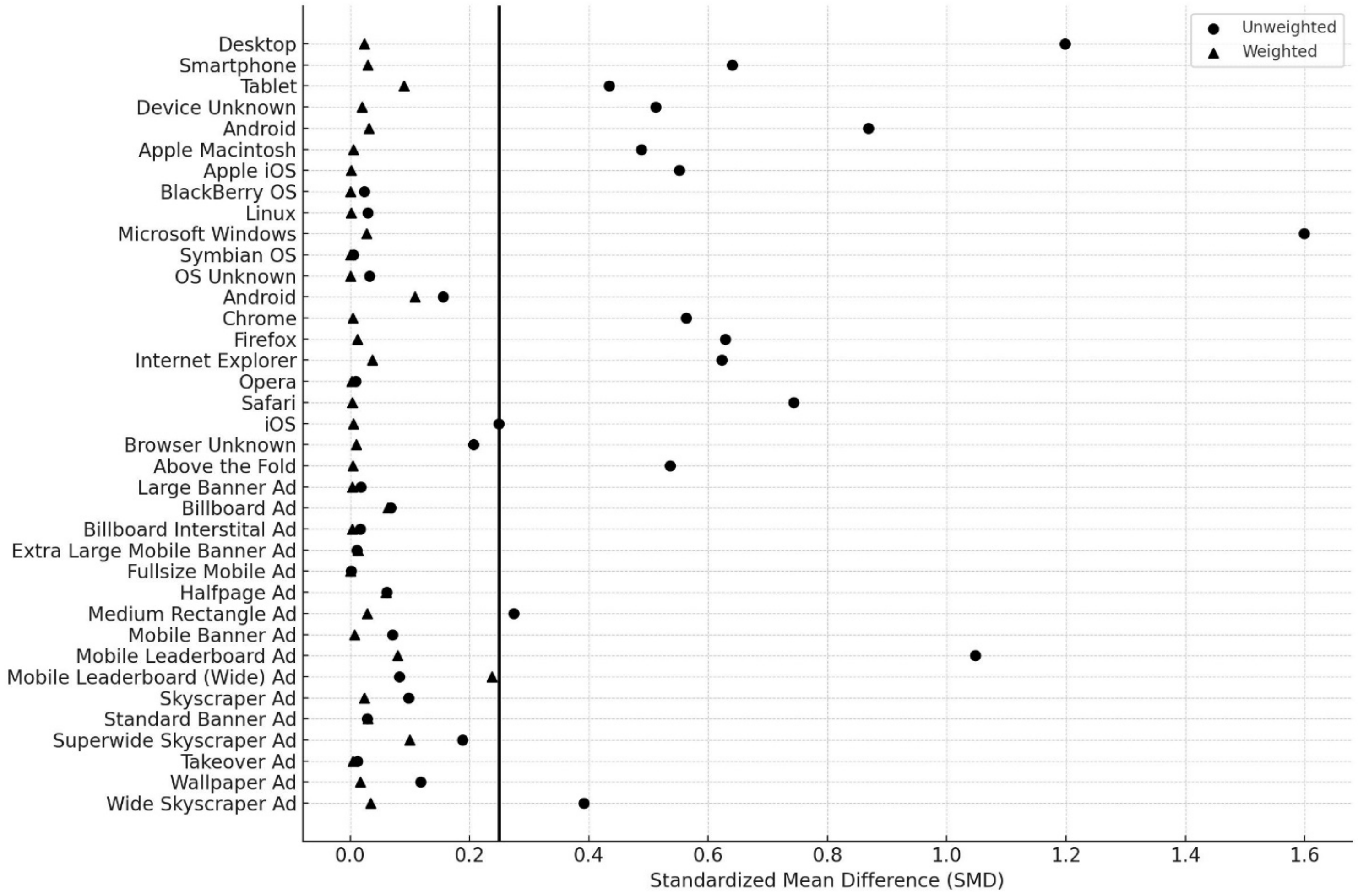
Table A4. Study 1a: Results Estimation of Probability of User Tracking Presence

