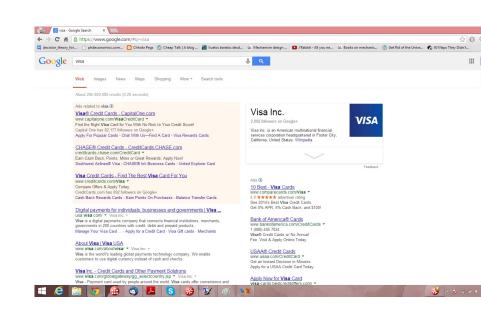
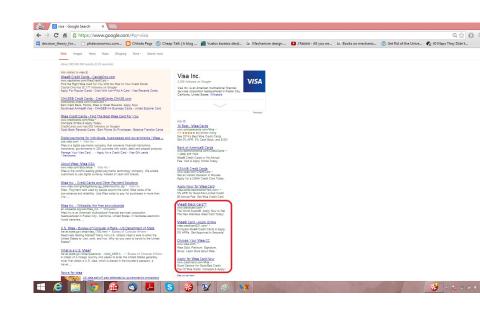
A Structural Model of Sponsored Search Advertising Auctions

S. Athey and D. Nekipelov

presented by Marcelo A. Fernández

February 26, 2014





Remarks

► The search engine does not sell specific positions on the page.¹

- ▶ The search engine only gets paid if you click on the ad.
- Instead of selling a lottery, it is selling a contingent good, which it does not own or control.

 $^{^{1}}$ or at least not explicitly, it can be the case in equilibrium. $\stackrel{>}{\scriptstyle{\bullet}}$ $\stackrel{>}{\scriptstyle{\bullet}}$ $\stackrel{>}{\scriptstyle{\bullet}}$ $\stackrel{>}{\scriptstyle{\bullet}}$ $\stackrel{>}{\scriptstyle{\bullet}}$ $\stackrel{>}{\scriptstyle{\bullet}}$

Sponsored Search Auctions - Overview

- ▶ Bidders enter per-click (standing) bids into a database.
- ► Each time a user enters that query, the bids for that query are called from the database and enter an auction
- ▶ Bids are ranked(*) and per-click prices are determined.
- ► Ads are displayed in rank order for a fixed number of slots.

Sponsored Search Auctions - Rank

- ► The search engine wants to rank the alternatives such that it maximizes its profit.
- ► This does not necessarily imply that the highest bidder should get the first position on the webpage.
 - ► The number of clicks received by an ad and the position it occupies in the web page are highly correlated empirically.
 - Suppose highest bidder gets the first position, but is never clicked upon. The search engine sees zero income, since the contingent good (click) was not delivered.
 - The search engine may want to allocate the first position to a lower bid that receives a high number of clicks.
- ▶ Rank ads by "expected revenue", in practice attach a weight to the bids, named "quality scores.", estimation of click-through-rates, fraction of the queries that click on the ad if its on the first position.

Sponsored Search Auctions - Prices

- Generalized Second Price Auction:
 - ► After being ranked by their quality-score-adjusted bids, a bidder pays the minimum per-click price required to keep the bidder in her position.

a priori difference from previous literature

- Incorporate real-world observation: bids change slower than quality scores and the rate at which queries arrive.
- Compare the model that takes this observations into account (SEU) against the one that doesn't (NU).

Goal

- Estimate valuation of bidders.
- Perform Counterfactual experiments:
 - Evaluate models by their out of sample predictions.
 - ▶ Re-evaluate statistics on bidders profits and competition.
 - Compare efficiency and revenue between Vickrey and SEU.
 - ► Evaluate different quality scores.

Basic Setup

- Static model.
- ▶ Each $i \in I$ posts bid b_i for a single phrase for a fixed period of time (e.g. a week)
- Fixed J slots available.
- ▶ Consumer search is a random process: \bar{c}_{ij} is the probability that a consumer clicks on a particular ad i in position j, same for all consumers.
- $ightharpoonup \bar{c}_{ij} = \alpha_j \gamma_i$
- ▶ Per-click price P_{k_i} that bidder k_j in position j pays is:

$$\begin{aligned} p_{k_j} &= \min\{b_{k_j} : s_{k_j} b_{k_j} \ge s_{k_{j+1}} b_{k_{j+1}} \} \\ &= \frac{s_{k_{j+1}} b_{k_{j+1}}}{s_{k_i}} \end{aligned}$$

Equilibrium with No Uncertainty (NU)

