Selling Cookies

Dirk Bergemann[1] Alessandro Bonatti[2]

[1]Yale University

[2]MIT Sloan

Federal Trade Commission Microeconomics Conference

October 16th, 2014

Markets for Personal Data

Markets for Personal Data

Recent developments in data collection and processing online.

Markets for Personal Data

Recent developments in data collection and processing online.

Welfare effects, regulatory implications depend on:

Markets for Personal Data

Recent developments in data collection and processing online.

Welfare effects, regulatory implications depend on:

1. How is the demand for data determined?

Markets for Personal Data

Recent developments in data collection and processing online.

Welfare effects, regulatory implications depend on:

1. How is the demand for data determined?

2. How should a data provider price its information?

Markets for Personal Data

Recent developments in data collection and processing online.

Welfare effects, regulatory implications depend on:

1. How is the demand for data determined?
2. How should a data provider price its information?
3. How does (data) market structure affect the equilibrium price?

Markets for Personal Data

Recent developments in data collection and processing online.

Welfare effects, regulatory implications depend on:

1. How is the demand for data determined?
2. How should a data provider price its information?
3. How does (data) market structure affect the equilibrium price?
4. Implications for related markets (e.g. advertising).

Markets for Personal Data

Direct marketing: mailing lists, lead generation.

Data brokerage: demographics, household financial means.

Markets for Personal Data

Direct marketing: mailing lists, lead generation.

Data brokerage: demographics, household financial means.

Data management platforms (BlueKai, eXelate, Krux, Lotame):

Markets for Personal Data

Direct marketing: mailing lists, lead generation.

Data brokerage: demographics, household financial means.

Data management platforms (BlueKai, eXelate, Krux, Lotame):

- use cookies to track consumers' behavior online;

Markets for Personal Data

Direct marketing: mailing lists, lead generation.

Data brokerage: demographics, household financial means.

Data management platforms (BlueKai, eXelate, Krux, Lotame):

- use cookies to track consumers' behavior online;
- compile segments based on "intent" or "purchase history."

Markets for Personal Data

Direct marketing: mailing lists, lead generation.

Data brokerage: demographics, household financial means.

Data management platforms (BlueKai, eXelate, Krux, Lotame):

- use cookies to track consumers' behavior online;
- compile segments based on "intent" or "purchase history."
- Data point = (browser ID, user characteristics).

Markets for Personal Data

Direct marketing: mailing lists, lead generation.

Data brokerage: demographics, household financial means.

Data management platforms (BlueKai, eXelate, Krux, Lotame):

- use cookies to track consumers' behavior online;
- compile segments based on "intent" or "purchase history."
- Data point = (browser ID, user characteristics).

Endogenous information structure: *targeted* vs. *residual* users.

Step 1. Review

Step 1. Review   Step 2. Finish

Thank you for providing your online preferences using the BlueKai registry! This simple 3-step process will put you in control of what some marketers know about you. Based on partner feedback, below is a list of your online preferences for topics of interest. Please review and edit by clicking "Remove" next to a particular preference. Then, click on "Continue" to select benefits.

▶ Location & Neighborhood

▶ Things You May Want To Buy

- Boston - Logan Intl (BOS)
  In-Market > Travel > Air Travel > United States > Massachusetts > Boston
- Leonardo da Vinci-Fiumicino (FCO)
  In-Market > Travel > Air Travel > International > Europe > Italy > Rome (All
- US Domestic Flyers
  In-Market > Travel > Air Travel > US Domestic Flyers
- International Flyers
  In-Market > Travel > Air Travel > International Flyers
- Economy
  In-Market > Travel > Air Travel > Economy

1 2

Please note that preferences are noted based on collective activities from your computer. If your computer is shared, this may reflect interests from other members of your household. To see how this works, visit www.welovesports.us (a fictitious sports site that will appear in a new window), and refresh this page.

Continue

Market for Cookies

Cookie data used to target online advertising.

Market for Cookies

Cookie data used to target online advertising.

DMPs work with sellers of online ad space ("media partners").

Market for Cookies

Cookie data used to target online advertising.

DMPs work with sellers of online ad space ("media partners").

Buying data about a segment $\sim$ adding a targeting category.

E.g., "Overome Tours" may buy the IDs of:

- 25,000 consumers who *have bought* a ticket to Rome;
- 50,000 consumers who *intend to purchase* a ticket.