Dependent Variable: Log(Price)	Logit for AIPW	Probit for Heckman
Log(Google Trends "http cookie")	-	-0.004*** (0.001)
Device	Yes	Yes
OS	Yes	Yes
Browser	Yes	Yes
Time of Day, Weekday, Week	Yes	Yes
Publisher ID	Yes	Yes
AIC	20,939,784	20,993,120
N Ad Impressions	41,767,963	41,767,963

*p < 0.1; **p < 0.05; ***p < 0.01.
Notes: AIPW = Augmented Inverse Probability Weighting. Robust standard errors are in parentheses.

Value of User Tracking at the Ad Impression Level

- **Step 1: Treatment Model (Propensity Score Model)**



Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Value of User Tracking at the Ad Impression Level

- **Step 2: Outcome Models**

Table A5. Study 1a: Regression Results Robustness Checks

Dependent Variable: Log(Price)	Linear Regression (Full Data)	Linear Regression (98% Outlier Corrected Data)	Linear Regression (95% Outlier Corrected Data)	AIPW (Outcome Regression)	Heckman (Second Stage)
User Tracking (1/0)	0.200***	0.194***	0.194***	0.201***	0.202***
	(0.042)	(0.042)	(0.042)	(0.042)	(0.041)
Inverse Mills Ratio					0.452***
					(0.045)
Other User Data					
Device x Operating System × Browser	Yes	Yes	Yes	Yes	Yes
Time (Week, Weekday, Hour of Day)	Yes	Yes	Yes	Yes	Yes
Location of User (City)	Yes	Yes	Yes	Yes	Yes
Ad Slot Data					
Ad Position (Above Fold)	Yes	Yes	Yes	Yes	Yes
Ad Format	Yes	Yes	Yes	Yes	Yes
Publisher ID	Yes	Yes	Yes	Yes	Yes
Control					
Advertiser ID	Yes	Yes	Yes	Yes	Yes
Adj. R2	0.654	0.647	0.654	0.654	0.654
N Ad Impressions	41,767,963	40,937,021[A]	39,757,090[A]	41,430,997[B]	41,767,963

*p < 0.1; **p < 0.05; ***p < 0.01. Notes: Robust standard errors are in parentheses.
AIPW: Augmented Inverse Probability Weighting
[A] To account for the potential effect of outliers, we exclude all ad impressions with a price > $2.98 (= 830,942 ad impressions, 1.98%) in the 98% outlier corrected data set, and all ad impressions with a price > 1.55 (= 2,010,873, 4.8%) in the 95% outlier corrected data set.
[B] In line with previous research, we exclude 336,966 (0.80%) ad impressions for the AIPW estimate with a very high ($\hat{\pi}_i > 0.999$) or very low ($\hat{\pi}_i < 0.001$) treatment probability because they could yield very high or very low weights.

Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Value of User Tracking at the Ad Impression Level

- **Step 3: Determination of AIPW Estimator**

Table 4. Study 1a: Robustness Checks for the Average Treatment Effect

	Average Treatment Effect (ATE) All Ad Impressions				
	Linear Regression (Full Data)	**Linear Regression** (98% Outlier Corrected Data)	**Linear Regression** (95% Outlier Corrected Data)	**AIPW**	**Heckman**
E(Price| User Tracking =1)	0.652 $	0.570 $	0.529 $	0.632 $	0.652 $
E(Price| User Tracking =0)	0.533 $	0.470 $	0.441 $	0.507 $	0.531 $
ATE (in US$)	**-0.119 $**	**-0.100 $**	**-0.088 $**	**-0.125 $**	**-0.121 $**
ATE (%) / Relative Price Difference	**-18.3%**	**-17.5%**	**-16.6%**	**-19.8%**	**-18.5%**
N Ad Impressions	41,767,963	40,937,021[A]	39,757,090[A]	41,430,997[B]	41,767,963

