A Framework for Detection, Measurement, and Welfare Analysis of Platform Bias

Imke Reimers
Cornell University

Joel Waldfogel
University of Minnesota

DMA and Beyond

February 6, 2025

DMA: “Self-preferencing” is now illegal

- Regulatory action is ahead of research
- Urgent need to detect and measure welfare consequences of self-preferencing
- But: identifying *unwarranted* self-preferencing is not straightforward

- **Some basic questions:**
    - What is self-preferencing?
    - How do we detect it?
    - What is its welfare cost?

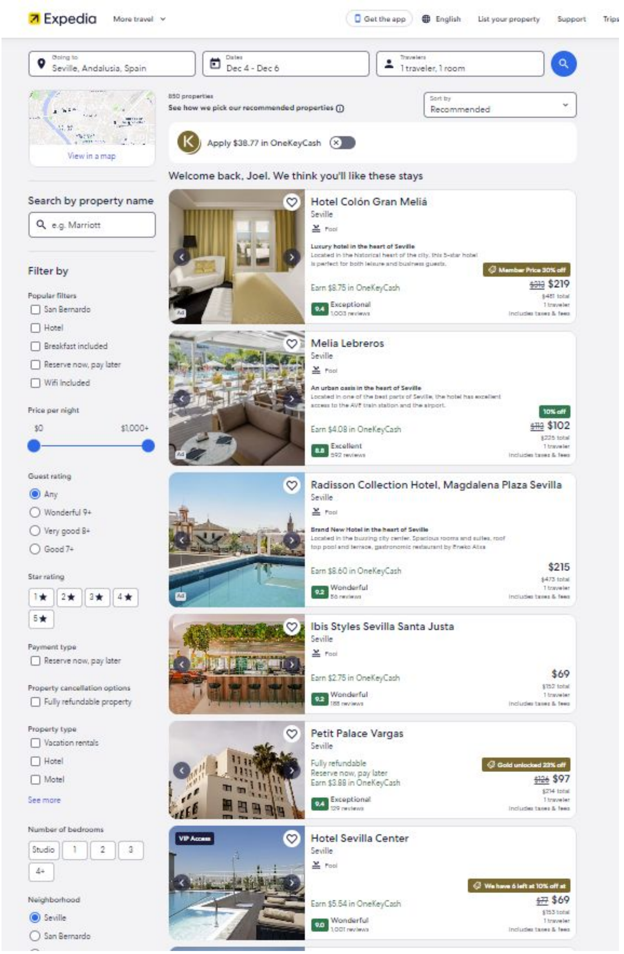
Generic setup: search result rankings

- Users choose among (ranked) product lists

- The platform chooses ranks to serve an objective
    - Consumers, sellers, the platform itself

- <u>One possible definition of self-preferencing:</u>

- The platform ranks its own products higher than would maximize some combination of seller and consumer surplus



Roadmap

- A simple theoretical framework
- Use the theory to compare 2 methods for bias detection
  - Conditioning on observables (COO) vs outcome-based (OB) tests
- Data and empirical comparison
  - …confirming conflicts between COO and OB
- Structural model and estimates of rank bias and welfare effects
  - … meaningful differences across settings

Model

Model

- Two parts:
  - <u>Consumers</u> choose among ranked products; better rankings → higher purchase probability
  - <u>The platform</u> chooses among $N!$ possible rank orderings

- The platform decides:
  - a) how to balance interests of consumers and sellers, and
  - b) how much to advance its own interest at the expense of consumers & sellers

- *Without self-preferencing:*
  - Search rankings lead to a welfare frontier between max CS and max PS

- *With platform bias (e.g., self-preferencing):*
  - rankings depart from the frontier

Implementation needs and model choice

- Need a way to map <u>product characteristics</u>, <u>prices</u> and <u>platform-chosen ranks</u> into **quantities sold**, **total revenue**, and **CS** for the choice set

- Various possible demand approaches
  - Search models (Ursu, Seiler, Honka (2023))
  - Limited information choice models (Goeree (2008), Abaluck and Adams-Prassl (2021))

- Here, we want to illustrate, simply
  - Logit (and NL)

Consumer side

- Consumer i chooses among J ranked products

$$u_{ij} = \delta_j^0 + \underbrace{\boxed{\gamma r_j}}_{\text{causal rank effect}} + \epsilon_{ij}$$