It *tailors* bids, budgets, messages, volumes to each segment.

Market for Cookies

Cookie data used to target online advertising.

DMPs work with sellers of online ad space ("media partners").

Buying data about a segment $\sim$ adding a targeting category.

E.g., "Overome Tours" may buy the IDs of:

- 25,000 consumers who *have bought* a ticket to Rome;
- 50,000 consumers who *intend to purchase* a ticket.

It *tailors* bids, budgets, messages, volumes to each segment.

DMP charges a linear price *per unique user* or *per use*.

Consumers, Advertisers and Matching

Unit mass of *user characteristics* i and firms (advertisers) j.

Match value (potential surplus) $v\left(i,j\right) \in V$.

Firm j would like to *tailor* actions to values $v(i,j)$.

Consumers, Advertisers and Matching

Unit mass of *user characteristics* i and firms (advertisers) j.

Match value (potential surplus) $v\left(i,j\right) \in V$.

Firm j would like to *tailor* actions to values $v(i,j)$.

Firm j knows mapping $i \rightarrow v(i,j)$ and distribution $F_j\left(v\right)$.

Consumers, Advertisers and Matching

Unit mass of *user characteristics* i and firms (advertisers) j.

Match value (potential surplus) $v(i, j) \in V$.

Firm j would like to *tailor* actions to values $v(i, j)$.

Firm j knows mapping $i \to v(i, j)$ and distribution $F_j(v)$.

Characteristics i unobservable: targeting requires user-level data.

Information and Data Provider

Monopolist data provider knows individual users' characteristics i.

Information and Data Provider

Monopolist data provider knows individual users' characteristics i.

Suppose firm j wants to identify all users with match value v.

Information and Data Provider

Monopolist data provider knows individual users' characteristics i.

Suppose firm j wants to identify all users with match value v.

Firm j requests IDs of all users with characteristics $i : v(i, j) = v$.

Information and Data Provider

Monopolist data provider knows individual users' characteristics i.

Suppose firm j wants to identify all users with match value v.

Firm j requests IDs of all users with characteristics $i : v(i, j) = v$.

"Firm j buys cookie v"

Information and Data Provider

Monopolist data provider knows individual users' characteristics i.

Suppose firm j wants to identify all users with match value v.

Firm j requests IDs of all users with characteristics $i : v(i, j) = v$.

"Firm j buys cookie v"

Let $\mu(A_j)$ denote the measure of users i with $v(i, j) \in A_j$.

Information and Data Provider

Monopolist data provider knows individual users' characteristics i.

Suppose firm j wants to identify all users with match value v.

Firm j requests IDs of all users with characteristics $i : v(i, j) = v$.

"Firm j buys cookie v"

Let $\mu(A_j)$ denote the measure of users i with $v(i, j) \in A_j$.

Firm j can choose any measurable subset of users $A_j \subset V$.

Information and Data Provider

Monopolist data provider knows individual users' characteristics i.

Suppose firm j wants to identify all users with match value v.

Firm j requests IDs of all users with characteristics $i : v(i,j) = v$.

"Firm j buys cookie v"

Let $\mu(A_j)$ denote the measure of users i with $v(i,j) \in A_j$.

Firm j can choose any measurable subset of users $A_j \subset V$.

Data about individual users sold at a constant linear price,

$$p(A_j) \triangleq p \cdot \mu(A_j).$$

Selling Access to a Database

Selling Cookies (today's paper):

- Unit of sale: individual queries (realizations of a r.v.).
- Linear price per query.
- More elaborate market environment.

Selling Experiments (tomorrow's paper):

- Unit of sale: arbitrary information structures.
- Menu pricing of information.
- Abstracts from source of value.

Advertising as Matching

Advertisers generate value by choosing **match intensity** q:

$$\pi(v, q) = v \cdot q - c \cdot m(q).$$

Advertising as Matching

Advertisers generate value by choosing **match intensity** q:

$$\pi(v, q) = v \cdot q - c \cdot m(q).$$

Matching cost function $m(q)$ strictly increasing and convex.

Advertising as Matching

Advertisers generate value by choosing **match intensity** q:

$$\pi(v, q) = v \cdot q - c \cdot m(q).$$

Matching cost function $m(q)$ strictly increasing and convex.

