Structural Models and Market Design in the Age of Experimentation

Susan Athey, Harvard University

Today's Talk

- Market design problem
  - Design auction-based marketplace
  - Extent of participation and engagement crucial
- Methods
  - Theory, structural models and experiments
  - Complements, not substitutes
- Case studies
  - Timber
  - Internet Search Advertising

Research v. Practice

Online v. Offline

Methods: Theory

- Conceptual framework based on theory literature very useful for empirical research and practice

- Fixed participation
  - Single-unit auctions: Myerson, Maskin/Riley
  - Online ads: Varian, Edelman-Ostrovsky-Schwarz

- Endogenous participation
  - Single-unit auctions: many…
  - Online ads: Athey-Ellison (QJE 2011), this paper

Methods:
Experiments and Structural Models

Structural models

- Structural models essential for evaluating market design and algorithms
  - Impact of policies on welfare/profit needed to predict long-run responses
  - Evaluate efficiency v. revenue tradeoffs
  - Consider market designs not currently in use

Field experiments

- Compare design alternatives
- Estimation of models
  - Create data with exogenous variation
  - Generate data from alternative/new designs
- Model validation

	Structural Models	Field Experiments
Offline	Used in research, occasional policy evaluations for market design, e.g.: • Optimal reserve price modeling for softwood lumber trade dispute (Athey & Ingraham) • Comparing auction format in treasury auctions (e.g. Hortacsu, Kastl) • Detection and damages in bidder collusion (Marshall et al, Bajari, Porter, Pesendorfer, etc.) • Merger analysis in auction markets (Froeb)	Expensive, slow Relatively rare but can be very influential
Online	Algorithms used *are* structural models Long run impact of market design, algorithms: where experiments are expensive, slow, uninformative • E.g. advertiser reactions (Ostrovsky & Schwarz, Athey & Nekipelov, this paper)	Essential and integral to businesses • Algorithm training and real-time explore/exploit • Selection and evaluation of new algorithms: REQUIRED • Tuning algorithms and UI • Validation of structural models Researchers can create (e.g. eBay) or exploit (Einav & Levin, this paper)

Theory, Structural Models and Experiments Yield Market Design Insights

- "Size of the pie" v. distribution of rents
  - Short-run distribution affects long run size of pie and platform revenue
  - Direct: platform share of pie
  - Indirect: participation
- Indirect channel often dominates
  - Policy conclusion: design markets for participation
  - Preference policies for weak bidders may not hurt revenue
- Internet search advertising
  - Substantial tradeoffs between efficiency and short-run revenue
    - Incentives to price discriminate and favor market thickness over efficiency
    - Incentives to show too many irrelevant ads
  - Incorporating long-run bidder and user responses leads to different design choices

Case Study: Timber Auctions

Susan Athey, Jonathan Levin, and Enrique Seira, *QJE* 2011

Susan Athey, Dominic Coey, and Jonathan Levin, working paper, 2010

Overview: Market Design in Timber

- Market design problem
  - Efficient & revenue-generating market design
  - Preserve rents and employment for small, local timber firms
- Design elements
  - Open v. sealed bidding
  - Preference policies
  - Reserve prices
  - Packaging tracts (big v. small)
- Issues
  - Few large mills v. many small loggers
  - Collusion historical problem
  - Policy mandate to preserve volumes to small business
- Broader applications
  - US Procurement $500b/yr
  - US Goal: 23% small business
  - Participant structure fairly common
  - Collusion common in construction, procurement

Theoretical Framework: Single Unit Auctions

	Exogenous Entry	Endogenous Entry

Ascending v. 1st Price	• Revenue Equivalence Theorem:
	  • Same revenue, bidder profits
	• Key assumptions:
	  • Independent private values
	  • Risk neutral
	  • Symmetric bidders
	  • Competitive behavior
		• RET extends to entry since profits are the same

