Competition and Privacy

Yifei Wang

University of Pittsburgh

January 9, 2025

Research Question

Does More Competition Lead to More or Less Privacy Intrusion?

Motivation

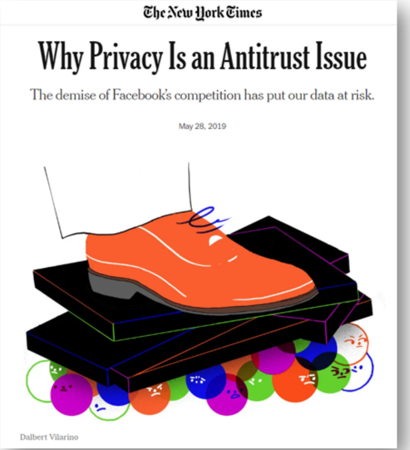
Privacy has become increasingly important in the mobile economy

- On average,[1] a mobile phone is located by apps 3691 times/day
- Photos and files data are accessed 2432 times/day

[1] Statistics published by Xiaomi privacy team, January 2021

Policy Debate Worldwide

- Regulators are concerned that insufficient competition leads to privacy abuse
  - US: DOJ vs. Google, 2020; FTC vs. Meta, 2021
  - EU: German antitrust regulators vs. Facebook, 2020



Policy Debate Worldwide

Regulators' Hypothesis

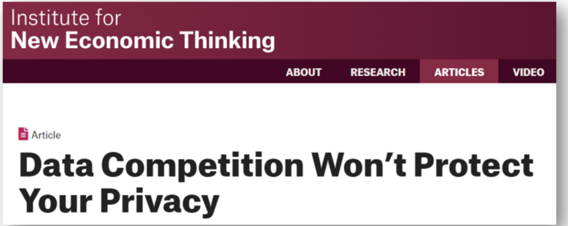
" Emboldened by the decline of market threats, Facebook revoked its users' ability to vote on changes to its privacy policies and then (almost simultaneously with Google's exit from the social media market) changed its privacy pact with users."

(The New York Times, 2019)

But Competition Might Not be the Cure

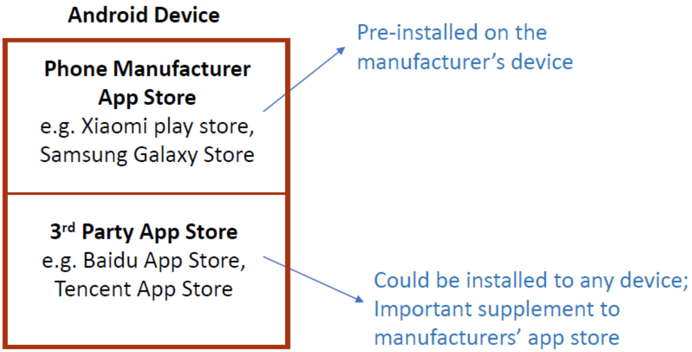
The Alternative Hypothesis

- “But as Breaking Away explores, more competition will not help when the competition itself is toxic. Here rivals compete to exploit us by discovering better ways to addict us, degrade our privacy, manipulate our behavior, and capture the surplus.” (Stucke, 2022)
- Large firms collect less sensitive data & invest more on privacy protection (Kummer and Schulte, 2019; Dulberg, 2021)



Outline

Chinese Android App Markets

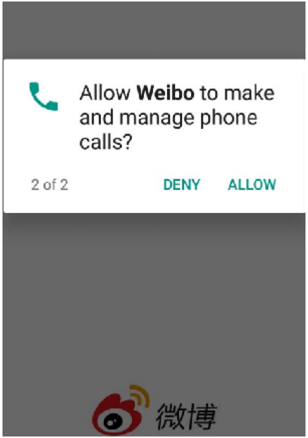


- App data from a major 3rd party Android app store (top 5)

Measuring Privacy

- Privacy intrusion: the number of permissions requested (Krafft et al., 2017; Kesler et al., 2017; Kummer and Schulte, 2019)
  - Apps need to request permissions to Android if they want to access users' data or control device functions
  - The number of permissions: the level of data access and device control of an app
  - Particularly 'dangerous' permissions defined by Android

- Permission $\neq$ user consent



Measuring Competition

- Focus on a policy shock to competition which
  - Reduced the number of available products in treated markets
  - Increased market concentration in treated markets