$q =$ consumer awareness probability;

$m(q) =$ required amount of advertising space;

$c =$ unit price of advertising space.

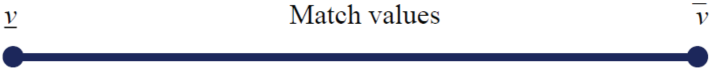
Choice of Information Structure

Suppose firm j buys cookies in the *targeted set* A.

- Complete-information action $q^*(v)$ for each $v \in A$.
- Constant action $q^*(A^C)$ for all $v \notin A$.

Each firm chooses a *targeted set* A to maximize

$$\int_A (\pi(v, q^*(v)) - p)\, \mathrm{d}F(v) + \int_{A^C} \pi(v, q^*(A^C))\mathrm{d}F(v).$$



Choice of Information Structure

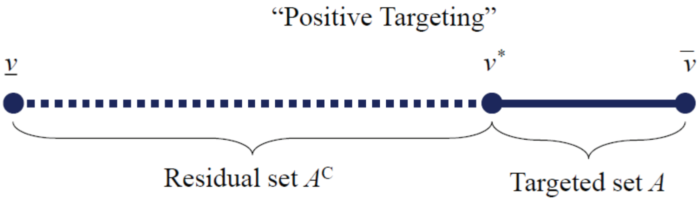
Suppose firm j buys cookies in the *targeted set* A.

Complete-information action $q^*(v)$ for each $v \in A$.

Constant action $q^*(A^C)$ for all $v \notin A$.

Each firm chooses a *targeted set* A to maximize

$$\int_A (\pi(v, q^*(v)) - p) \, dF(v) + \int_{A^C} \pi(v, q^*(A^C)) dF(v) .$$



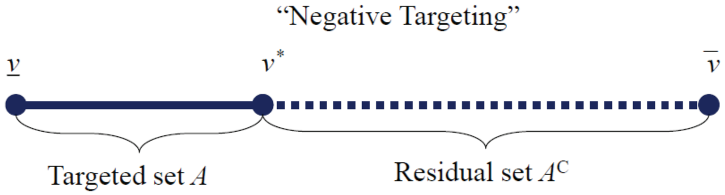
Choice of Information Structure

Suppose firm j buys cookies in the *targeted set* A.

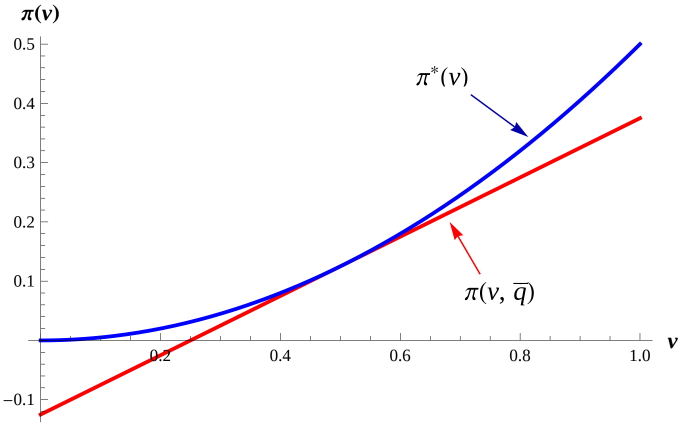
- Complete-information action $q^*(v)$ for each $v \in A$.
- Constant action $q^*(A^C)$ for all $v \notin A$.

Each firm chooses a *targeted set* A to maximize

$$\int_A (\pi(v, q^*(v)) - p)\, \mathrm{d}F(v) + \int_{A^C} \pi(v, q^*(A^C))\mathrm{d}F(v).$$



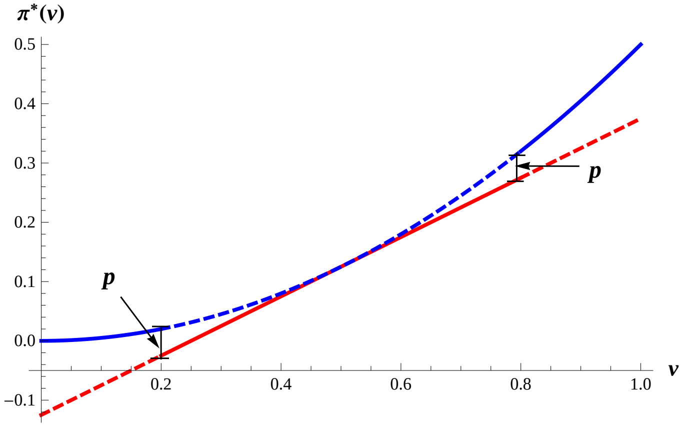
Prior- and Full-Information Benchmarks



Linear profits under prior information.

