Data and Competition

A Simple Framework

Alexandre de Cornière[1] & Greg Taylor[2]

[1]Toulouse School of Economics
[2]Oxford Internet Institute, University of Oxford

5 April 2024, JFTC Seminar

Roadmap

Introduction and motivating example

Competitive effects of data

Applications

Data-driven mergers

Introduction

Data is at the centre of the ongoing digital revolution.

- Much rhetoric about danger to competition.

Among concerns/questions:

- Exploitative behaviour: lack of privacy, (price-)discrimination.
- Data as barrier to entry / market tipping.
- How to deal with data-driven mergers?
- Data externalities.

Introduction

Data is at the centre of the ongoing digital revolution.

- Much rhetoric about danger to competition.

Among concerns/questions:

- Exploitative behaviour: lack of privacy, (price-)discrimination.
- Data as barrier to entry / market tipping.
- How to deal with data-driven mergers?
- Data externalities.

But, also potential efficiencies:

- product improvement, elimination of asymmetries, better matching, real-time feedback etc.

Introduction

Data is at the centre of the ongoing digital revolution.

- Much rhetoric about danger to competition.

Among concerns/questions:

- Exploitative behaviour: lack of privacy, (price-)discrimination.
- Data as barrier to entry / market tipping.
- How to deal with data-driven mergers?
- Data externalities.

But, also potential efficiencies:

- product improvement, elimination of asymmetries, better matching, real-time feedback etc.

In short: (when) is data a pro- or anti-competitive force?

- Many types of data, uses of data, and business models.

Motivating example

Suppose firm has a dataset of size δ.

- Data improves the firm's product: utility is $u = v(\delta) - p$.

Motivating example

Suppose firm has a dataset of size δ.

- Data improves the firm's product: utility is $u = v(\delta) - p$.
- Consumer demand is $D(u)$.

Motivating example

Suppose firm has a dataset of size δ.

- Data improves the firm's product: utility is $u = v(\delta) - p$.
- Consumer demand is $D(u)$.
- The firm therefore solves

$$\max_p \pi(p, \delta) = pD[v(\delta) - p].$$

Motivating example

Suppose firm has a dataset of size δ.

- Data improves the firm's product: utility is $u = v(\delta) - p$.
- Consumer demand is $D(u)$.
- The firm therefore solves

$$\max_p \pi(p, \delta) = pD[v(\delta) - p].$$

Question: if we give the firm more data, are consumers better or worse off?

- Two effects: quality effect and price effect.

Motivating example

Suppose firm has a dataset of size δ.

- Data improves the firm's product: utility is $u = v(\delta) - p$.
- Consumer demand is $D(u)$.
- The firm therefore solves

$$\max_{p} \pi(p, \delta) = pD[v(\delta) - p].$$

Question: if we give the firm more data, are consumers better or worse off?

- Two effects: quality effect and price effect.

We can answer this question by solving the firm's problem.

- But solution will be specific to this use of data/business model.
- Instead, let's look at another approach...

Motivating example, an alternative approach

Notice that

$$u = v(\delta) - p$$

Motivating example, an alternative approach

Notice that

$$u = v(\delta) - p \implies p = v(\delta) - u$$

Motivating example, an alternative approach

Notice that

$$u = v(\delta) - p \implies p = v(\delta) - u \equiv r(u, \delta).$$

Motivating example, an alternative approach

Notice that

$$u = v(\delta) - p \implies p = v(\delta) - u \equiv r(u, \delta).$$

Thus, we can reformulate the firm's problem as

$$\max_u \pi(u, \delta) = r(u, \delta) D(u).$$

Motivating example, an alternative approach

Notice that

$$u = v(\delta) - p \implies p = v(\delta) - u \equiv r(u, \delta).$$

Thus, we can reformulate the firm's problem as

$$\max_{u} \pi(u, \delta) = r(u, \delta) D(u).$$

Plan:

1. study the effects of δ in this setup (depend on shape of r).
2. study how different types and uses of data influence the shape of r.