Asymmetric Bidders	• Efficiency v. revenue tradeoff
	• Optimal auction (Myerson):
	  • Bias in favor of weak bidder
	• Ascending v. 1st Price (Maskin/Riley)
	  • Sealed often does better, weak bidders shade less, win more
		• 1st price attracts more weak bidders, reinforcing revenue advantage

Collusion	• Ascending vulnerable to collusion • Possible to respond immediately to deviation	• If strong bidders collude perfectly on bidding, weak bidder entry unaffected
Preferences	• Set-asides reduce revenue • Subsidies to weak bidders can mimic optimal auction	• In set-asides, additional entry can increase revenue over unrestricted • Subsidies can encourage entry, so larger subsidies optimal

Evaluating Market Design: Field Experiments and Settings with Selection on Observables

- Open v. Sealed
  - Randomized field experiment (ID/MT)
    - Probabilities vary by region, based on observables
  - Volume-based policy (CA)
- Set-asides
  - Large sales excluded
  - Policy to select representative tracts

The Impact of Design Choices

Treatment Effect of Auction Method				
Dependent Variable:	ln(Logger Entry)	Loggers/ Entrants	Logger Wins	ln(Revenue)
Avg. 1st Price Effect	0.097	0.058	0.038	0.099
	(0.036)	(0.016)	(0.027)	(0.039)

Treatment Effect of Small-Business Set-Aside Provision		
Dependent Variable:	ln(Entrants)	ln(Revenue)
Avg. Set-Aside Effect	0.083	-0.001
	(0.050)	(0.045)

Treatment effect methods enable evaluation of impact on entry and revenue

Cannot assess whether relative magnitudes are consistent with theory

Cannot assess efficiency/revenue tradeoffs

Methodological Steps

1. Identification & estimation with one market design / behavioral assumption
2. Predict out of sample to alternative design
3. Validation and model selection
   - Compare predictions from alternative models to actual outcomes
4. Do counterfactual analysis with validated model

Structural Model: Set-Up and Identification

	Exogenous Entry	Endogenous Entry
Primitives	• Value distributions (IPV) for large and small bidders • Distribution of unobserved auction heterogeneity	• Entry costs
Behavioral Assumptions	• Bayesian Nash equilibrium • Alternative: Large bidder collusion	• Large bidders fixed number • Small bidders: free entry until zero profit
Observables	• Bids (1st Price) • Drop-out points (ascending)	• Auction-level entry decisions
Identification	• 1st Price: nonparametric identification • Ascending: *not* identified with unobs. heterogeneity	• Given identification of values, infer profits and thus entry costs for zero profit
Set-asides	• Cannot observe or identify anything about large bidders in set-aside auctions	

Model Validation and Testing

<table>
<tr><th></th><th>Unrestricted participation</th><th>Set-aside for small bidders</th></tr>
<tr><td>1st Price (competitive bidders)</td><td>Estimate
• Value distributions (large and small bidders)
• Entry costs</td><td rowspan="3">Predict out of sample
• Bids, profit cond’l on entry
• Entry

Test model
• Actual entry, revenue v. predictions</td></tr>
<tr><td>Ascending (competitive bidders)</td><td rowspan="2">Predict out of sample
• Bids, profit cond’l on entry
• Entry
Test model
• Actual entry v. prediction
Model selection
• Competitive v. collusion differ in bids/prices
• Select model based on fit</td></tr>
<tr><td>Ascending (collusive bidders)</td></tr>
</table>

Experiments and quasi-experiments:

Generate 1st price data for identification/estimation of structural model (ascending was standard practice)