Strictly convex profits under full information.

Positive and Negative Targeting



p = marginal value of information.

Optimal Residual Set

Proposition (Convexity of Residual Set)

For any $c, p > 0$*, the optimal residual set* $A^{\mathsf{C}}(c, p)$ *is a non-empty interval* $[v_1(c, p), v_2(c, p)]$*.*

Optimal Residual Set

Proposition (Convexity of Residual Set)

For any $c, p > 0$, the optimal residual set $A^{C}(c, p)$ is a non-empty interval $[v_1(c, p), v_2(c, p)]$.

Optimal *targeted set* trades off:

- gains from adaptation to values v;
- likelihood of each realization v.

Optimal Residual Set

Proposition (Convexity of Residual Set)

For any $c, p > 0$*, the optimal residual set* $A^{C}(c, p)$ *is a non-empty interval* $[v_1(c, p), v_2(c, p)]$*.*

Optimal *targeted set* trades off:

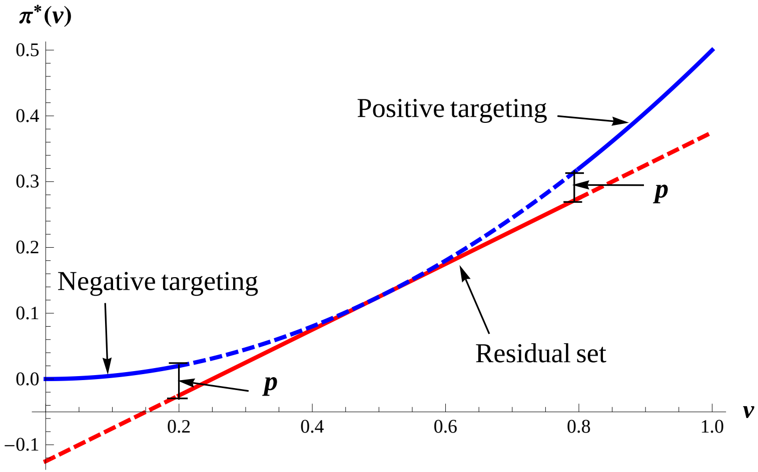
- gains from adaptation to values v;
- likelihood of each realization v.

Proposition (Positive and Negative Targeting)

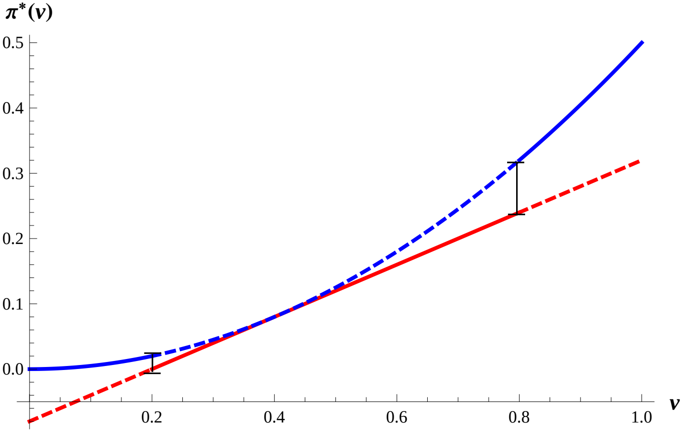
With symmetrically distributed match values and quadratic matching costs, the optimal residual set is given by:

$$A^{C}(c, p) = \left[\mathbb{E}[v] - \sqrt{cp},\ \mathbb{E}[v] + \sqrt{cp}\right].$$