(Or the outside good, $u_{i0} = 0$)

Note: δ_j^0 is related to r_j beyond the causal effect of platform rank choice: the platform ranks better products higher

- Mean utility: $\bar{u}_j = \delta_j^0 + \gamma r_j$

- *Rank-independent* mean utility: δ_j^0

Outcomes depend on ranking R via $e^{\gamma r_j}$ term

- Probability of purchasing product j:

$$s_j(R) = \frac{e^{\delta_j^0} e^{\gamma r_j}}{1 + \Sigma e^{\delta_j^0} e^{\gamma r_j}}$$

- Gross seller surplus across products:

$$PS = PS(R)$$

- Consumer surplus:

$$CS = CS(R)$$

- (No evidence of price changes when rank algos change)
- (still, we allow for price changes later)

The platform's ranking choice

- Big combinatoric problem (N! choices)
  - See Compiani, Lewis, Peng, Wang (2021)

- Simplify, starting with two welfare frontier extremes:
- a) <u>Maximize CS</u>: rank in descending order of rank-independent mean utility δ_j^0
- b) <u>Maximize PS</u>: rank by rank-independent var. profit $(p_j - c_j)e^{\delta_j^0}$

→ The welfare frontier comes from weighted sums of these two

For exposition, assume $c = 0$

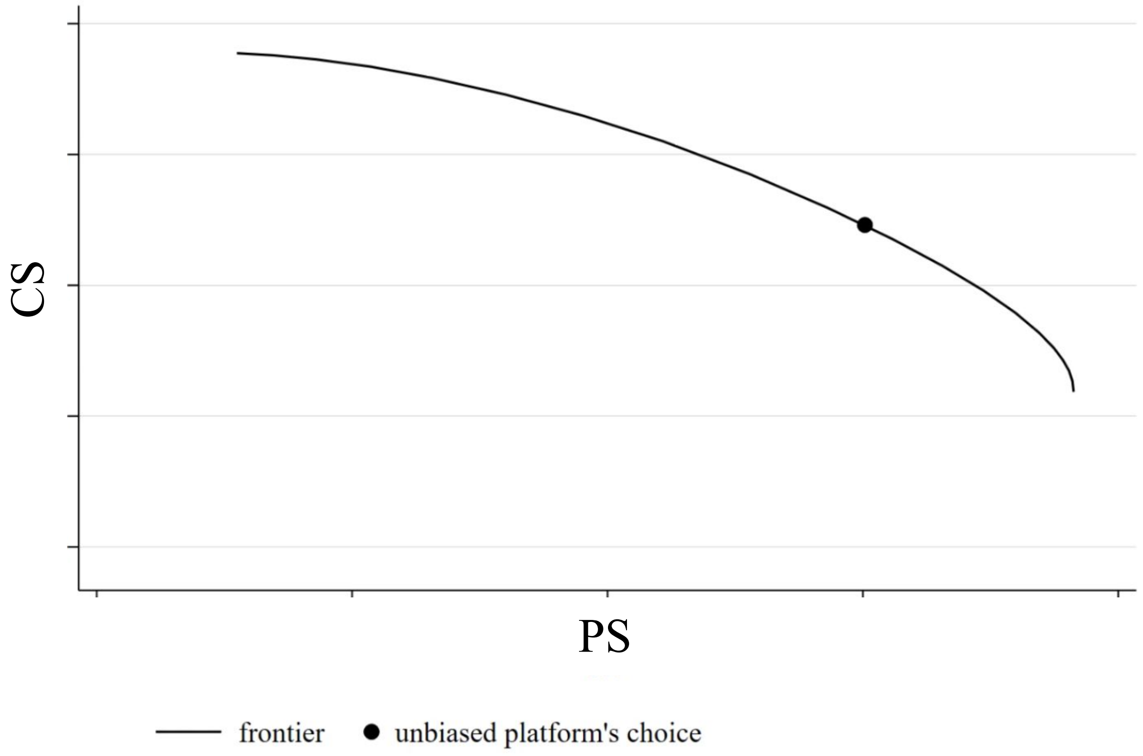
Welfare frontier

- Ranks according to $p_j e^{\delta_j^0}$ maximize revenue (PS)
- Ranks according to $e^{\delta_j^0}$ maximize CS