Enable model validation and selection among alternatives

Model Validation: Set-Asides

	Actual Outcome	Predicted Outcomes		
		Unrestricted	Set-Aside	*In Sample Validation:* Actual/ Predicted Unrest.
Panel A: Tracts sold by unrestricted sale				
Avg. Small Bidder Entry	2.71	2.69	4.93	1.01
Avg. Big Bidder Entry	1.55	1.57	0.00	0.99
Avg. Total Entry	4.26	4.26	4.93	1.00
Avg. Prices	95.12	98.51	95.87	0.97
% Sales won by Small Bidders	52.44	54.14	97.93	0.97
Avg. Sale Surplus (per mbf)		119.16	100.11	
				Out of Sample Validation: Actual/ Predicted Set-Aside
Panel B: Tracts sold by set-aside sale				
Avg. Small Bidder Entry	4.50	2.57	4.71	0.95
Avg. Big Bidder Entry	0.00	1.71		
Avg. Total Entry	4.50	4.28		
Avg. Prices	90.74	97.29	92.01	0.99
% Sales won by Small Bidders	100.00	50.63		
Avg. Sale Surplus (per mbf)		119.21	98.39	

Model Validation and Selection: Auction Format

1st Price Sales		Actual	Predicted	*In Sample Validation:* Actual/ Predicted
	Avg. Bid	59.6	57.4	1.04
	Avg. Sale Price ($/mbf)	69.4	70.4	0.99
	Avg. Logger Entry	3.2	3.2	1.00
Ascending Sales				***Out of Sample Validation and Model Selection:*** **Actual/ Predicted**
	Avg. Sale Price (Competition)	63.3	67.8	0.93
	Avg. Sale Price (Collusion)	63.3	44.1	1.44
	Avg. Logger Entry	2.75	2.67	1.03

Both competition and collusion are rejected at 5% level.

Better fit from a partially collusive model (e.g. collusion in 18% of auctions), consistent with industry anecdotes.

Model Selection: Additional Evidence on Collusion

Actual versus Predicted Sale Prices by Mill Participation			
	Actual	**Predicted**	**Actual/ Predicted**
No Mills			
1st Price	51.7	51.4	1.0
Ascending	49.8	47.1	1.1
Mills Present			
1st Price	89.8	91.0	1.0
Ascending	73.8	81.6	0.9

Observed revenue shortfalls occur only in auctions where mills are present, consistent with competitive behavior among small loggers and mild collusion by mills.

What is Important in Timber Market Design?

- Exogenous entry
  - Effects of auction format on surplus and allocation are small
  - Small revenue effects with competitive bidding (<1%).
  - Large revenue effects from reducing mill collusion (≈10%)
- Equilibrium entry increases differences
  - 1st price leads to >10% more loggers
  - Revenue advantage of 1st price is 3x larger when entry accounted for
- Preference policies
  - 46% of unrestricted sales won by large firms
  - Endogenous entry implies that set-asides attract 75% more small entrants, leading to only modest revenue shortfall (6%)
  - Counterfactual subsidy policy can do better for all constituents: bidders, revenue and efficiency