Joint Positive and Negative Targeting

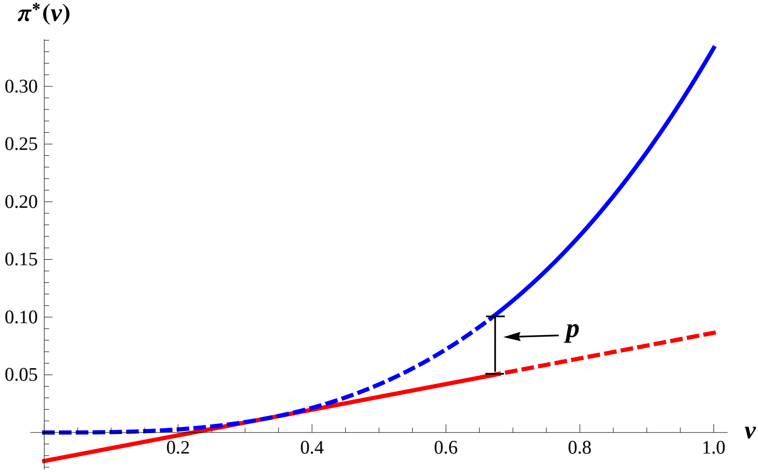


Decreasing Density $f(v)$



Value of information highest "at the top" of residual set.

Positive Targeting



Optimal targeted set is an interval, $A = [v^*(c,p), \bar{v}]$.

Demand for Data and Advertising: Summary

Positive vs. negative targeting vs. both depends on:

- advertising (matching) technology;
- distribution of consumers characteristics.

I.e. properties of complete-information profits alone.

Demand for Data and Advertising: Summary

Positive vs. negative targeting vs. both depends on:

- advertising (matching) technology;
- distribution of consumers characteristics.

I.e. properties of complete-information profits alone.

Demand for advertising:

- Differential spending levels within targeted set.
- Uniform (positive) spending level on residual set.

The Data Seller's Tradeoff

Assume *positive targeting* is optimal, $A(c, p) = [v^*, \bar{v}]$.

A monopolist seller chooses the threshold v^* to maximize

$$\underbrace{\left(\pi\left(v, q^*\left(v\right)\right) - \pi\left(v, q^*\left(\mathbb{E}\left[\tilde{v} \mid \tilde{v} \leq v\right]\right)\right)\right)}_{\text{price}} \underbrace{\left(1 - F\left(v\right)\right)}_{\text{quantity}}.$$

The Data Seller's Tradeoff

Assume *positive targeting* is optimal, $A(c,p) = [v^*, \bar{v}]$.

A monopolist seller chooses the threshold v^* to maximize

$$\underbrace{\left(\pi\left(v, q^*\left(v\right)\right) - \pi\left(v, q^*\left(\mathbb{E}\left[\tilde{v} \mid \tilde{v} \le v\right]\right)\right)\right)}_{\text{price}} \underbrace{\left(1 - F\left(v\right)\right)}_{\text{quantity}}.$$

Effects of expanding supply (lowering v^*):

1. Lower marginal value of information (*i.e.* price) at $v = v^*$.
2. Lower match intensity with residual set $A^{C} = [\underline{v}, v^*]$.

Note: $\#2$ partially compensates for $\#1$.

Competing Data Sellers

Each seller has an exclusive over one consumer...
and sets the price of the corresponding cookie.

Competing Data Sellers

Each seller has an exclusive over one consumer. . .
and sets the price of the corresponding cookie.

Analogous to selling your own data, or data exchange.

Competing Data Sellers

Each seller has an exclusive over one consumer...
and sets the price of the corresponding cookie.

Analogous to selling your own data, or data exchange.

Each seller takes $q^*(A^{\mathsf{C}})$ as given.

Competing Data Sellers

Each seller has an exclusive over one consumer. . .
and sets the price of the corresponding cookie.

Analogous to selling your own data, or data exchange.

Each seller takes $q^*(A^C)$ as given.

Proposition (Data Sales Fragmentation)

The symmetric equilibrium price with a continuum of data sellers $\bar{p}$ exceeds the monopoly price p^*.

Result extends to n independent and exclusive data sellers.

Incomplete Database

Monopolist data provider has information about $\beta \leq 1$ users.

$v \sim F(v)$ both in the database and outside.

Incomplete Database

Monopolist data provider has information about $\beta \leq 1$ users.

$v \sim F(v)$ both in the database and outside.

