ADVERSE SELECTION AND AUCTION DESIGN FOR INTERNET DISPLAY ADVERTISING

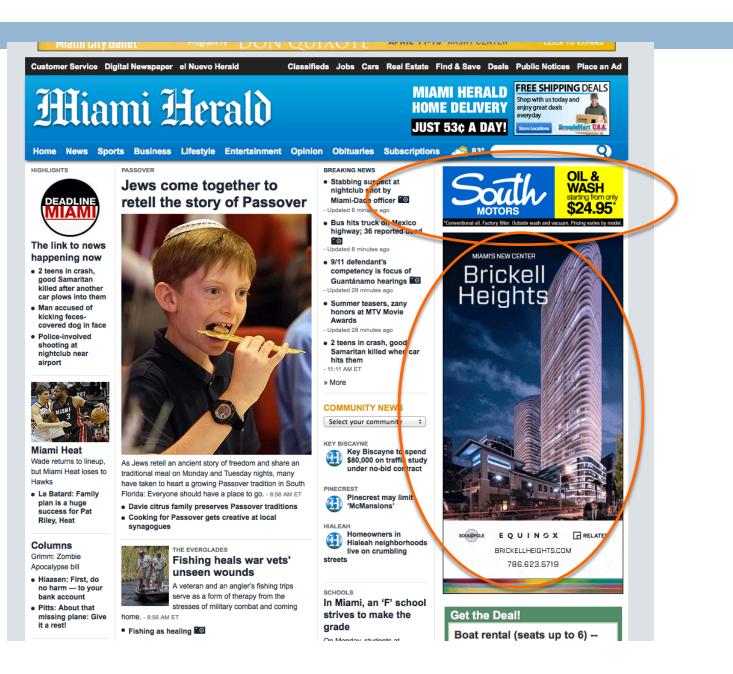
NICK ARNOSTI, MARISSA BECK AND PAUL MILGROM NOVEMBER 2015

Old Advertisers & New

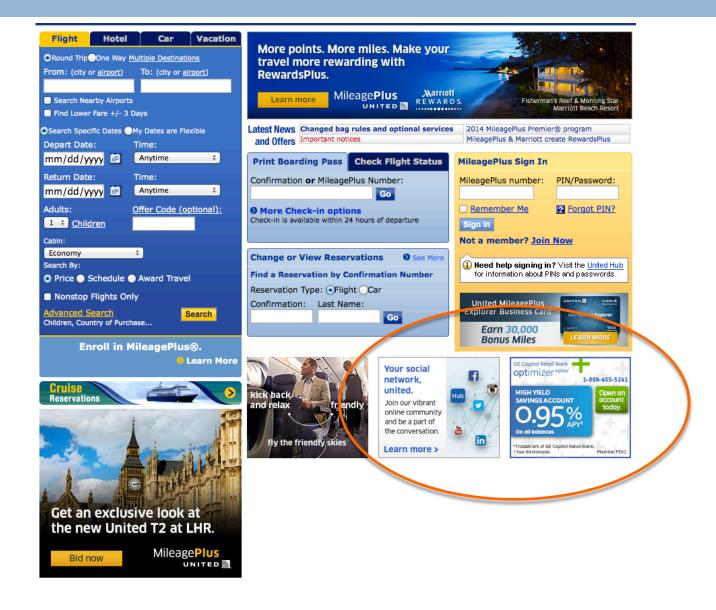
"Half the money I spend on advertising is wasted; the trouble is, I don't know which half."

- John Wanamaker, Advertising pioneer

Old-Fashioned "Brand" Ads



New-Fashioned "Performance" Ads



Display Advertisement Types

Brand Ads

- Goal: awareness and image
 - Reach and repetition.
- Common Characteristics
 - Targeted to a large group
 - Large number of Impressions
 - Guaranteed delivery
- Sample Advertisers
 - Ford (weekend auto sale)
 - Disney (movie openings)
 - Shopping Center (location)

Performance Ads

- Goal: measurable action now
 - Click, fill form, or buy.
- Common Characteristics
 - Targeted to an individual (based on cookies)
 - Smaller number of impressions
 - Bought one by one
- Sample Advertisers
 - Hertz (car rental)
 - Amazon (re-targeting)
 - Quicken mortgage (refinance)

Danger of Adverse Selection

Brand Advertisers

- May select impressions en masse ("road block" ads)
- Receive deferred, aggregated data about performance of the whole ad campaign
- Cannot easily distinguish lowperforming ads and publishers

Performance Advertisers

- Mostly use private cookies to select impressions
- Receive immediate, detailed data about the performance of individual ads
- Can quickly identify lowperforming ads and publishers

If the value of ad impressions is positively correlated for both types of advertisers, then brand advertisers may suffer adverse selection.