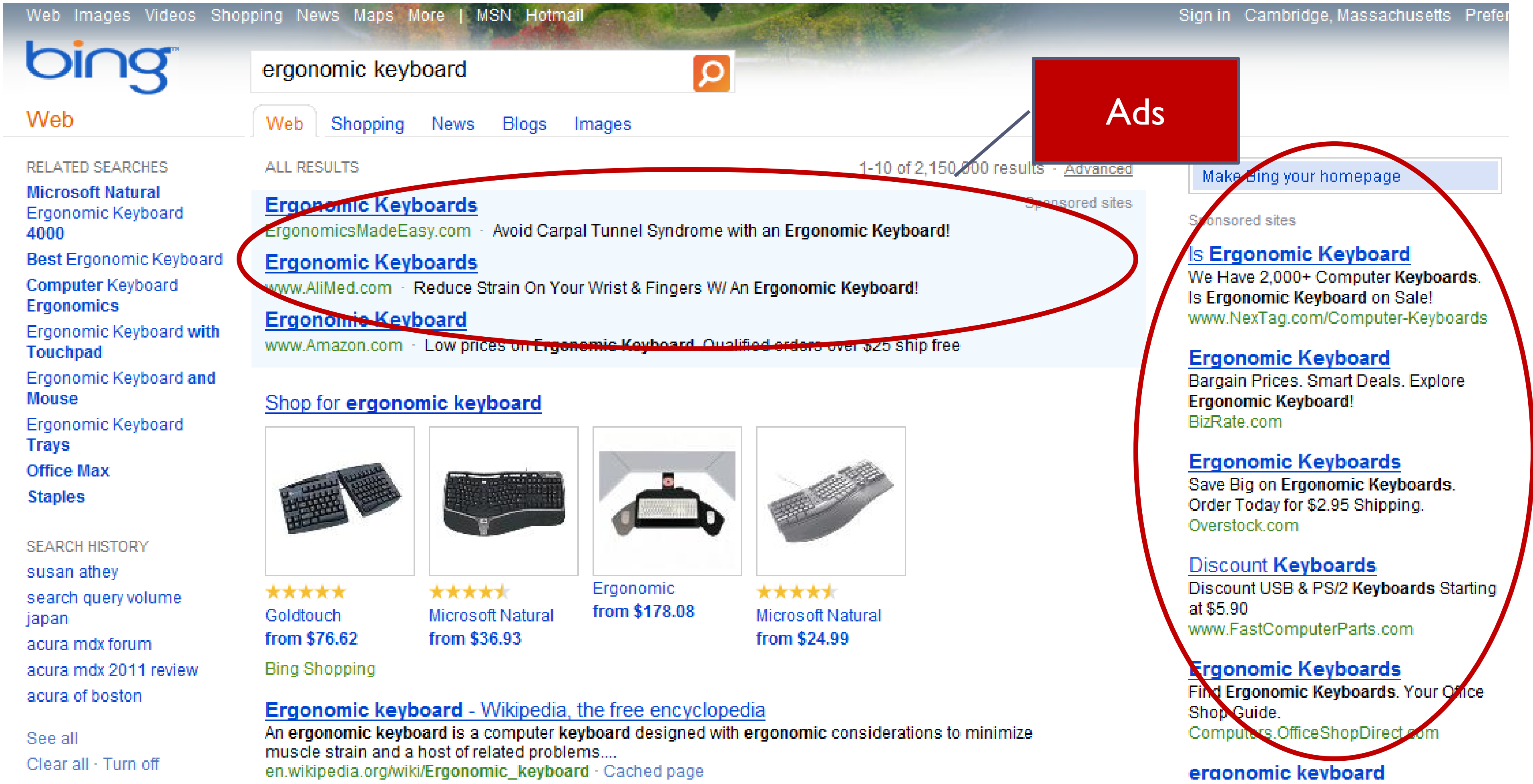
Case Study: Internet Search Advertising

Athey and Ellison, *QJE* 2011

Athey and Nekipelov, working paper

New applications and extensions

Internet Search and Search Advertising



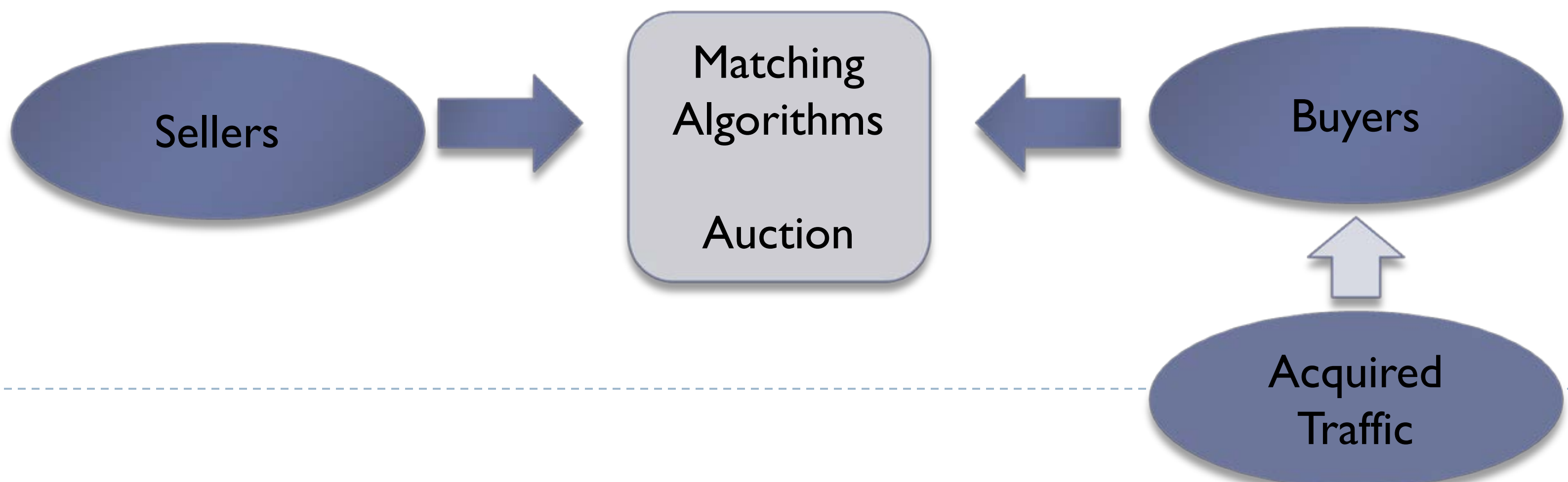
Auction-based Platforms

- Platform markets
  - Media markets, credit cards, dating, video games, operating systems…
  - Two groups of customers, externalities (typically indirect network effects)
    - Platform can help internalize externalities
- Auction-based platforms
  - Online advertising
  - Used cars