Roadmap

Introduction and motivating example

Competitive effects of data

Applications

Data-driven mergers

Framework: competition in utilities

- $n \geq 1$ firms choose mean utility u_i.
  - price, quality, ad load,...
- Implied demand $D_i(u_i, \mathbf{u}_{-i})$.

Framework: competition in utilities

- $n \geq 1$ firms choose mean utility u_i.
  - price, quality, ad load,...
- Implied demand $D_i(u_i, \mathbf{u}_{-i})$.
- Associated mark-up: $r(u_i, \delta_i)$ (we often say "per-consumer revenue").

$$\frac{\partial r(u_i, \delta_i)}{\partial \delta_i} \geq 0$$

Framework: competition in utilities

- $n \geq 1$ firms choose mean utility u_i.
  - price, quality, ad load,...
- Implied demand $D_i(u_i, \mathbf{u}_{-i})$.
- Associated mark-up: $r(u_i, \delta_i)$ (we often say "per-consumer revenue").

$$\frac{\partial r(u_i, \delta_i)}{\partial \delta_i} \geq 0$$

- Fixed cost $C(u_i)$.
  - Can be zero (constant returns to scale).

Framework: competition in utilities

- $n \geq 1$ firms choose mean utility u_i.
  - price, quality, ad load,…
- Implied demand $D_i(u_i, \mathbf{u}_{-i})$.
- Associated mark-up: $r(u_i, \delta_i)$ (we often say "per-consumer revenue").

$$\frac{\partial r(u_i, \delta_i)}{\partial \delta_i} \geq 0$$

- Fixed cost $C(u_i)$.
  - Can be zero (constant returns to scale).
- Profit:

$$r(u_i, \delta_i) D_i(u_i, \mathbf{u}_{-i}) - C(u_i)$$

Competitive effect of data

FOC:

$$\frac{\partial \pi_i}{\partial u_i} = 0.$$

This gives us firm i's best-response function.

Competitive effect of data

FOC:

$$\frac{\partial \pi_i}{\partial u_i} = 0.$$

This gives us firm i's best-response function.

Is data pro- or anti-competitive?

- Comparative statics exercise: how does increase in δ_i change i's choice of u_i?
- Firm i's best-response shifts up $\iff \frac{\partial^2 \pi_i}{\partial u_i \partial \delta_i} > 0$.

Competitive effect of data

$$\frac{\partial^2 \pi_i}{\partial u_i \partial \delta_i} > 0 \iff \frac{\partial r}{\partial \delta_i}\frac{\partial D_i}{\partial u_i} + \frac{\partial^2 r}{\partial u_i \partial \delta_i} D_i > 0$$

We say data is *unilaterally pro-competitive* (UPC) if this condition is satisfied.

Data is *unilaterally anticompetitive* (UAC) if the inequality is reversed.

Competitive effect of data

$$\frac{\partial r}{\partial \delta_i}\frac{\partial D_i}{\partial u_i} + \frac{\partial^2 r}{\partial u_i \partial \delta_i} D_i > 0$$

Interpretation, part 1: mark-up effect

- $\frac{\partial r_i}{\partial \delta_i}\frac{\partial D_i}{\partial u_i} > 0.$
- Data makes marginal consumer more valuable.
- $\implies$ Provides extra incentive to compete.

Competitive effect of data

$$\frac{\partial r}{\partial \delta_i}\frac{\partial D_i}{\partial u_i} + \frac{\partial^2 r}{\partial u_i \partial \delta_i} D_i > 0$$

Interpretation, part 2: surplus-extraction effect

- Sign of $\frac{\partial^2 r_i}{\partial u_i \partial \delta_i}$ is not *a priori* obvious.
- Reflects how data changes the opportunity cost of providing utility.
- E.g., more data $\implies$ better ad targeting $\implies$ more costly to reduce ad load.