Notes: AIPW = Augmented Inverse Probability Weighting.
[A] To account for the potential effect of outliers, we excluded all ad impressions with a price > $2.98 (= 830,942 ad impressions, 1.98%) in the 98% outlier corrected data set, and all ad impressions with a price > $1.55 (= 2,010,873 ad impressions, 4.8%) in the 95% outlier corrected data set.
[B] In line with previous research, we excluded 336,966 (0.80%) ad impressions for the AIPW estimate with a very high ($\hat{\pi}_i > 0.999$) or very low ($\hat{\pi}_i < 0.001$) treatment probabilities because they could yield very high or very low weights.

Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Results for Heterogeneity Across Users (RQ2)

Advertising Components of Value of User Tracking

- **Value of Identifier**
- **Browsing History**
- **Ad Recency**
- **Ad Frequency**

Dependent Variable: Log(Price)	Model 4.1
Identifier available	0.317*** (0.010)
Car Website Visits[A]	-0.318*** (0.090)
Computer & Technology Website Visits[A]	0.583*** (0.143)
Dating Website Visits[A]	-1.499*** (0.525)
Entertainment Website Visits[A]	0.101 (0.072)
Games Website Visits[A]	-0.077** (0.037)
Health & Medicine Website Visits[A]	-0.677*** (0.063)
Lifestyle & Shopping Website Visits[A]	1.351*** (0.386)
Finance & Real Estate Website Visits[A]	0.130*** (0.039)
News & Information Portal Website Visits[A]	-0.239 (0.429)
Sports Website Visits[A]	-0.260*** (0.049)
Student's Interest Websites Visits[A]	0.448*** (0.226)
Travel Website Visits[A]	0.158* (0.085)
Women's Interest Website Visits[A]	0.288** (0.135)
Ad Recency Groups[B]	Yes
Ad Frequency Groups[B]	Yes
Other User Data, Publisher Data & Controls	
Device x Operating System x Browser, Time (Week, Weekday, Hour of Day), Location of User (Continent, City), Ad Position (Above Fold), Ad Format, Publisher ID, Advertiser ID	Yes
Adj. R2	0.661
N	41,767,963

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.
Notes: Robust standard errors are in parentheses.
[A] We include the standardized browsing history variables in the estimation.
[B] See Web Appendix Table A9 for all ad recency and ad frequency group coefficients.

Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Value of User Data Categories (Based upon Model 4.1)

Table 8. Study 1a: Scenarios and Expected Prices for Value of Data Categories

Scenario Number	*Description of Scenario*	*Expected Price (in US$)*	*Differences in Expected Prices*
Scenario 1	No (data from) user tracking	0.536	-
Scenario 2	Only user ID available	0.737	Scenario 2 – Scenario 1: 0.737 – 0.536 = 0.201
Scenario 3	Cookie id and browsing history available	0.739	Scenario 3 – Scenario 2: 0.739 – 0.737 = 0.002
Scenario 4	Cookie id and ad recency available	0.727	Scenario 4 – Scenario 2: 0.727 – 0.737 = -0.010
Scenario 5	Cookie id and ad frequency available	0.709	Scenario 5 – Scenario 2: 0.709 – 0.737 = -0.028

Notes: All numbers are rounded to 3 decimals. N = 35,515,448 (= all ad impressions with user tracking).

Value of User Data Categories (Based upon Model 4.1)

Table 9. Study 1a: Price Differences for Value of Data Categories

		Price Differences (In US$)							
Value Added by Data Categories	Calculation:	Mean	Sd.	Quantiles					
				0%	50%	75%	90%	98%	100%
Value added by Browsing History	Scenario 3 - Scenario 2	0.002	0.007	-2.391	0.002	0.003	0.004	0.009	2.131
Value added by Ad Recency	Scenario 4 - Scenario 2	-0.010	0.029	-0.893	0.000	0.000	0.000	0.000	0.000
Value added by Ad Frequency	Scenario 5 - Scenario 2	-0.028	0.046	-1.110	0.000	0.000	0.000	0.000	0.000

Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Results for
Heterogeneity Across Publishers (RQ3)

Publisher Components of Value of User Tracking

- **Type of ads**
  - Above fold
  - Large ads (obtrusiveness)
- **Type of publisher**
  - Premium publisher
  - Thematic-focused
  - Publisher Size
  - Publisher Topic

Table 6. Study 1a: Value of (A) Ad Obtrusiveness and Visibility, (B) Premium, Thematic Content, and Size

	(A)		(B)	
Dependent Variable: Log(Price)	Model 2.1	Model 2.2	Model 3.1	Model 3.2
User Tracking	0.303***	0.258***	0.191**	0.244***
	(0.036)	(0.032)	(0.074)	(0.056)
Ad Above Fold	0.039***	-0.074**		
	(0.007)	(0.025)		
Ad Above Fold × User Tracking		0.127***		
		(0.023)		
Large Ad	0.131***	0.400***		
	(0.015)	(0.006)		
Large Ad × User Tracking		-0.295***		
		(0.014)		
Premium			0.277***	0.548***
			(0.075)	(0.005)
Premium × User Tracking				-0.106*
				(0.009)
Thematic-Focused			0.253***	0.512***
			(0.016)	(0.014)
Thematic-Focused × User Tracking				-0.258***
				(0.021)
Publisher Size			-0.003	-0.184***
			(0.035)	(0.039)
Publisher Size x User Tracking				0.094***
				(0.035)
Publisher Topic Category			Yes	Yes
Device x Operating System x Browser	Yes	Yes	Yes	Yes
Time (Hour, Day, Week)	Yes	Yes	Yes	Yes
Location of User (City)	Yes	Yes	Yes	Yes
Advertiser ID	Yes	Yes	Yes	Yes
Adj. R^2	0.552	0.553	0.573	0.582
N Ad Impressions	41,767,963	41,767,963	41,767,963	41,767,963

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$. Notes: Robust standard errors are in parentheses.

Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Differences of the Value of User Tracking across Publishers

	Price Predictions and Treatment Effects					
	Premium Publisher	Non-Premium Publisher	Thematic-focused Publisher	Thematic-broad Publisher	Large Publisher	Small Publisher
E(Price| User Tracking =1)	0.725	0.621	1.060	0.644	0.647	0.722
E(Price| User Tracking =0)	0.578	0.444	1.050	0.473	0.477	**0.630**
ATE (in US$)	**-0.147**	**-0.177**	**0.010**	**-0.171**	**-0.17**	**-0.092**
ATE (%) / Relative Price Difference	**-20.3%**	**-28.5%**	**-0.9%**	**-26.6%**	**-26.3%**	**-12.7%**
N Ad Impressions	10,344,836	31,423,127	350,502	41,447,461	41,628,978	138,985
	41,767,963		41,767,963		41,767,963	

Notes: ATE = Average Treatment Effect. Correlation of premium publisher and thematic-focused publisher (-0.170), correlation of premium publisher and large publisher (0.390), correlation of large publisher and thematic-focused publisher (-0.320). Number of premium publishers = 24 (22%). Number of thematic-focused publishers = 35 (32%). Number of large publishers = 55 (50%).

Setup of Second Empirical Study

Description of Data Set

- **Data set from demand-side platform (DSP) in programmatic mobile ad market**
  - DSP receives bid request with following features (among others)
    - operating system of device (i.e., Apple or Android)
    - availability of a device ID
    - country of the user (e.g., a European (EU) country or US)
    - date and time of bid request
    - In case of bidding, DSP receives winning price

- **31,890 publisher instances (publisher x device x operating system x ad format)**
  - 10,526 publishers (apps)

- **About 218 million ad impressions observed over six weeks (mid-September to end-October 2023)**
  - Share of impressions with device ID
    - Android (91%)
    - Apple (17%)