Outline

Identification Challenge

- Competition in a market is endogenous
  - Confounds affect both competition and firms' privacy intrusion
  - E.g. market size, product features

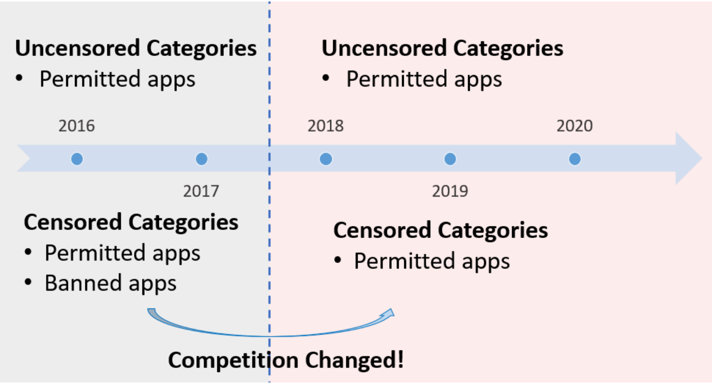
Identification Strategy

Internet Censorship and Censorship Circumvention

- In 2003, China launched the Golden Shield Project
  - Some of the most popular global websites and their apps are banned (Facebook, YouTube, NYT...)
- People can still use the banned apps with censorship circumvention tools (e.g. VPNs)
- In 2017, new regulation prohibits such tools not authorized by the government

Policy Change Affects Competition

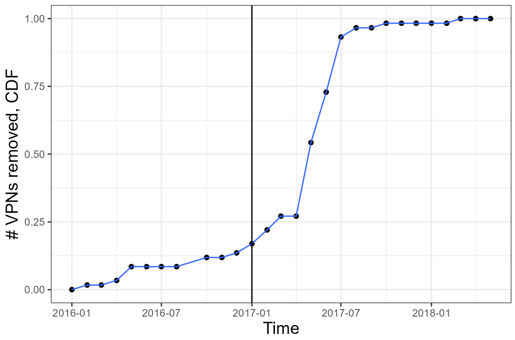
Defining Treated and Control



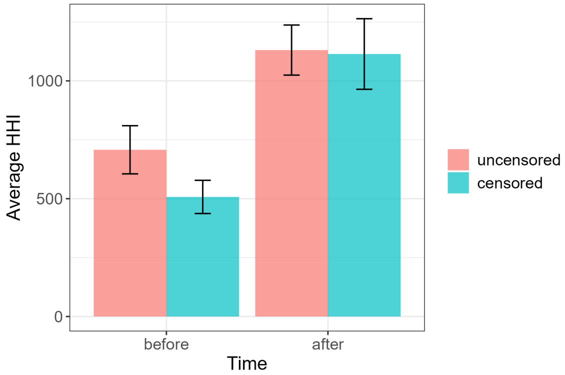
- Treated: permitted apps in censored categories
- Control: permitted apps in uncensored categories

Evidence of Policy Change

- Removal of popular VPNs



- Change in market concentration



Outline

Data

App Data

- A list of 17,001 apps available on the 3rd party Android app store
- Downloaded all historical versions 2014 - 2022 for apps in sample (327,734 App Installation packages)
- Permissions, app functions, and revenue model data from the App Installation Kit (APK) packages

Censorship Data

- Censorship status and blocking time data from greatfire.org
- A category is censored if ≥ 1 top non-Chinese apps were banned before 2017 (20 out of 87 categories)

Model: SynthDID

- Used the synthetic differences-in-differences design (SDID) (Arkhangelsky et al., 2021)
- SDID combines desirable features of DID and synthetic control

Variable Definition

$$(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \underset{\tau, \mu, \alpha, \beta}{\text{argmin}} \{\sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid}\}$$