6 Matching with Adverse Selection

...and "Not-Quite-Optimal" Market Design

Model

- $\Box \quad \text{There are } N + 1 \text{ advertisers, with } N \geq 2$
- \Box The value of an impression to advertiser *i* is $X_i = CM_i$
- \Box C is the (random) **common value factor** and
 - *M_i* is the (random) *match value factor* for bidder *i*

Key Assumptions

- 1. Advertiser 0 (the "brand advertiser") does not observe X_0
- 2. Performance advertisers n = 1, ..., N observe their values X_n Define $X = (X_1, ..., X_n)$.
- 3. The common value factor C is statistically independent of $M \stackrel{\text{def}}{=} (M_0, \dots, M_N)$

A Market Design Challenge

- Compare the restricted-worst-case performance of different mechanisms on efficiency grounds
- □ The mechanisms considered are:
- 1. A benchmark: "Omniscient" mechanism with C observed
- 2. "Optimal" (expected-efficiency maximizing) mechanisms
- 3. Second-price auction
- 4. "Modified second-bid auction" in which the highest performance bidder wins if the ratio of the highest to second-highest performance bid exceeds a threshold.

9 The Omniscient Benchmark

OMN, in which the auctioneer observes both the bids and C

OMN Benchmark

If the auctioneer could separately gather perfect information about the common factor C and decide the allocation accordingly (no incentive constraints), it could achieve this value:

$$V(OMN) = E[\max(C \cdot E[M_0], X_1, \dots, X_n)]$$

Performance of other mechanisms will be compared to V(OMN).

11 Bayesian Optimal Mechanism

OPT ... and its drawbacks

Optimal Mechanism Formulation

- $\Box z_i(X)$ is probability that *i* wins, given X
- $\square p_i(X)$ is *i*'s expected payment, given X
- Efficiency Objective
 - Goal is to maximize $E[\sum_{i=0}^{n} X_i z_i(X)]$
 - subject to dominant-strategy incentive constraints and participation constraints
 - Let OPT be the mechanism that does that

OPT in an Example

13

 \square Assume that M_1, \ldots, M_n are IID and that...

$$P\{C = 1\} = P\{C = 2\} = \frac{1}{2}$$

$$P\{M_n = 1\} = P\{M_n = 2\} = P\{M_n = 4\} = \frac{1}{3}$$

 $3 < E[M_0] < 4$

□ So, optimally, only a performance advertiser n with $M_n = 4$ ought to be assigned this impression.

Example Solved

- \Box The expected-efficiency-maximizing assignment with N = 2 is:
 - □ If $X_{(1)} \in \{1,2\}$, then $M_{(1)} \leq 2 < E[M_0] \Rightarrow$ brand advertiser
 - If $X_{(1)} = 8$, then $M_{(1)} = 4 > E[M_0] \Rightarrow$ top performance advertiser
 - If $X_{(1)} = 4$, assignment hinges on whether $E[M_{(1)}|X_{(1)}, X_{(2)}] \ge E[M_0]$.
 - If $X_{(2)} = 1$, then $M_{(1)} = 4 \Rightarrow$ top performance advertiser
 - If $X_{(2)} = 2$, then $E[M_{(1)}|X_{(1)}, X_{(2)}] = 3 < E[M_0] \Rightarrow$ brand advertiser
 - In this case, $\Pr\{C = 1, M_{(1)} = 4, M_{(2)} = 2 | X_{(1)}, X_{(2)}\} = \Pr\{C = 2, M_{(1)} = 2, M_{(2)} = 1 | X_{(1)}, X_{(2)}\} = \frac{1}{2}.$
 - If $X_{(2)} = 4$, then $E[M_{(1)}|X_{(1)}, X_{(2)}] = 3 < E[M_0]$, \Rightarrow brand advertiser
 - In this case, $\Pr\{C = 1, M_{(1)} = M_{(2)} = 4 | X_{(1)}, X_{(2)}\} = \Pr\{C = 2, M_{(1)} = M_{(2)} = 2 | X_{(1)}, X_{(2)}\} = \frac{1}{2}.$