- Hence, define a frontier based on

$$I_j^* = \kappa_1 \ln(p_j) + \kappa_2 \delta_j^0$$

- Endpoints
  - $\kappa_1 = \kappa_2 > 0 \Leftrightarrow$ PS max
  - $\kappa_1 = 0, \kappa_2 > 0 \Leftrightarrow$ CS max



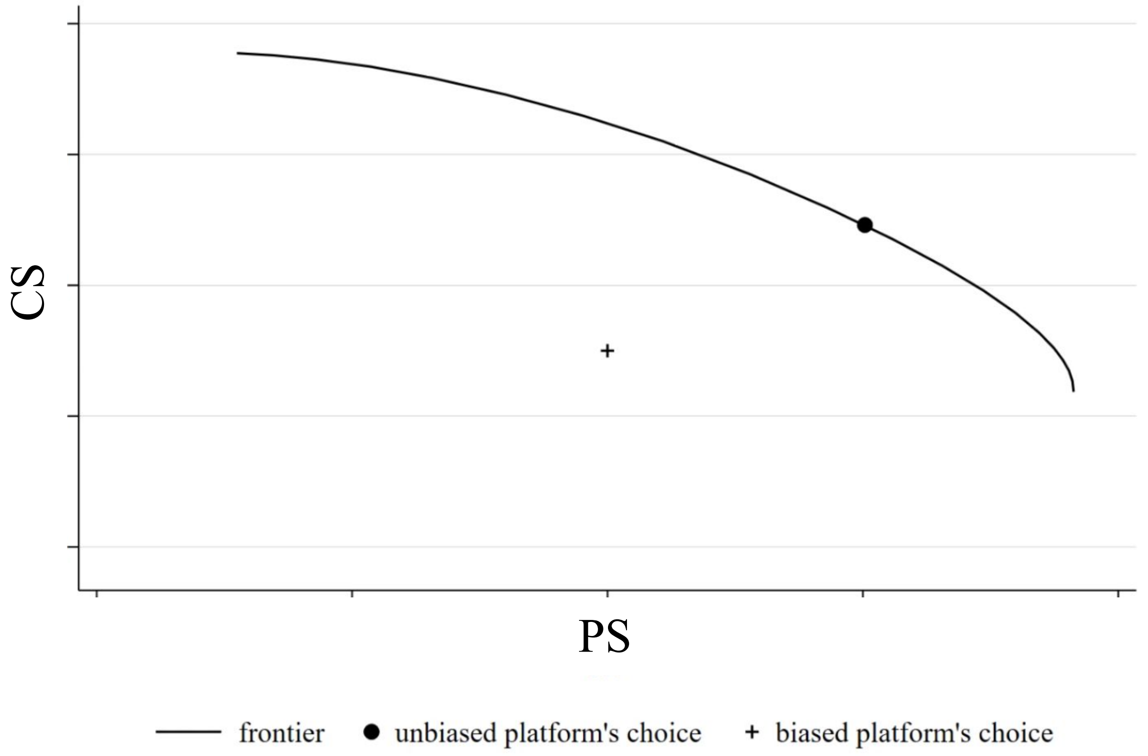
Add possible bias

- Ranks according to $p_j e^{\delta_j^0}$ maximize revenue (PS)
- Ranks according to $e^{\delta_j^0}$ maximize CS
- **Possible bias**

- Hence, the platform supply function:

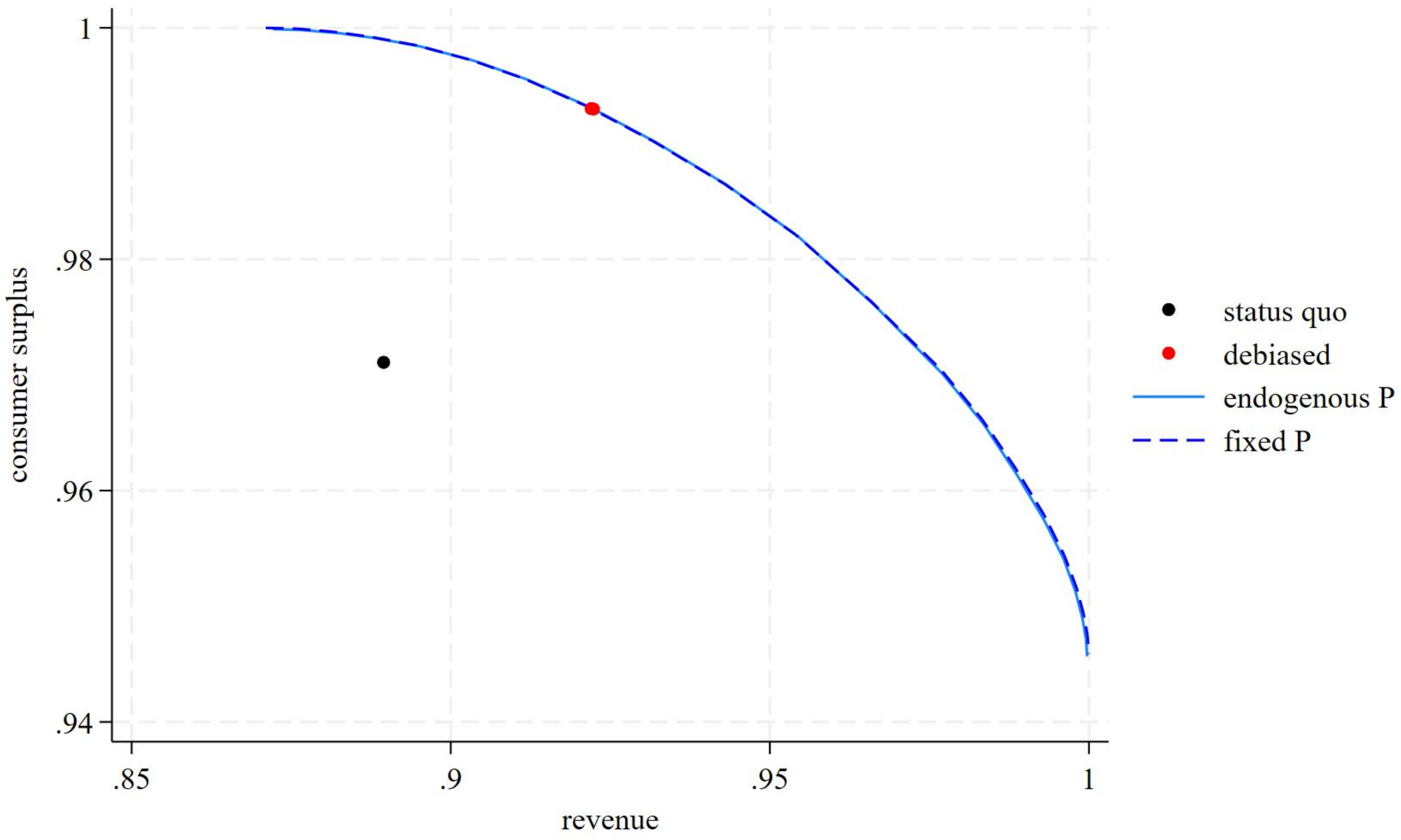
$$I_j = \kappa_1 \ln(p_j) + \kappa_2 \delta_j^0 + \boxed{\psi D_j}$$

- $\psi \neq 0$ is bias, and it changes ranking
- Delivers a solution interior to frontier



Finally, a note on the full equilibrium

- The above leaves out firm responses to ranking algorithms
  - They might change prices if they know they will be ranked more highly
  - But it turns out prices wouldn't change much



From theory to bias tests

Using the platform supply function

Supply function and bias detection: COO

- $$I_j = \kappa_1 \ln(p_j) + \kappa_2 \delta_j^0 + \psi D_j + \varepsilon_j$$
- Suppose the index is **cardinal** and **linear**
- Then:

$$r_j = \kappa_1' \ln(p_j) + \boxed{\kappa_2' \delta_j^0} + \psi' D_j + \varepsilon_j$$