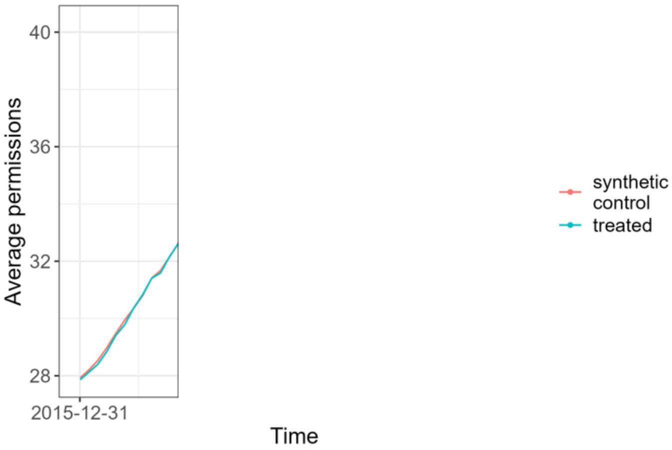
- Unit of analysis: a version of a *permitted* app (monthly)
- Outcome (Y_{it}): the number of permissions by app i at month t
- Treatment (W_{it}): 1 if an app is in a censored category *and* the observation was after the policy shock
- Controlled for app and month fixed effects (α_i, β_t), and unit and time weights ($\hat{\omega}_i$, $\hat{\lambda}_t$)

Summary statistics SDID weights

Outline

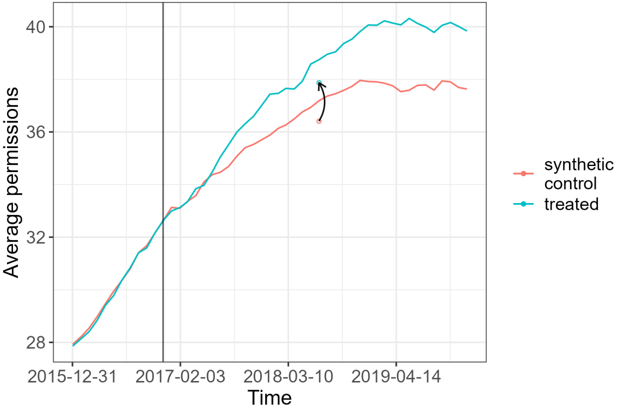
SynthDID Results

All Permissions, Before



SynthDID Results

Lower Competition Leads to More Intrusion



Numerical Estimates

Table: Estimated Treatment Effects

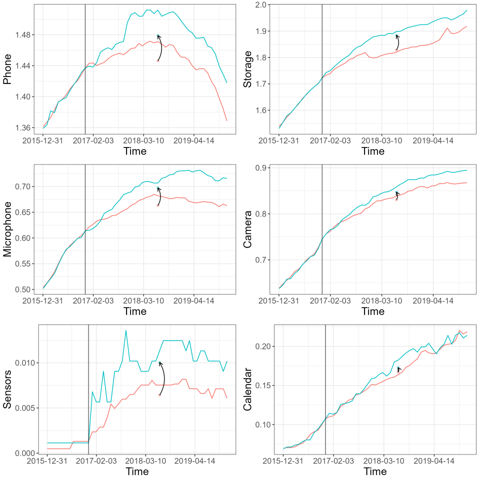
	DID	*Synth. DID*	*Synth. DID*
	(1)	(2)	(3)
	All Permissions	All Permissions	Dangerous Permissions
Treated	1.927***	1.463***	0.310***
Std. Error	(0.503)	(0.482)	(0.078)

Notes: Clustered standard error by market and by month is reported for the DID estimate. Jackknife standard errors are reported for synthetic DID estimates. The number of observations is 137,788.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

- Robustness on (a) dangerous permissions & (b) effect is due to competition

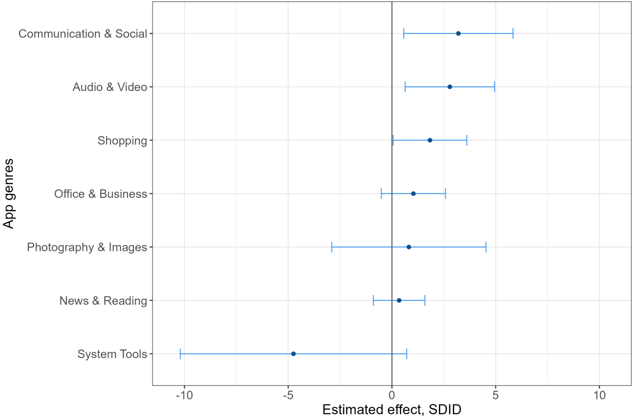
Examples of Affected Permissions

Effect Comes from Various Permissions...