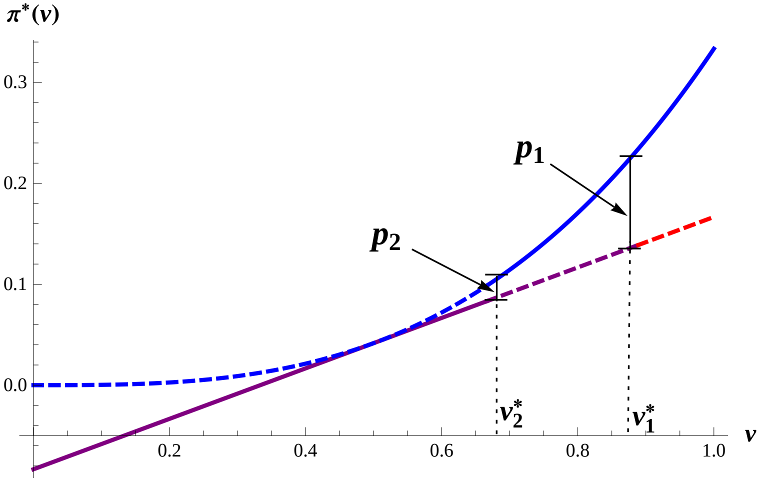
Proposition (Reach of the Database)

1. *Then the advertisers marginal willingness to pay $p(A;\beta)$ is increasing in β for all A.*
2. *(Under suitable conditions) the monopoly price is strictly decreasing in the reach β.*

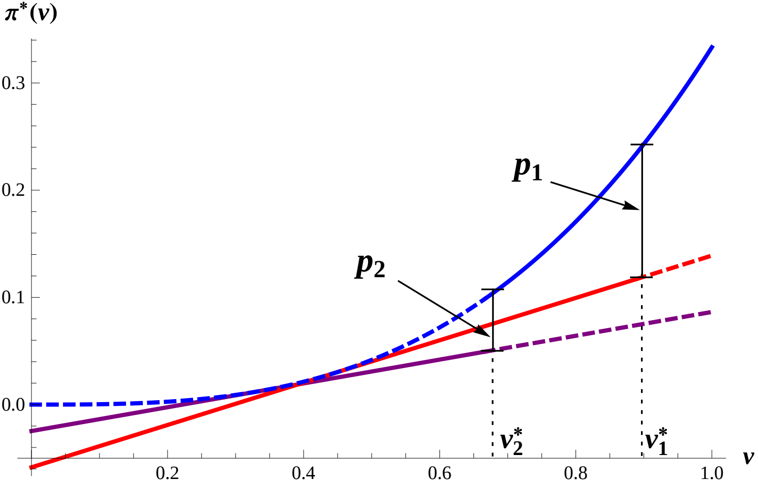
Low *reach* $\beta \Rightarrow q^*(A^{\mathsf{C}})$ not *responsive* to A.

Lower incentives to expand supply of data.

Composition Effect



Composition Effect



Monopoly Pricing: Summary

Advertisers are interested in multiple user profiles.

Willingness to pay for *targeted* users depends on the *residual set*.

Monopoly Pricing: Summary

Advertisers are interested in multiple user profiles.

Willingness to pay for *targeted* users depends on the *residual set*.

Data provider can influence the *composition* of the residual set.

Composition effect provides incentive to *lower* price.

Monopoly Pricing: Summary

Advertisers are interested in multiple user profiles.

Willingness to pay for *targeted* users depends on the *residual set*.

Data provider can influence the *composition* of the residual set.

Composition effect provides incentive to *lower* price.

Lower prices imply efficiency gains (matching).

Monopoly Pricing: Summary

Advertisers are interested in multiple user profiles.

Willingness to pay for *targeted* users depends on the *residual set*.

Data provider can influence the *composition* of the residual set.

Composition effect provides incentive to *lower* price.

Lower prices imply efficiency gains (matching).

Market structure or data availability may limit composition effect.

Concluding Remarks

A model of markets for personal data:

- specific information structures;
- implications of linear “cookie pricing.”

Concluding Remarks

A model of markets for personal data:

- specific information structures;
- implications of linear "cookie pricing."

Extensions:

- Cross-market externalities: availability of data may increase equilibrium price of advertising space.
- Consumer surplus: advertising as matching $(+)$ vs. division of total surplus (?).
- Value of privacy and endogenous data availability.