Hard to observe

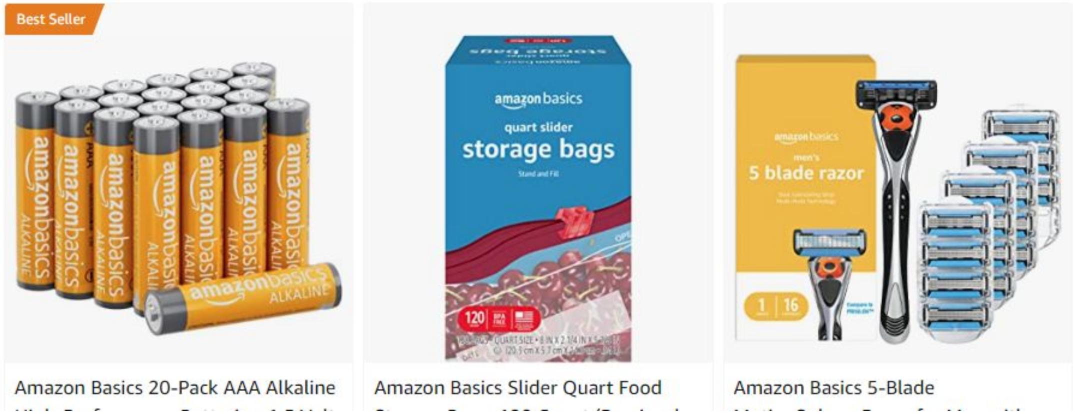
- Regress ranks on $\ln(p_j)$ and δ_j^0, then the coefficient on D_j reflects bias
- This is the "**conditioning on observables**" (COO) approach

COO implementation in practice

- Normally, we don't observe δ_j^0
- Instead, regress ranks on "controls" and D_j:

$$r_j = X_j\beta + \kappa \ln(p_j) + \psi D_j + \varepsilon_j$$

- ψ provides a measure of **bias** if X_j controls for *all* effects on demand
- But ψ could also reflect unobserved platform brand characteristics



Supply function and bias detection: OB approach

- Suppose a platform & a non-platform product have the same index value I_j
  - Platform product: $\kappa_1 \ln(p_j) + \kappa_2 \delta_j^0 + \psi D_j$
  - Non-platform product: $\underbrace{\kappa_1 \ln(p_k) + \kappa_2 \delta_k^0}_{\text{welfare frontier}}$

Monte Carlo simulations: OB test more reliable than COO

- If $\psi > 0$, then the platform product at the same rank is "worse"

$$\ln(q_j) = \mu_r + \kappa \ln(p_j) + \psi D_j + \varepsilon_j$$

Proxy for δ_j^0

rank FEs

Implementation and data needs

Rankings, platform identifier, and …

- **Conditioning on observables approach:**
  - characteristics legitimately predictive of ranks/sales
- **Outcome-based approach:**
  - outcomes caused by the ranks
- (Welfare analysis:
  - The above, plus causal rank effect estimates)

“Real-world” illustrations

(Illustrative) data and contexts

- Amazon Kindle Daily Deals 2022
  - 50 ranked titles each day
  - ≈20% published through Amazon: possibility of self-preferencing
- Expedia hotel searches 2013
  - 399,342 searches and 8,624,781 listings (121,545 randomized searches)
  - No self-preferencing. Possible bias with respect to chain hotels?
- Spotify New Music Friday 2017
  - 20 (of 50) ranked songs x 26 countries x ≈35 weeks
  - 18,489 listings; 6,637 appearing in top 200
  - Possible bias with respect to major labels?

Compare COO and OB: Amazon (an illustration)

- Both indicate self-preferencing
- Rank magnitudes differ

Panel A: Amazon Kindle Daily deals

	COO	Outcome-based
	(1)	(2)
	ln rank	ln quantity
preferred	-0.269***	-0.696***
	(0.025)	(0.042)
ln rank		
ln price	0.049***	-0.306***
	(0.012)	(0.020)
Observations	6796	6826

Compare COO and OB: Expedia

Panel B: Expedia hotels

	COO	Outcome-based
	(1) rank	(2) pr(buy)
chain	$\boxed{0.68104^{***}}$	$\boxed{0.00811^{***}}$
	$\boxed{(0.00866)}$	$\boxed{(0.00021)}$
rank		
promoted	0.14393***	
	(0.03203)	
price	0.01484***	-0.00017***
	(0.00006)	(0.00000)
Observations	6048717	6048717

Both find chain hotels are ranked “too low”

Compare COO and OB: Spotify

Panel C: Spotify New Music Monday

	COO	Outcome-based
	(1)	(2)
	rank	log streams
preferred	-1.212***	0.385***
	(0.087)	(0.036)
rank		
Observations	18233	18489

Opposite findings:

- COO: majors are ranked *too high*
- OB: majors are ranked *too low*

Bottom line: field data confirm Monte Carlo results (and general concerns about COO approach)