  - eBay
- Market design matters
  - Participation is crucial

Visit our eBay Store

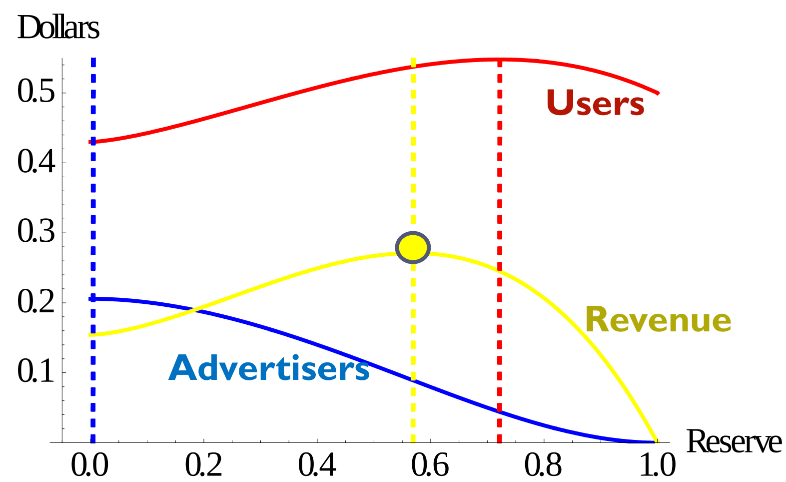
Auction-Based Platforms

- Like other two-sided markets, except…
  - Auction matches buyers and sellers, determines transaction price
  - Method for sorting as well as a method for platform to extract revenue
  - Limits discretion of market-maker
  - Still, rules, reserve prices and fees play an important role in determining the size and distribution of "the pie"



Balancing Constituents in Search Advertising: Endogenous Ad Clicking and Ad Footprint