Main Concerns about OPT

The example highlights three concerns about OPT

- 1. Sensitivity: OPT is sensitive to detailed distributional assumptions.
- 2. False-name bidding: Performance advertiser n with value $X_n = 4$ can only benefit by submitting a additional, false-name bid of $X_{\hat{n}} = 1$ (because that leads the auctioneer to infer that $M_n = 4$.)
- Adverse selection: The brand advertiser wins 4/9 of high-value impressions, but 7/9 of low-value ones.
 - Most problematic if the brand advertiser feels uninformed about the impressions and who else may be bidding.

16 MSB Characterization

Modified Second Bid auction characterized by its properties

Properties for Characterization

A mechanism is

- anonymous among performance advertisers if...
- strategy-proof if...
- fully strategy-proof if, in addition, it is both
 - bidder false-name proof: no bidder can benefit by submitting multiple bids, and
 - publisher false-name proof: the seller cannot benefit by submitting "low" bids (below all performance bids)
- adverse-selection free if for every joint distribution on (C, M) such that C and M are independent, $z_0(X)$ is statistically independent of C.

Characterization Theorem

- Definition. A direct mechanism is a modified second bid auction if for some $\alpha \ge 1$,
 - If $\frac{X_{(1)}}{X_{(2)}} > \alpha$, then the highest performance advertiser wins & pays $\alpha X_{(2)}$.
 - If $\frac{X_{(1)}}{X_{(2)}} \leq \alpha$, then the brand advertiser wins (and pays its contract price).
- Theorem. A deterministic mechanism (z, p) is anonymous, fully strategy-proof, and adverse selection free *if and only if* it MSB.

19 Comparing MSB_{α} and SP_{r} to OMN

 MSB_{α} : modified second-bid auction

 SP_r : second-price auction with reserve

Assumptions for Comparison

- □ Evaluate MSB_{α} and SP_{r} mechanisms in worst case over a limited family of environments, in which...
 - \blacksquare M_1, \ldots, M_N are IID from a distribution F.
 - $\square C$ is drawn from distribution G.
 - $N \ge 2$ and $E[M_0] \ge 0$ are arbitrary.

Efficiency Performance

- 21
- **Theorem.** (Comparing SP_r and MSB_{α} to OMN)
- Assuming Nash equilibrium bidding by the brand advertiser, both MSB and SP have similar worst case performance:

$$\inf_{\substack{F,G,N\geq 2, E[M_0]\geq 0}} \max_{\alpha} \frac{V(MSB_{\alpha})}{V(OMN)} = \frac{1}{2}$$
$$\inf_{\substack{F,G,N\geq 2, E[M_0]\geq 0}} \max_{r} \frac{V(SP_{r})}{V(OMN)} = \frac{1}{2}$$

2. Further restricting F and/or G to be drawn from power law distributions \mathcal{P} ,

$$\inf_{\substack{F \in \mathcal{P}, G \in \mathcal{P}, N \ge 2, E[M_0] \ge 0 \\ F \in \mathcal{P}, G, N \ge 2, E[M_0] \ge 0 }} \max_{\alpha} \frac{V(SP_r)}{V(OMN)} = \frac{1}{2}}{V(OMN)} \approx 0.948$$

Revenue Performance

- 22
- **Theorem.** Fix a number of bidders N and assume that the publisher shares in the rents from brand advertising in any fixed proportions, say $(\delta, 1 \delta)$.
- □ If *F* is a power law distribution, then there is some α such that MSB_{α} achieves at least 94.8% of the expected revenue from the corresponding expected-revenue-maximizing strategy-proof auction *REVMAX*.

Conclusion

- Adverse selection can be neutralized, even without encouraging false-name bidding, provided that X_n = CM_n and C and M are independent.
- The cost of doing that, even without observing the common value factor C, is low provided that the tails of the distribution are fat (power law).
- □ For real applications, we need to evaluate...
 - Is adverse selection important?
 - Are values independent?
 - Are match-value distributions fat-tailed?