Structural approach

Platform preferences and welfare implications of self-preferencing

Structural model: Amazon

- **Demand** estimated as plain logit

$$\ln(s_j) - \ln(s_0) = x_j\beta + \alpha p_j + \gamma r_j + \xi_j$$

  - X contains an Amazon dummy and pre-promotion sales
  - → The estimated values, *minus (causal)* γr_j, give us $\boxed{\hat{\delta}_j^0}$

- **Supply**: rank-ordered logit

$$r_j = \kappa_1 \ln(p_j) + \kappa_2 \boxed{\hat{\delta}_j^0} + \psi D_j + \epsilon_j$$

Amazon estimates

- Demand
  - Causal part of rank effect = 0.335 from title FE approach
  - Allows calculation of δ_j^0

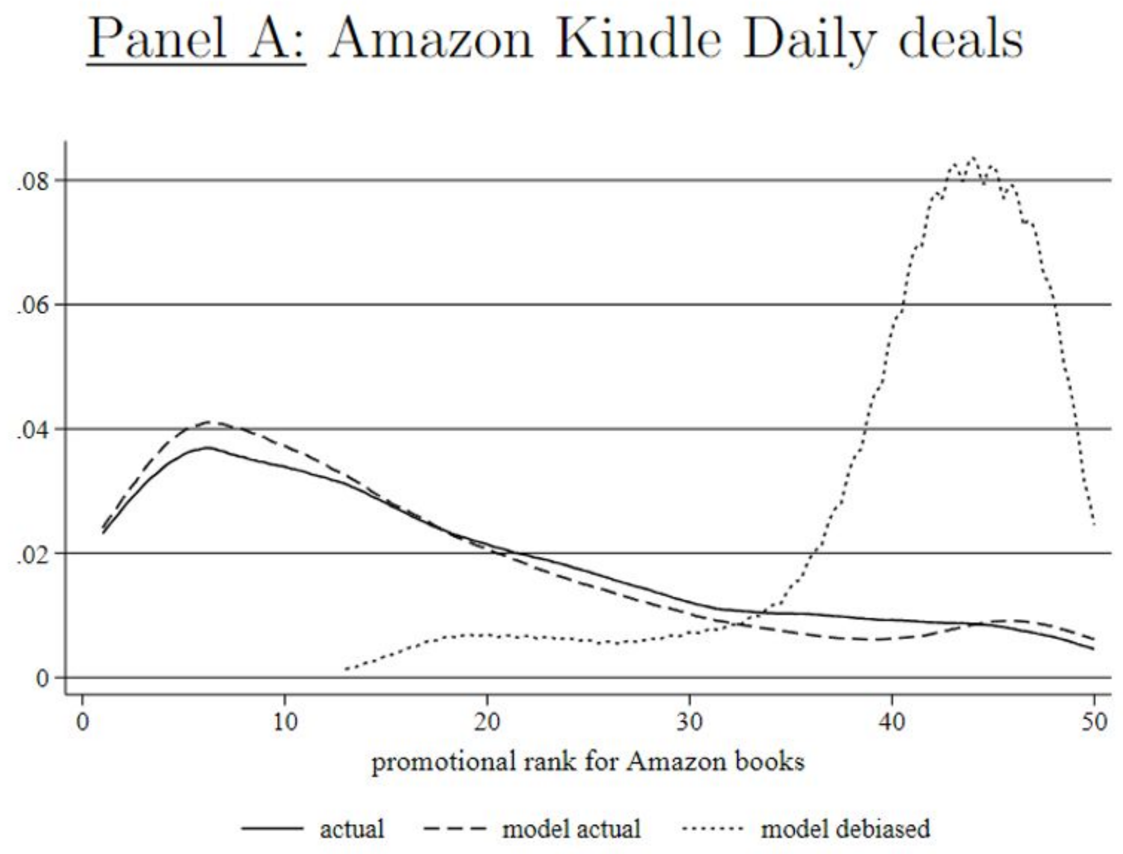
- Supply
  - Linear (intuitive): higher δ_j^0 gets better rank, as does higher price
  - Platform product gets better rank
  - Same pattern in rank ordered logit

