

Data Broker Competition and Downstream Entry

Laura Abrardi, Carlo Cambini, Flavio Pino

Discussion by Frank Verboven, KU Leuven

Cresse, Rhodes

June 2023

Overview: model

- Two upstream data brokers
 - Sell consumer data to downstream firms
 - Their data are not perfectly accurate
 - DB1 is more accurate than DB2: measure of competition
- Multiple downstream firms
 - Sell products to consumers
 - Can perfectly price discriminate for share of consumers where broker has information
 - Are differentiated on Salop circle, and enter freely given fixed costs

How does the level of competition in the DB market affect downstream entry and competition?

Overview: findings

- Base findings
 - Firms buy data from DB1, but DB2 exerts a competitive pressure
 - Firm entry is lower than in standard Salop model
 - Consumer surplus higher than in standard Salop circle if β sufficiently high
- Findings if DB2 data set does not completely overlap with DB1
 - If both datasets have synergies (superadditive), both DBs have more market power
 - DB prices increase, firm entry falls and consumers are harmed

General comments

- Interesting paper
- Contributes to small new literature on data brokers, how they enable price discrimination and affect welfare
- Founded in older literature since Thisse and Vives (1988) on price discrimination
 - Surplus extraction effect: tailored prices raise profits
 - Competition effect: reduced base prices reduces profits
- Overall, the paper presents a fairly positive policy view on data brokers
 - Focus on granting level playing field between brokers,
 - instead of limiting information accuracy
- My main comments
 - Broader context
 - Modeling assumptions

Data brokers: who, how and why?

- Very heterogeneous and complex market (Reviglo, Internet Policy Review, 2022)
 - Obtain data from ad-tech groups, consumer analytics firms, credit agencies, ...
 - Sell data to advertisers, employers, bankers, insurers, police departments, ...
 - Examples:
 - Acxiom, Oracle, Datalogix, Experian, ...
 - But also Facebook, Google, ...
- Sources of information:
 - individuals (app use), cookies, credit cards, third parties, governments, scraping,
- Uses (Ferraro, 2022):
 - Marketing/advertising
 - Fraud detection
 - Risk mitigation
 - People search sites

It would be useful to provide more of this context, to interpret the paper's focus

Modeling assumptions (1)

1. Consumers observe a single price

- In practice, arbitrage may happen
- If consumer sees high price, she may suspect it is targeted and look for other price (on other PC/phone)

2. Information accuracy is exogenous

- In practice, how do data brokers obtain their information?
- They may invest to avoid becoming DB2 with the less accurate data set

3. Data brokers sell their data set only to firms

- In practice, data brokers may also interact with consumers, including direct selling (Facebook, Google)
- What are incentives for “vertically integrated” platforms to sell their data to data brokers?

Modeling assumptions (2)

4. Data brokers sell a single “partition” of their data

- In practice, data brokers may sell menu of packages: base package, extended
- Data brokers may also not observe where the firm is located on the circle, partitioning more complicated

5. Synergic datasets are of specific kind (supermodular)

- Buying from **both** DB1 and DB2 raises accuracy
- This gives extra market power to both data brokers, almost by construction

Alternative approach:

- Data brokers cannot provide partitioned data that are centered around the firm
- Allow firms to buy from a single data broker, with partially overlapping partitions