Competitive effect of data

$$\frac{\partial r}{\partial \delta_i}\frac{\partial D_i}{\partial u_i} + \frac{\partial^2 r}{\partial u_i \partial \delta_i} D_i > 0$$

Interpretation, part 2: surplus-extraction effect

- Sign of $\frac{\partial^2 r_i}{\partial u_i \partial \delta_i}$ is not *a priori* obvious.
- Reflects how data changes the opportunity cost of providing utility.
- E.g., more data $\implies$ better ad targeting $\implies$ more costly to reduce ad load.

Observation: if $\frac{\partial^2 r_i}{\partial u_i \partial \delta_i} \geq 0$ then data is UPC (sufficient condition).

Competitive effects of data

We can make more headway if $C'(u_i) = 0$.

Competitive effects of data

We can make more headway if $C'(u_i) = 0$.
The FOC simplifies to:

$$\frac{\partial r}{\partial u_i} D_i + r_i \frac{\partial D_i}{\partial u_i} = 0.$$

UPC condition:

$$\frac{\partial r}{\partial \delta_i} \frac{\partial D_i}{\partial u_i} + \frac{\partial^2 r}{\partial u_i \partial \delta_i} D_i > 0.$$

Competitive effects of data

We can make more headway if $C'(u_i) = 0$.
The FOC simplifies to:

$$\frac{\partial r}{\partial u_i} D_i + r_i \frac{\partial D_i}{\partial u_i} = 0.$$

UPC condition:

$$\frac{\partial r}{\partial \delta_i} \frac{\partial D_i}{\partial u_i} + \frac{\partial^2 r}{\partial u_i \partial \delta_i} D_i > 0.$$

Combining these two expressions eliminates the terms involving D and yields a new UPC condition:

$$-\frac{\partial r}{\partial u_i} \frac{\partial r}{\partial \delta_i} + \frac{\partial^2 r}{\partial u_i \partial \delta_i} r > 0$$

Competitive effects of data

We can make more headway if $C'(u_i) = 0$.
The FOC simplifies to:

$$\frac{\partial r}{\partial u_i} D_i + r_i \frac{\partial D_i}{\partial u_i} = 0.$$

UPC condition:

$$\frac{\partial r}{\partial \delta_i} \frac{\partial D_i}{\partial u_i} + \frac{\partial^2 r}{\partial u_i \partial \delta_i} D_i > 0.$$

Combining these two expressions eliminates the terms involving D and yields a new UPC condition:

$$-\frac{\partial r}{\partial u_i} \frac{\partial r}{\partial \delta_i} + \frac{\partial^2 r}{\partial u_i \partial \delta_i} r > 0 \iff \frac{\partial^2 \ln[r(u_i, \delta_i)]}{\partial u_i \partial \delta_i} > 0$$

(a necessary and sufficient condition).

Competitive effects of data

Sufficient condition

Data is UPC if r is supermodular.

Necessary and sufficient condition ($C'(u_i) = 0$)

Data is UPC if and only if r is log-supermodular.

In many cases ($C'(u_i) = 0$, or $\frac{\partial^2 r_i}{\partial u_i \partial \delta_i} \geq 0$), no information about D required to check whether data is pro or anti-competitive.

- What matters is the technology by which revenue is extracted from data, summarised by r.

Competitive effects of data

Sufficient condition

Data is UPC if r is supermodular.

Necessary and sufficient condition ($C'(u_i) = 0$)

Data is UPC if and only if r is log-supermodular.

In many cases ($C'(u_i) = 0$, or $\frac{\partial^2 r_i}{\partial u_i \partial \delta_i} \geq 0$), no information about D required to check whether data is pro or anti-competitive.