Table 2: Amazon demand and supply estimates

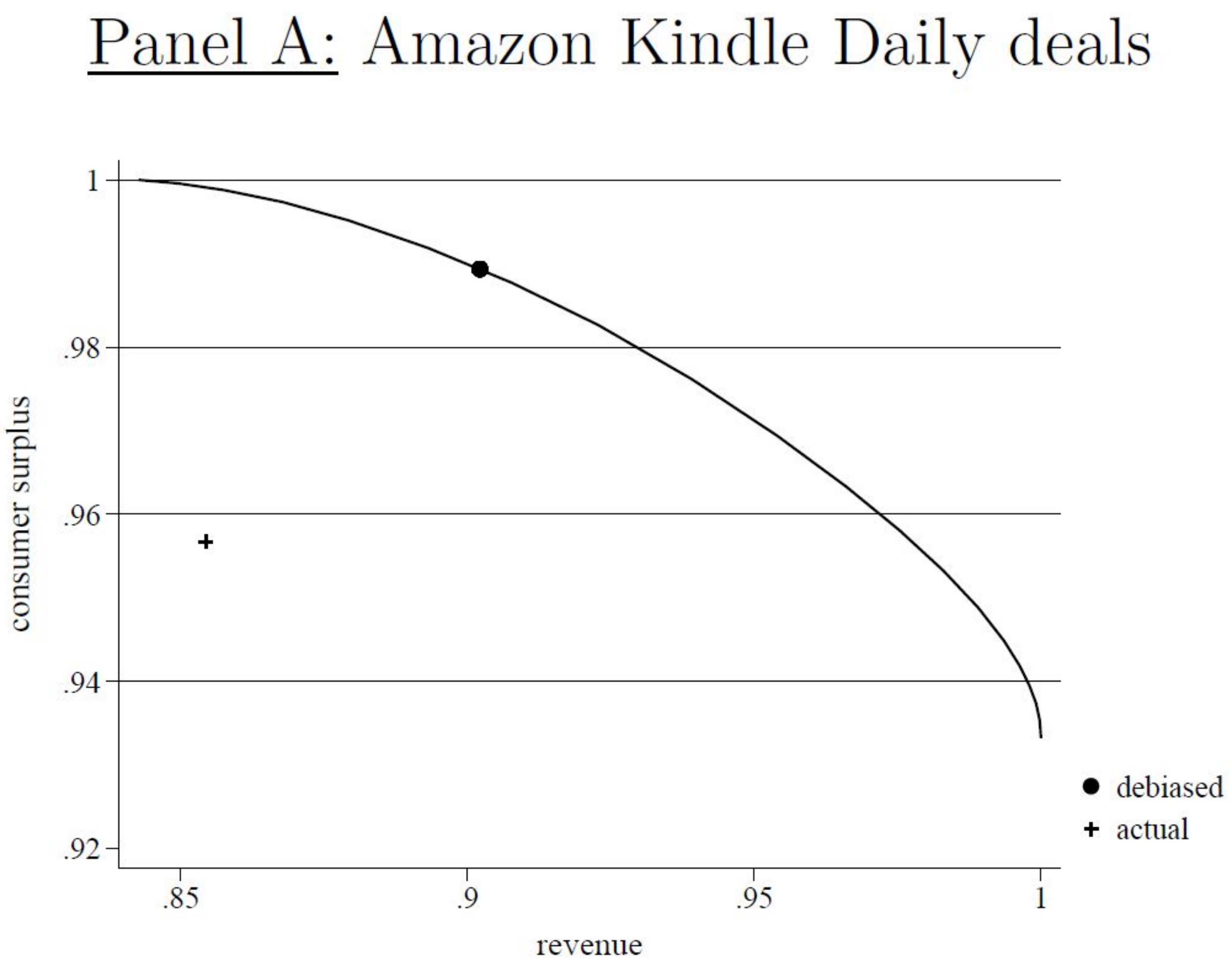
	demand
	(1)
	logit
price	-0.0391***
	(0.00435)
ln rank	-0.405***
	(0.0272)
platform product	-0.669***
	(0.0452)
ln daily pre-promo sales	0.0406*
	(0.0226)
ln price	
rank-indep mean util	
Observations	6826