Heterogeneity by App Genre

- Genres are defined by the app store



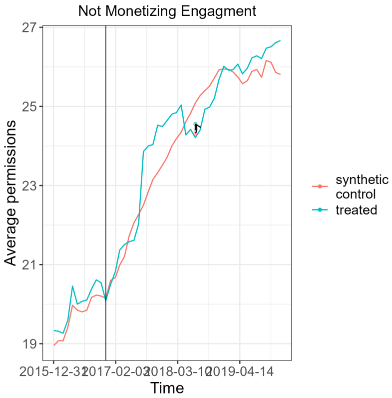
Mechanism: What Are the Increased Permissions for?

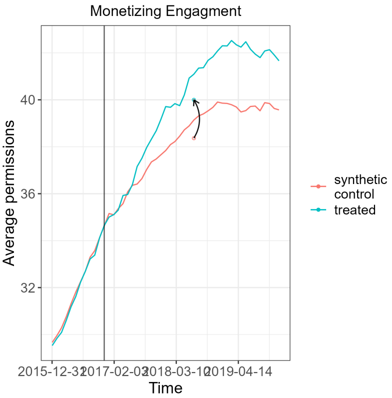
Intrusion for Engagement

- Analyzed each of the 150 most commonly used permissions
  - Treated apps increased requests in 114 of the 150 permissions
  - But some of the most affected permissions are specifically designed to engage users, especially through direct marketing
- This suggests treated apps substantially increased their effort to engage consumers

Does Increased Engagement Effort Explain the Effect?

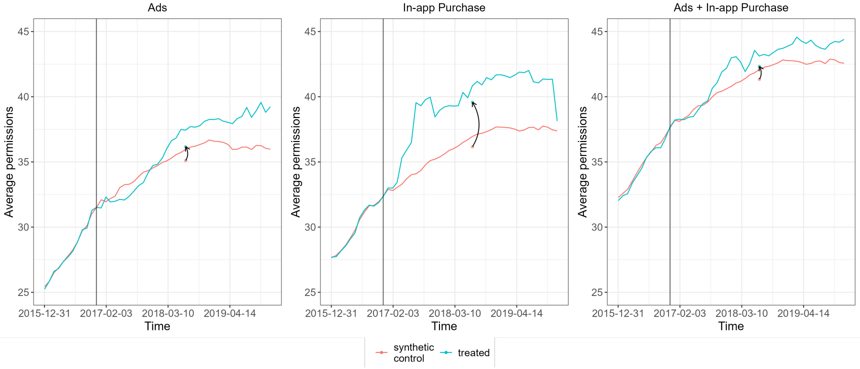
- If yes, effect should be larger when firms have stronger incentives to maximize engagement
- In particular, when engagement is profitable





Stratifying by Pre-treatment Monetization Model

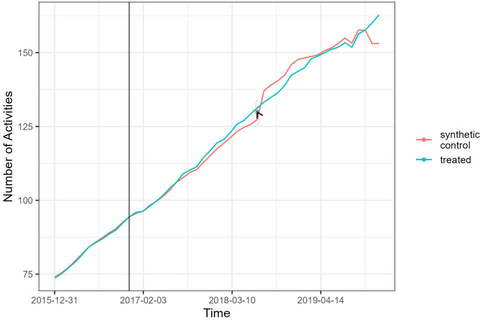
- The largest effect comes from the subsample with an in-app purchase model



What Does Not Explain the Effect?

Alternative Explanations

- Treated apps increased permissions to develop more functions? No.
- ... because they are facing more or different entrants? No. Alt.explanation: entrants
- ... because they use data to improve ad targeting? No. Alt.explanation: ads



(Est. effect = 0.19, p = 0.93)

To Summarize

- I study a policy shock that reduced competition in some app categories but not others
- I find that reducing competition leads to more privacy intrusion from treated apps
- This is one of the first empirical study on the relationship between market competition and firms' privacy invasion
- For regulators, this research suggests an important way to protect privacy is by restricting market power and encouraging competition

Thank you. Questions?

- Email: yiw386@pitt.edu
- Paper link: `https://shorturl.at/FDC9g`
- Paper QR code:

This Research in One Page

Does More Competition Lead to More or Less Privacy Intrusion?