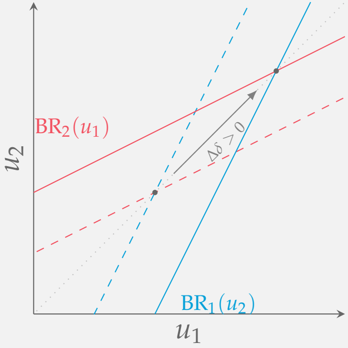
- What matters is the technology by which revenue is extracted from data, summarised by r.

Equilibrium: symmetric environment

If we assume firms are symmetric (including $\delta_i = \delta_{-i}$), these unilateral results extend to equilibrium.

Equilibrium: symmetric environment

If we assume firms are symmetric (including $\delta_i = \delta_{-i}$), these unilateral results extend to equilibrium. Two-firm example:

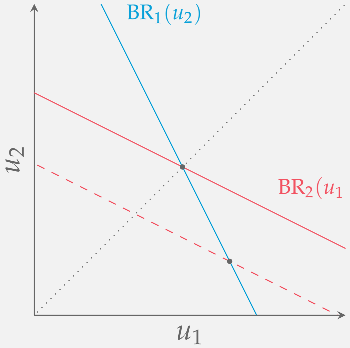


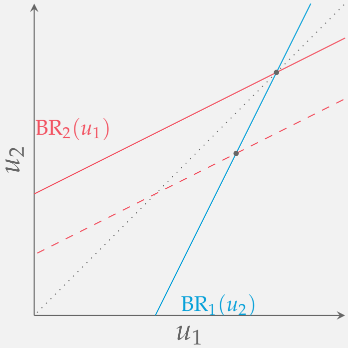
When $\delta_i \equiv \delta$, data increases equilibrium utility offers if and only if it is UPC.

Equilibrium with asymmetries

If firms are asymmetric, then

1. effect for focal firm given by UPC/UAC,
2. effect for its rivals determined by strategic complementarity/substitutability.





Equilibrium and congruence/conflict

Payoffs are strategic complements iff

$$0 < \frac{\partial^2 \pi_i}{\partial u_i \partial u_j}$$

Equilibrium and congruence/conflict

Payoffs are strategic complements iff

$$0 < \frac{\partial^2 \pi_i}{\partial u_i \partial u_j} = r(u_i)\frac{\partial^2 D(u_i, \mathbf{u}_{-i})}{\partial u_i \partial u_j} + \frac{\partial r(u_i)}{\partial u_i}\frac{\partial D(u_i, \mathbf{u}_{-i})}{\partial u_j}.$$

Equilibrium and congruence/conflict

Payoffs are strategic complements iff

$$0 < \frac{\partial^2 \pi_i}{\partial u_i \partial u_j} = r(u_i)\frac{\partial^2 D(u_i, \mathbf{u}_{-i})}{\partial u_i \partial u_j} + \frac{\partial r(u_i)}{\partial u_i}\frac{\partial D(u_i, \mathbf{u}_{-i})}{\partial u_j}.$$

Useful result in Hotelling duopoly: $D(u_i, u_{-i}) = \frac{t+u_i-u_J}{2t} \implies \frac{\partial^2 D(u_i, \mathbf{u}_{-i})}{\partial u_i \partial u_j} = 0.$

Equilibrium and congruence/conflict

Payoffs are strategic complements iff

$$0 < \frac{\partial^2 \pi_i}{\partial u_i \partial u_j} = r(u_i)\frac{\partial^2 D(u_i, \mathbf{u}_{-i})}{\partial u_i \partial u_j} + \frac{\partial r(u_i)}{\partial u_i}\frac{\partial D(u_i, \mathbf{u}_{-i})}{\partial u_j}.$$

Useful result in Hotelling duopoly: $D(u_i, u_{-i}) = \frac{t+u_i-u_J}{2t} \implies \frac{\partial^2 D(u_i,\mathbf{u}_{-i})}{\partial u_i \partial u_j} = 0$.

- u_1 & u_2 are strategic complements iff $r'(u_i) < 0$ (*conflict*).