Model: actual vs debiased ranks

Can re-calculate rankings after setting the "bias" parameter to zero



CS vs PS & bias



- Debiased point near CS max, further from rev max

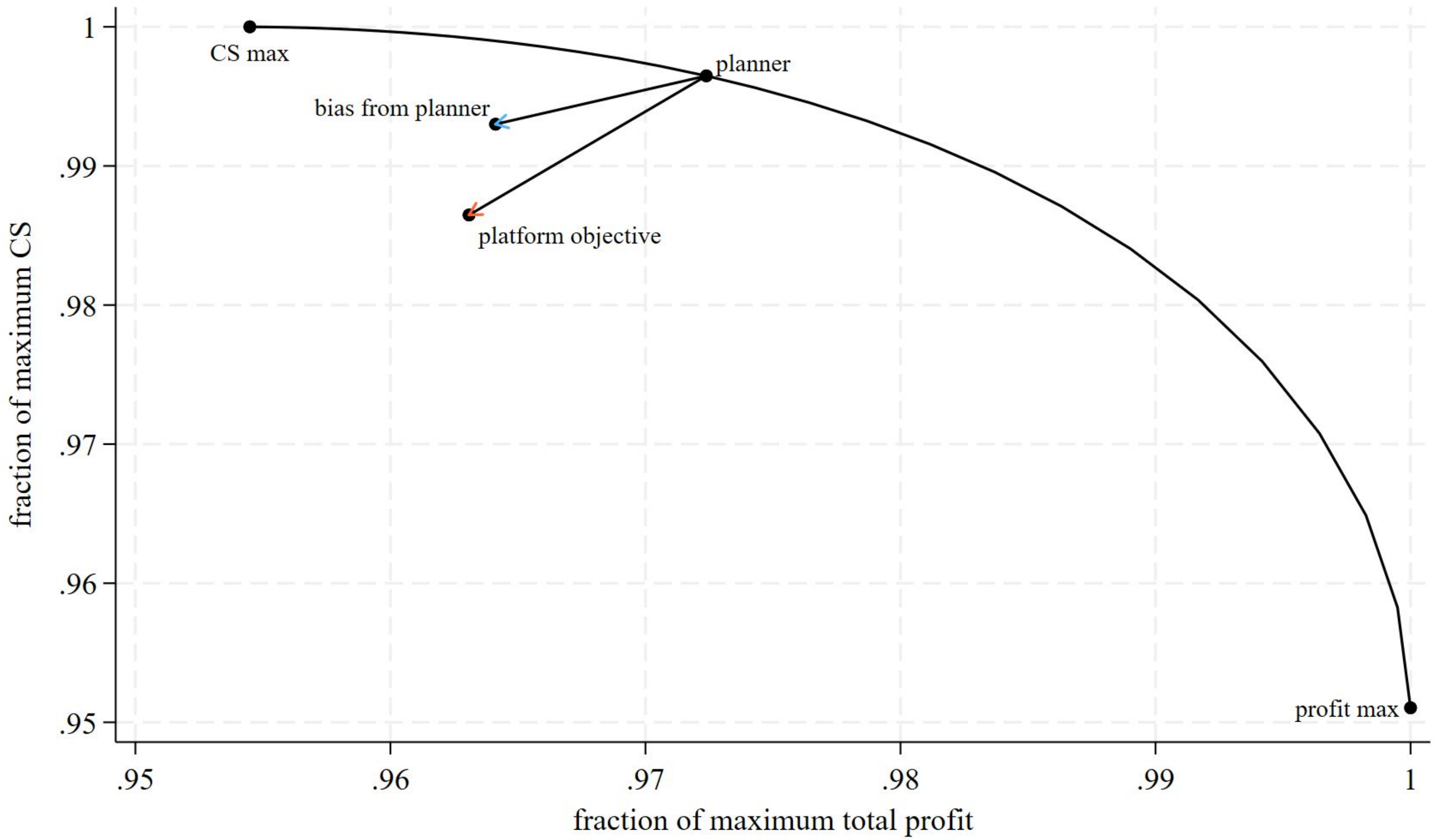
- Bias forgoes 3.3% of debiased CS, 5.3% of debiased PS

Note: where can “bias” come from?

- Source 1: the platform wants to give its own products preferential treatment
  - “Naked” self-preferencing

- Source 2: the platform cares about its commission
  - If compensated at proportional commission, the platform likes revenue
  - □ Is this illegal under the DMA?

Illustration: different platform objectives



Where else can “bias” come from?

- Source 3: the platform cares about things other than PS and CS
  - E.g., star ratings, return policies, …
  - These can be “accidentally” correlated with the platform dummy
  - □ Is this illegal under the DMA?

Conclusion

- Platform regulation: We need ways to test for, and evaluate, possible bias

- This paper presents a possible definition
    - As well as ways to test for, and measure welfare effects of, such bias

- Data access is hard for outsiders, but we hope this framework is useful for regulators