Prices of Trackable and Untrackable Ad Impressions

(CPM) Price in US$	Mean	Standard Deviation	Median	Min	q95	q98	Max	N (= Ad Impressions)
Panel A: EU (N = 10,433,115)								
With User Tracking	7.817	6.904	6.862	0.040	18.648	28.770	86.480	4,915,925
Without User Tracking	5.860	7.579	4.313	0.034	19.741	29.001	120.509	5,517,190
Relative Price Difference	-25.0%	9.8%	-37.1%	-15.0%	5.9%	0.8%	39.3%	-
Panel B: United States (N = 207,961,593)								
With User Tracking	8.577	13.443	1.550	0.087	35.742	50.107	224.581	74,483,959
Without User Tracking	4.720	9.843	0.510	0.059	21.933	36.712	224.446	133,477,634
Relative Price Difference	-45.0%	-26.8%	-67.1%	-32.2%	-38.6%	-26.7%	-0.1%	-

Notes: N_{Total} = 218,394,708; CPM Price = price for 1,000 ad impressions, min = minimum, max = maximum, q95 = 95% quantile, q98 = 98% quantile.

Results for Average Value of User Tracking (RQ 1)

Regression Results on the Value of User Tracking

	EU			US		
Dependent Variable: Log(Price)	**Model 6.1**	**Model 6.2**	**Model 6.3**	**Model 7.1**	**Model 7.2**	**Model 7.3**
User Tracking (1/0)	0.861*** (0.001)	0.368*** (0.001)	0.266*** (0.001)	1.052*** (0.000)	0.641*** (0.029)	0.628*** (0.015)
Other User Data						
Operating System = iOS		-0.714*** (0.001)	0.010 (0.001)		-0.674*** (0.032)	0.007*** (0.000)
Publisher Data						
Ad Format = Interstitial			2.942*** (0.001)			3.040*** (0.012)
Ad Format = Rewarded			3.462*** (0.001)			3.682*** (0.018)
Adj. R^2	0.066	0.088	0.771	0.076	0.089	0.771
N Publisher-Instances / N Publishers / N Ad Impressions	3,412 / 1,225 / 10,433,115	3,412 / 1,225 / 10,433,115	3,412 / 1,225 / 10,433,115	28,478 / 9,301 / 207,961,593	28,478 / 9,301 / 207,961,593	28,478 / 9,301 / 207,961,593

Derivation of Treatment Effects

Panel B: Price Predictions and Treatment Effects		
	EU	**US**
	Model 6.3	**Model 7.3**
E(Price| User Tracking =1)	8.62 $	9.16 $
E(Price| User Tracking =0)	6.61 $	4.89 $
ATE (in US$)	**-2.15 $**	**-4.27 $**
ATE (%) / Relative Price Difference	**-23.3%**	**-46.6%**
N Publisher-Instances / N Publishers / N Ad Impressions	3,412 / 1,225 / 10,433,115	28,478 / 9,301 207,961,593

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$. Notes: ATE = Average Treatment Effect. The ATE in percentages corresponds to the relative price change from Equation (1). Robust standard errors are in parentheses.

Summary and Implications

Summary and Implications

- **RQ1: Ad prices are, on average, lower without user tracking**
  - Study 1: -18.3%
  - Study 2
    - Europe: -23.3%
    - US: -46.6%
  - Quantity and quality of free content for users at risk

- **RQ2: Differences across users**
  - Mainly driven by identifier
    - Enables ad performance measurement
  - Browsing history hardly generates economic value
    - Enables ad targeting

- **RQ3: Differences across publishers**
  - Higher value of user tracking for
    - Non-premium publishers
    - Publishers with broader content
    - Large publishers


DELETE COOKIES?!
gegen-den-strich.com
FERNANDEZ

The Impact of Privacy Regulation on the Online Advertising Market

gdpr-impact.com

Laub, Miller, Skiera (2024), The Economic Value of User Tracking for Publishers.

Thank You for Your Attention!

Klaus Miller
millerk@hec.fr

Paper Download


2025
HEC
PARIS
2025