Equilibrium and congruence/conflict

Payoffs are strategic complements iff

$$0 < \frac{\partial^2 \pi_i}{\partial u_i \partial u_j} = r(u_i)\frac{\partial^2 D(u_i, \mathbf{u}_{-i})}{\partial u_i \partial u_j} + \frac{\partial r(u_i)}{\partial u_i}\frac{\partial D(u_i, \mathbf{u}_{-i})}{\partial u_j}.$$

Useful result in Hotelling duopoly: $D(u_i, u_{-i}) = \frac{t+u_i-u_J}{2t} \implies \frac{\partial^2 D(u_i,\mathbf{u}_{-i})}{\partial u_i \partial u_j} = 0.$

- u_1 & u_2 are strategic complements iff $r'(u_i) < 0$ (*conflict*).

If δ_i increases:

	Data	
Payoffs	UAC	UPC
Conflicting	$\downarrow u_i^*, \downarrow u_j^*$	$\uparrow u_i^*, \uparrow u_j^*$
Congruent	$\downarrow u_i^*, \uparrow u_j^*$	$\uparrow u_i^*, \downarrow u_j^*$

Taking stock

Some immediate implications:

- Policies like mandated data sharing effective if data is UPC.

Taking stock

Some immediate implications:

- Policies like mandated data sharing effective if data is UPC.
- Data a barrier to entry if and only if it is UPC.
  - Implies a short-run/long-run tension.

Taking stock

Some immediate implications:

- Policies like mandated data sharing effective if data is UPC.
- Data a barrier to entry if and only if it is UPC.
  - Implies a short-run/long-run tension.
- Data externalities are negative if data is UAC.

Taking stock

Some immediate implications:

- Policies like mandated data sharing effective if data is UPC.
- Data a barrier to entry if and only if it is UPC.
  - Implies a short-run/long-run tension.
- Data externalities are negative if data is UAC.

Now let's look at some applications:

- Product improvement.
- Targeted advertising.
- Price-discrimination.
- Contract design.

Roadmap

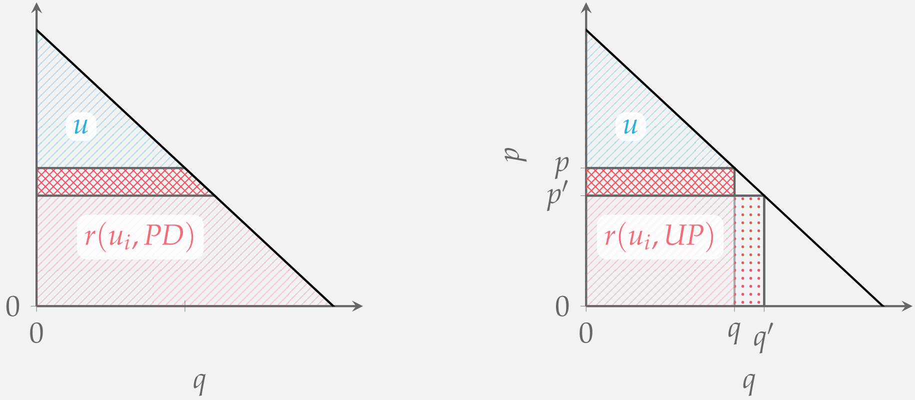
Product improvement

- In the motivating example we had $r(u,\delta) = v(\delta) - u$.
- It is immediate that $\frac{\partial^2 r}{\partial u \partial \delta} = 0 \implies$ data is UPC.
- This is an example where the surplus extraction effect is inactive because the firm can extract surplus efficiently through the price for any δ.

Price-discrimination

Multi-product firms, one-stop shoppers.