- Context: Chinese Android app markets, 2016 - 2020
  - Privacy: permissions requested by apps
  - Market: a group of similar-functioned apps (e.g. news apps)
  - Competition: degree of concentration in a market
- Challenge: Competition is endogenous
- Solution:
  - Exogenous policy shock that reduced the competition in censored markets
  - Synthetic diff-in-diff: censored vs. uncensored markets over time
- Findings:
  - Decrease in market competition $\Rightarrow$ more privacy-intrusive behavior

Summary Statistics

Table: Distribution of Top 10 App Categories by Treatment Status

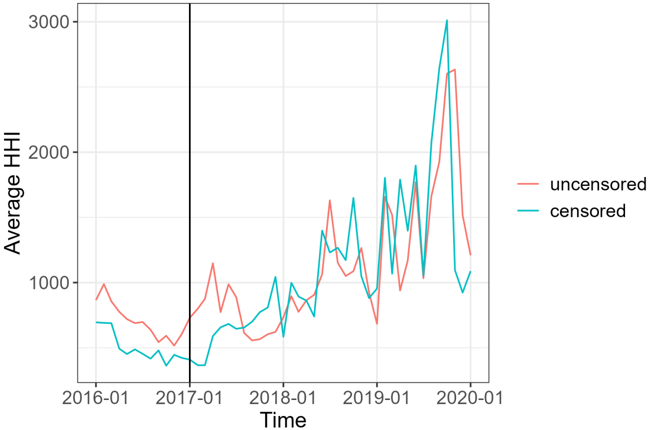
Group	Category	Observations	% within group
Treated	Productivity	41203	17.24%
Treated	E-shopping	38409	16.07%
Treated	Communities	25381	10.62%
Treated	Information	20932	8.76%
Treated	Videos	16271	6.81%
Treated	Chat	16209	6.78%
Treated	News	15905	6.65%
Treated	Novel	9315	3.90%
Treated	Music	8083	3.38%
Treated	Livestream	7976	3.34%
Control	Tools	41552	7.60%
Control	Learning	40112	7.34%
Control	Office software	29741	5.44%
Control	Examination	25918	4.74%
Control	Discounts	23353	4.27%
Control	Medical	23055	4.22%
Control	Early childhood education	22868	4.18%
Control	Games (children)	20829	3.81%
Control	Cars	17425	3.19%
Control	Car renting	16052	2.94%

Variable Definition

Backup Slide 1: Competition

Policy Change Affects Competition

- The average HHI of censored markets increased by 183.36 compared to the uncensored markets ‣ HHI bar chart



Backup Slide 2: SynthDID Weights

- Larger weights on control units that are 'similar' to treated units, and pre-treatment periods that are 'similar' to post-treatment periods

$$(\hat{\omega}_0, \hat{\omega}^{sdid}) = \underset{\omega_0 \in \mathbb{R}, \omega \in \Omega}{\operatorname{argmin}} \left\{ \sum_{t=1}^{T_{pre}} (\omega_0 + \sum_{i=1}^{N_{co}} \omega_i Y_{it} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} Y_{it})^2 + \zeta^2 T_{pre} \|\omega\|_2^2 \right\}, \tag{1}$$

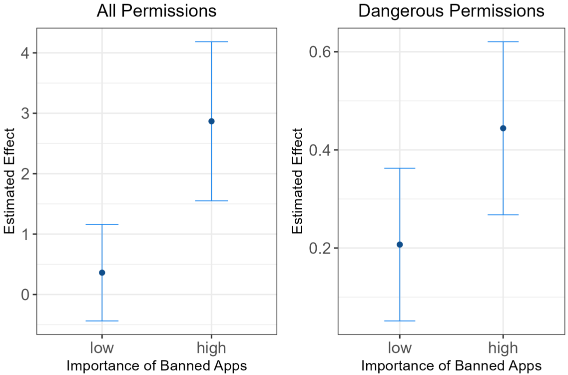
$$(\hat{\lambda}_0, \hat{\lambda}^{sdid}) = \underset{\lambda_0 \in \mathbb{R}, \lambda \in \Lambda}{\operatorname{argmin}} \left\{ \sum_{i=1}^{N_{co}} (\lambda_0 + \sum_{t=1}^{T_{pre}} \lambda_t Y_{it} - \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it})^2 \right\},$$

Variable Definition

Backup Slide 3: Competition

If this is really about competition, then...