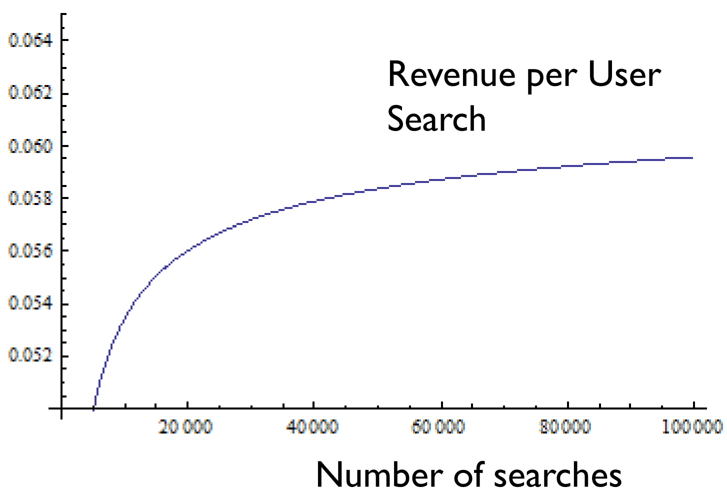
- Key features of medium-run model
  - Users
    - Number of searches fixed
    - Search costs of clicking on ads
    - Users click more on ads when avg quality is higher
    - Rationally anticipate higher avg quality when reserve price is higher
    - Opportunity cost of ad space: when algo content is pushed down the page, users see less benefit from it
  - Advertisers
    - Entry /budget fixed in short run; choose bids
    - Profit when user makes a purchase; probability of match (per click) goes up when reserve price rises
  - Extension of Athey and Ellison, *QJE* 2011
- Impact of raising reserve prices
  - Initially, allows better-quality algo results to appear
  - Eventually, ads are better than algo and users harmed
  - Advertisers and medium-term revenue in conflict



Balancing Constituents in Search Advertising: Long Run Participation Determined by Medium-Run Welfare

Long run competition for users

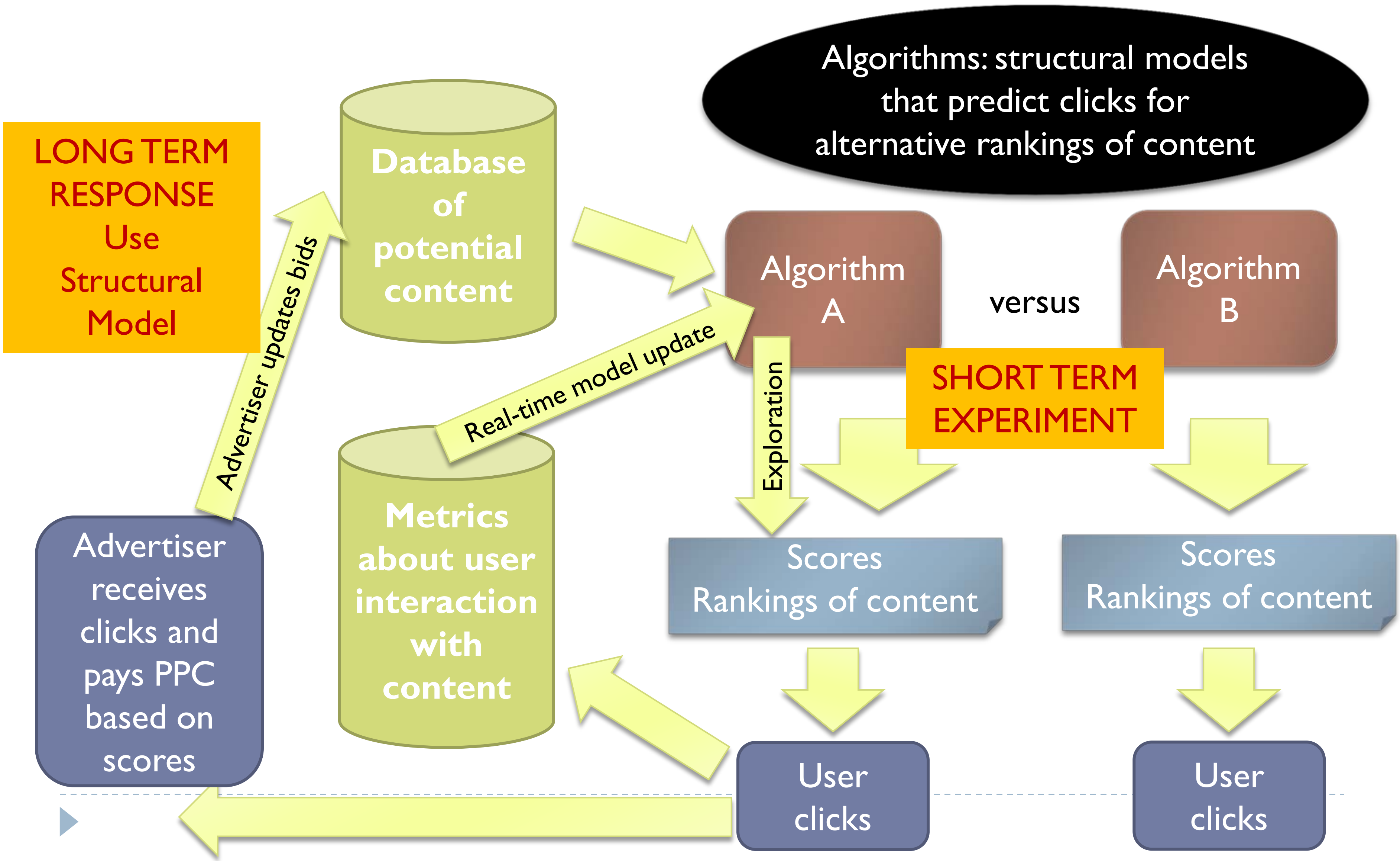
- Number of searches depends on
  - Perceived average quality (#, quality of ads)
  - Idiosyncratic preferences or convenience
- Comparative Statics
  - Platform with better baseline quality typically more willing to distort page quality