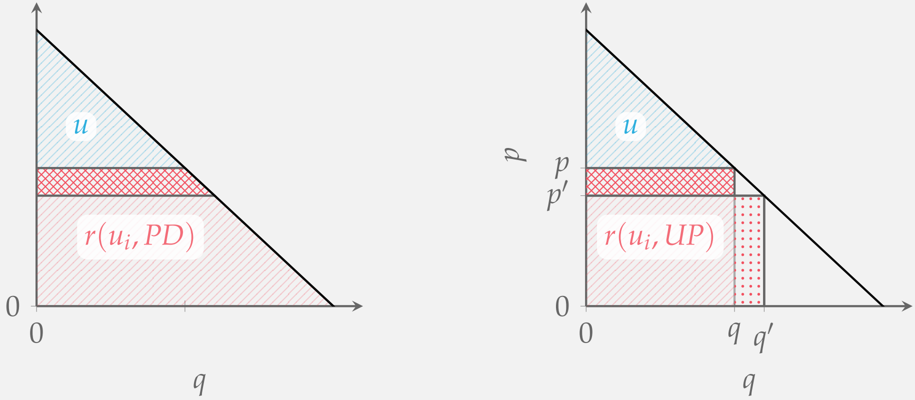
- Uniform list price, plus personalised discount for δ products.



Price-discrimination

Multi-product firms, one-stop shoppers.

- Uniform list price, plus personalised discount for δ products.



Surplus extraction effect is negative: we use the log supermodularity condition to show data is UAC in this kind of environment.

Contract design

- Consumer faces a risk and can exert effort $e \in \{0,1\}$ in the hope of avoiding it.
- Consider an insurer that can generate a signal that proves high consumer effort w.p. δ.

Contract design

- Consumer faces a risk and can exert effort $e \in \{0, 1\}$ in the hope of avoiding it.
- Consider an insurer that can generate a signal that proves high consumer effort w.p. δ.
- Insurance contract, $\mathcal{C} = \{p, X_+, X_-\}$, where X_+ is the amount reimbursed after a positive signal, X_- after a negative signal.
- Contract gives consumers utility $U(\mathcal{C}, e)$ for effort level e.

Contract design

- Consumer faces a risk and can exert effort $e \in \{0, 1\}$ in the hope of avoiding it.
- Consider an insurer that can generate a signal that proves high consumer effort w.p. δ.
- Insurance contract, $\mathcal{C} = \{p, X_+, X_-\}$, where X_+ is the amount reimbursed after a positive signal, X_- after a negative signal.
- Contract gives consumers utility $U(\mathcal{C}, e)$ for effort level e.
- Firm's revenue:

$$r(u, \delta) = \max_{\mathcal{C}} \{p - (1-\alpha)(\delta X_+ + (1-\delta)X_-)\}$$

$$\text{s.t.} \quad U(\mathcal{C}, 1) = u \quad \text{and} \quad U(\mathcal{C}, 1) \geq U(\mathcal{C}, 0).$$

Contract design

- Consumer faces a risk and can exert effort $e \in \{0,1\}$ in the hope of avoiding it.
- Consider an insurer that can generate a signal that proves high consumer effort w.p. δ.
- Insurance contract, $\mathcal{C} = \{p, X_+, X_-\}$, where X_+ is the amount reimbursed after a positive signal, X_- after a negative signal.
- Contract gives consumers utility $U(\mathcal{C}, e)$ for effort level e.
- Firm's revenue:

$$r(u,\delta) = \max_{\mathcal{C}} \{p - (1-\alpha)(\delta X_+ + (1-\delta)X_-)\}$$

$$\text{s.t.} \quad U(\mathcal{C},1) = u \quad \text{and} \quad U(\mathcal{C},1) \geq U(\mathcal{C},0).$$

- Can show that for CARA preferences, r is supermodular (surplus extraction effect is positive).
  - Data is UPC because it *reduces* the opportunity cost of utility.

Targeted advertising 1

- Suppose data is used to improve the targeting of ads.
- Implies inverse ad demand $P(a, \delta)$.