- Categories where banned apps have larger market shares should be more affected by the shock ▸ Robustness checks
  - $\geq$ 2 banned apps in the top 1000 app rank in China in 2016
  - Video, communication, news... vs. emails, browsers, storage...



Backup Slide 4: Dangerous Permission Types

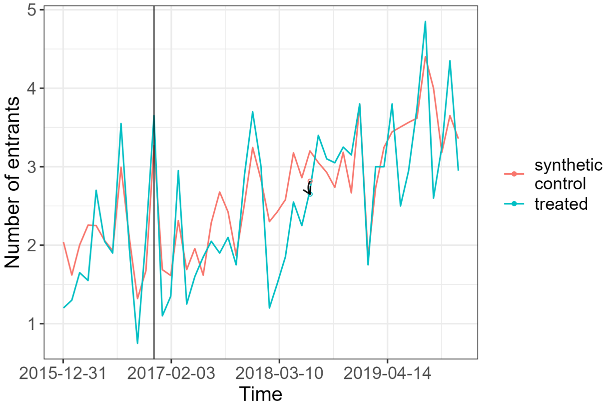
- Phone (7 permissions)
  - phone status, number, identifier
  - directly call phone numbers
- Storage (6 permissions)
  - access stored images and files from *other* apps

- Microphone (1 permission)
- Camera (1 permission)
- Body sensors (2 permissions)
- Calendar (2 permissions)
- Call log (3 permissions)
- SMS (6 permissions)
- Location (3 permissions)
- Contacts (3 permissions)

Backup Slide 5: Entry

Can the Effect of the Incumbents Affected by Entrants?

- The number of entrants is not significantly different
- The behavior of entrants is not significantly different



Backup Slide 6: Advertising

Is It Really Not about Advertising?

- The treatment does not affect whether an app uses an ad model, or the number of in-app ads.
- So treated apps did not increase how many ads they display.

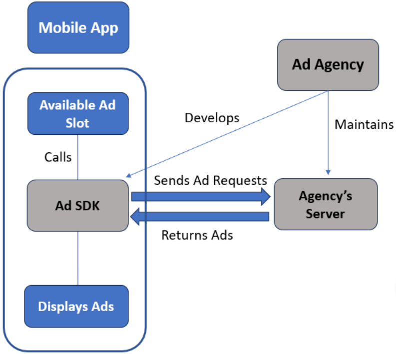
	(1) Has Ad	(2) Number of Ad Activities	(3) Permissions (Apps w/o Ads)
Treated	-0.008	-0.059	2.672**
Std. Error	(0.013)	(0.543)	(1.268)

Notes: N. Obs. is 137,788; 137,788; 101,577, respectively

Backup Slide 6: Advertising (Continued)

Is It Really Not about Advertising?

- But could they collect data to delivery better-targeted ads? No.
- The effect survives for treated apps that had no ads throughout.



Alternative Explanations

Why Does Competition Affect Firms' Incentives for Engagement?

A Theoretical Explanation

- An engaged consumer is more profitable with less competition
  - Intuitively: apps need to capture and monetize engagement
  - Apps have a stronger incentive to capture user engagement when it is more monetizable
  - Consistent with empirical findings on monetization

Arkhangelsky, D., S. Athey, D. A. Hirshberg, G. W. Imbens, and S. Wager (2021). Synthetic difference-in-differences. *American Economic Review 111*(12), 4088–4118.

Dulberg, R. (2021). Why the world's biggest brands care about privacy. *UX Collective*.

Kesler, R., M. E. Kummer, and P. Schulte (2017). Mobile applications and access to private data: The supply side of the android ecosystem. *ZEW-Centre for European Economic Research Discussion Paper* (17-075).

Krafft, M., C. M. Arden, and P. C. Verhoef (2017). Permission marketing and privacy concerns—why do customers (not) grant permissions? *Journal of interactive marketing 39*(1), 39–54.

Kummer, M. and P. Schulte (2019). When private information settles the bill: Money and privacy in google's market for smartphone applications. *Management Science 65*(8), 3470–3494.

Stucke, M. (2022). Data competition won't protect your privacy. *Institute for New Economic Thinking*.