Long run competition for advertisers

- Advertiser choices
  - Entry, campaign management:
    - Compare costs to profit conditional on entry/investment
  - Shift limited budgets, campaign management resources to most profitable platform
- Comparative statics
  - Platform with more users attracts more advertisers
  - Diminishing marginal returns to additional advertisers in terms of efficiency and revenue
    - Both depend primarily on best few advertisers (top order statistics)
  - Platform with more users can provide the same total profit to advertisers (users x advertiser profit per user) at higher reserve prices

Experiments and Models in Internet Search: Click Prediction Case Study



Methodology for Evaluating Design Change

1. Structural model of user choice among ads
   - Use *user-level experiments* to identify and estimate models
2. User model used to predict short-term CF impact on clicks of changing ad rankings with new algorithm
3. Structural model of advertiser response to new algorithm
   - Takes user model as input
   - Athey-Nekipelov model:
     - Values identified from past auction data
     - New equilibrium can be computed
4. Validate model after new algorithm adopted
   - Compare predicted bids to actual bids
   - See Ostrovsky & Schwarz for large-scale application at Yahoo!

User Model of Clicks

- Experiment only: too expensive, not sufficient
  - Too many permutations of ads
- Dedicated experiment + model
  - For new, very different algorithm
- Re-use existing experiments + model
  - Any experiment that generates exogenous variation in ad rankings is a candidate

- Basic model

$$1\{click_t\} = \sum \alpha_i \gamma_j 1\{ad_t = j,\ pos_t = i\} + \epsilon_t$$

- Identification
  - Problem with observational data
    - More "clickable" ads in high positions
    - Ad ranking respond to indiv. users
  - User-level experiments provide identification
    - Need experiments to shift probability of appearing in each position
    - Random assignment of users to experiments
- Two-step approach
  - Estimate $c_{ij} = \alpha_i \gamma_j$ using past user-level experiments x ads as instruments
  - Decompose c_{ij} into components

User Model of Clicks: Results from Spring 2009 Experiments

Clicks as a Fraction of Top First Position Clicks

Search phrase:		iphone			viagra	
Model:		OLS	IV		OLS	IV
Top Position 2		0.66	0.67		0.28	0.66
Top Position 3		0.40	0.55		0.14	0.15
Side Position 1		0.04	0.39		0.04	0.13

IV estimates show smaller position impact than OLS, as expected.