Targeted advertising 1

- Suppose data is used to improve the targeting of ads.
- Implies inverse ad demand $P(a, \delta)$.
- If consumers dislike ads then $u'(a) < 0$.

Targeted advertising 1

- Suppose data is used to improve the targeting of ads.
- Implies inverse ad demand $P(a, \delta)$.
- If consumers dislike ads then $u'(a) < 0$. Can write $a(u)$ for the ad load that gives utility u.

Targeted advertising 1

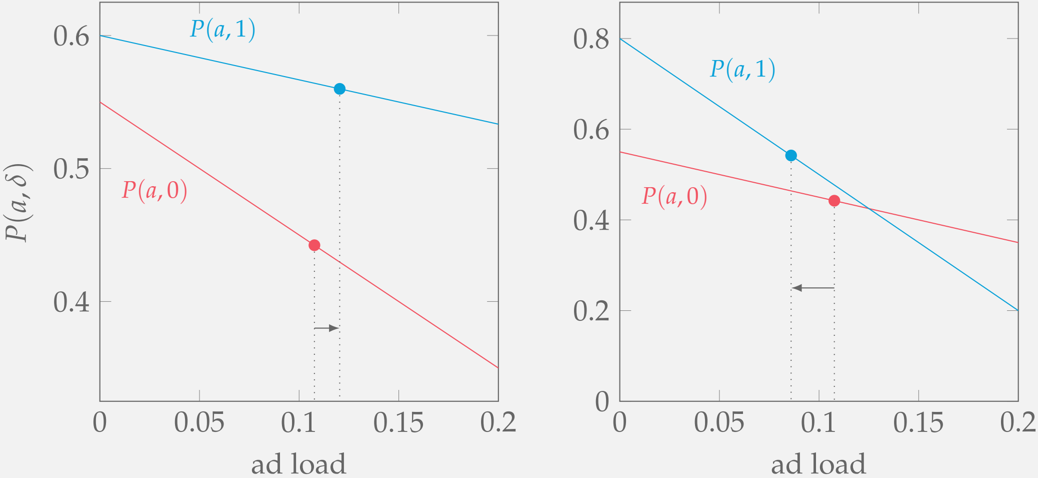
- Suppose data is used to improve the targeting of ads.
- Implies inverse ad demand $P(a, \delta)$.
- If consumers dislike ads then $u'(a) < 0$. Can write $a(u)$ for the ad load that gives utility u.
- $r(u, \delta) = a(u)P(a(u), \delta)$.

Targeted advertising 1

- Suppose data is used to improve the targeting of ads.
- Implies inverse ad demand $P(a, \delta)$.
- If consumers dislike ads then $u'(a) < 0$. Can write $a(u)$ for the ad load that gives utility u.
- $r(u, \delta) = a(u)P(a(u), \delta)$.
- Effects depend on the ad technology (i.e., how targeting affects P).
- Using the log-supermodularity condition, data is UPC if and only if it makes P more elastic.

Targeted advertising 2

Data can be informative about a category (left) or a brand's (right) match with a consumer.



Roadmap

Introduction and motivating example

Competitive effects of data

Applications

Data-driven mergers

Data-driven mergers

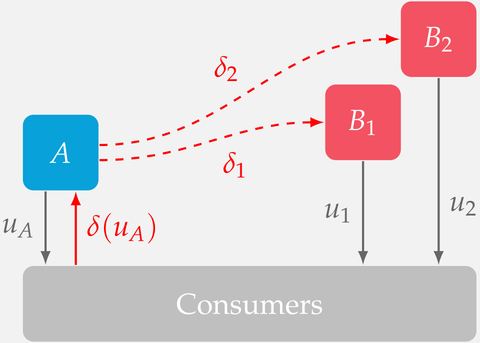
Several recent high-profile mergers:

- Facebook/Instagram/WhatsApp
- Microsoft/LinkedIn
- Google/Fitbit

Features

- Data obtained as by-product of activity in one market (IG, WA,LI)
- Used in other market (FB: targeted ads, Msft: personalized CRM software)

Data-connected markets



Data-driven mergers

In a companion paper (de Cornière and Taylor, forthcoming ManSci), we use our framework to study such mergers.

- UAC/UPC condition tells us the effect of data in market B.
- Merger also affects A's incentives to gather data, implying welfare effects in both markets.

Data-driven merger, no trade

Suppose that data trade is NOT possible absent merger.

- After merger, A internalises the value of data for $B1$: $\frac{\partial \pi_A(\delta)}{\partial \delta} + \frac{\partial \pi_{B1}(\delta,0)}{\partial \delta_1} = 0$.
- A thus collects more data: u_A must go up.
- Effect of merger in market B is positive if data is UPC, negative if UAC.

Data-driven merger, trade

Merger

- Marginal incentive to collect data: $\frac{\partial \pi_A(\delta)}{\partial \delta} + \frac{\partial \pi_{B1}(\delta,0)}{\partial \delta_1}$

Data-driven merger, trade

Merger

- Marginal incentive to collect data: $\frac{\partial \pi_A(\delta)}{\partial \delta} + \frac{\partial \pi_{B1}(\delta,0)}{\partial \delta_1}$

No merger

- Value of data = price at which it is sold to B firms

Data-driven merger, trade

Merger

- Marginal incentive to collect data: $\frac{\partial \pi_A(\delta)}{\partial \delta} + \frac{\partial \pi_{B1}(\delta,0)}{\partial \delta_1}$

No merger

- Value of data = price at which it is sold to B firms
- Price $= \pi_B(\delta, 0) - \pi_B(0, \delta)$
- Marginal incentive to collect data: $\frac{\partial \pi_A(\delta)}{\partial \delta} + \frac{\partial \pi_{B1}(\delta,0)}{\partial \delta_1} - \frac{\partial \pi_{B1}(0,\delta)}{\partial \delta_2}$

Data-driven merger, trade

Merger

- Marginal incentive to collect data: $\frac{\partial \pi_A(\delta)}{\partial \delta} + \frac{\partial \pi_{B1}(\delta,0)}{\partial \delta_1}$

No merger

- Value of data = price at which it is sold to B firms
- Price $= \pi_B(\delta, 0) - \pi_B(0, \delta)$
- Marginal incentive to collect data: $\frac{\partial \pi_A(\delta)}{\partial \delta} + \frac{\partial \pi_{B1}(\delta,0)}{\partial \delta_1} - \frac{\partial \pi_{B1}(0,\delta)}{\partial \delta_2}$
- If data is UPC then $\frac{\partial \pi_{B1}(0,\delta)}{\partial \delta_2} < 0$.
  - Stronger incentive to collect data without merger.
  - The opposite if data is UAC.

Summary effects of the merger

	data is UPC	data is UAC
Pre-merger data trade	$\downarrow u_A, \downarrow u_B$	$\uparrow u_A, \downarrow u_B$
No pre-merger data trade	$\uparrow u_A, \uparrow u_B$	$\uparrow u_A, \downarrow u_B$

Conclusion

- Data is at the centre of a fierce policy debate in tech.
- But competitive implications are ambiguous.
- We show that these can be understood through a simple condition, often without knowing about demand.
- Applications to various markets.
- These insights can also inform our understanding of merger policy and market structure.

Conclusion

- Data is at the centre of a fierce policy debate in tech.
- But competitive implications are ambiguous.
- We show that these can be understood through a simple condition, often without knowing about demand.
- Applications to various markets.
- These insights can also inform our understanding of merger policy and market structure.

Link to paper:
`https://drive.google.com/open?id=1p0mZDc5sEKa_Iz3tzA_wQJhiRK2zc5Pv`

Link to data-drive mergers companion paper:
`https://drive.google.com/file/d/1DqqtKiH8Vw-a7NGnaA-qH3SSix3JAkBt`