Since position discounts affect advertiser bid shading, this is very important for counterfactual advertiser models.

Click-weighted generalized second-price auction

- Pay per click v. pay per position/impression
  - Selling *impressions* in different positions leads to thin markets
  - Selling *clicks* reduces transaction costs, makes heterogeneous goods homogeneous, thickens market
  - "Conflation" (Levin and Milgrom, 2010)
- Price for position *m* determined using $m + 1^{st}$ revenue per *impression*
  - Multiply per-click bid by "clickability" score, *s*, to get per-impression bid
  - "Clickability" is the click-through rate if ad were to be shown in top position
- Pay minimum to maintain position

Per-Click Bid	Rank Score: Estimated Revenue Bid (normalized to 1st position)	Price Per Click	Estimated Revenue (including position discounts α_i)
b_1	$b_1 s_1$	$b_2 (s_2 / s_1)$	$b_2 s_2$
b_2	$b_2 s_2$	$b_3 (s_3 / s_2)$	$\alpha_2 b_3 s_3$
b_3	$b_3 s_3$	R / s_3	$\alpha_3 R$

More Accurate Click Prediction Increases Efficiency, But Can Decrease Revenue

Per-Click Bid	Estimated Revenue Bid (normalized to 1st position)	Price Per Click	Platform Revenue (including position discounts)	Expected Clicks (including position discounts)
"Coarse" click predictor with two bidders with equal bids, avg. clickability				
b	$b\ s$	b	$b\ s$	s
b	$b\ s$	R/s	$\alpha_2\ R$	$\alpha_2\ s$
"Granular" click predictor identifies user types: Half of users like A better so true score is $(1+d)\ s$ for A and $(1-d)\ s$ for B Half of users like B better so true score is $(1+d)\ s$ for B and $(1-d)\ s$ for A				
b	$b\ (1+d)\ s$	$b\ (1-d)\ /(1+d)$	$b\ (1-d)\ s$	$(1+d)\ s$
b	$b\ (1-d)\ s$	$R/((1-d)\ s)$	$\alpha_2\ R$	$\alpha_2(1-d)\ s$
Differences in outcomes: "Granular" – "Coarse"				
			$-d\ s$	$(1-\alpha_2)d\ s$

Reducing Click Prediction Accuracy Has Competing Effects: Reduced Welfare, Increased Revenue

Algorithm:	**Before Coarsening**	**After Coarsening**	**After Coarsening**
Bids:	**Original**	**Original**	**Counterfactual**
Welfare	6.959	6.713	6.715
Revenue	2.091	2.034	2.136
Advertiser Profit	4.869	4.679	4.579

Revenue falls in short run experiment

Revenue rises when bids adjust

Crucial role for advertiser modeling:
Short and long run revenue predictions go in opposite directions.

Conclusions

- Market design
  - Extent of participation and engagement crucial
  - Tradeoffs between efficiency, participant welfare, and revenue
- Methods
  - Theory, structural models and experiments
  - Complements, not substitutes
  - In online world, models and experiments are ubiquitous, integral and essential
- Academic Research v. Practice in Online Market Design
  - Issues, methods and debates surprisingly similar in two arenas
  - Debates about validity of structural models for long-term predictions
  - Differences
    - No debate that both experiments & models crucial in practice
    - Behavioral assumptions more controversial in practice (e.g. profit max)
    - Experimental methods more developed in practice
    - Econometric insights about causality and best ways to use observational data not fully appreciated in practice