- bidders known:
 - ▶ the set of competitors,
 - the score-weighted bids of opponents.
- ex post Nash equilibrium, typically not unique nor monotone
- ► Focus on equilibrium refinement "envy-free equilibrium" which are monotone:

$$s_{k_j}v_{k_j} \geq ICC_{j,j+1} \geq s_{k_{j+1}}v_{k_{j+1}}$$

where

$$ICC_{j,j+1} = \frac{s_{k_{j+1}}b_{k_{j+1}}\alpha_j - s_{k_{j+2}}b_{k_{j+2}}\alpha_{j+1}}{\alpha_j - \alpha_{j+1}}$$

Equilibrium with No Uncertainty (NU)

► NU-EFLB (envy free lower bound):

$$s_{k_j}v_{k_j}=ICC_{j,j+1}$$

► NU-EOS (lowest revenue for auctioneer and coincides with Vicrey):

$$ICC_{j,j+1} = s_{k_{j+1}} v_{k_{j+1}}$$

bidding strategies are not truthful.

Equilibium with Uncertainty (SEU)

- Uncertainty in:
 - Score statistical algorithm updates faster than bid change.

$$s_i = \bar{s_i} \epsilon_i$$

where ϵ_i is a shock tot the score by a random variation of the algorithm

- Bidder Entry budget limits, multiple ads, demographic targeting:
 - $C^i \in \tilde{C}^i$ is the realization of the random subset of participants other than i

Equilibium with Uncertainty (SEU)

Assume bidders correctly anticipate:

- ▶ distribution of C_i
- mean of opponent's $b_i \bar{s}_i$
- distribution of score shocks (i.i.d. and usual "desirable" properties)

Equilibium with Uncertainty (SEU)

Bidder's problem is to maximize:

$$EU_i(b_i; b_{-i}, \overline{s}) \equiv v_i \cdot Q_i(b_i; b_{-i}, \overline{s}) - TE_i(b_i; b_{-i}, \overline{s})$$

where:

 $Q_i(b_i; b_{-i}, \bar{s})$ denotes the expected number of clicks to be received with bid b_i

 $TE_i(b_i; b_{-i}, \bar{s})$ denotes the expected total expenditure of the advertiser for the clicks received with bid b_i



Theorem 1

(Under a set of conditions) The equilibrium of the GSP auction in the SEU environment exists and is unique.

Equilibrium not necessarily envy-free nor monotone (although latter holds for enough large number bidders and sufficient uncertainty.)

Bidders incentives

The bidders problem can be rewritten as $\max_{q_i} q_i(v_i - AC_i(q_i))$ with FOC:

$$v_i=q_iAC_i'(q_i)+AC_i(q_i)\equiv MC_i(q_i)$$
 where $AC_i(q_i)=rac{TE_i(Q_i^{-1}(q_i))}{q_i}$ and $Q_i^{-1}(q_i)=\inf\{b_i:Q_i(b_i)\geq q_i\}$

Identification

- In NU model:
 - identify score-weighted valuations for each bidder that lie between the steps of the ICC curve. Generally bounds, not point-dentified:

$$s_{k_j} v_{k_j} \in [ICC_{j,j+1}, ICC_{j-1,j}]$$

- identification of valuations per query. Which allows the bounds not to be consistent across queries.
- ▶ In SEU model:
 - if $TE(\cdot)$ differentiable $v_i = MC_i(Q_i(b_i))$ since distributions required to calculate $MC_i(q_i)$ are observable.
 - ▶ if $TE(\cdot)$ not differentiable and $b_i \in [b'_i, b''_i)$ then $v_i \in [MC_i(Q_i(b'_i)), MC_i(Q_i(b''_i))]$

$$v_i = \frac{\frac{\partial TE_i(b_i; b_{-i}, \overline{s})}{\partial b_i}}{\frac{\partial Q_i(b_i; b_{-i}, \overline{s})}{\partial b_i}}$$

$$Q_{i}(b_{i}; b_{-i}, \bar{s}) = E_{\tilde{C}^{i}, \epsilon} \left[\sum_{j} \sum_{k \in \tilde{C}^{i}} Pr(\Phi_{ik}^{j}(b, \bar{s}, \epsilon; \tilde{C}^{i}) = 1) \cdot \alpha_{j} \gamma_{i} \right]$$

$$TE_{i}(b_{i}; b_{-i}, \bar{s}) = E_{\tilde{C}^{i}, \epsilon} \left[\sum_{j} \sum_{k \in \tilde{C}^{i}} Pr(\Phi_{ik}^{j}(b, \bar{s}, \epsilon; \tilde{C}^{i}) = 1) \cdot \alpha_{j} \gamma_{i} \frac{\bar{s}_{k} \epsilon_{k} b_{k}}{\bar{s}_{i} \epsilon_{i}} \right]$$

where Φ^j_{ik} is an indicator of bidder i is in slot j and bidder k is in slot j+1:

$$\Phi^j_{ik}\left(b,\overline{s},\varepsilon;C^i\right) = \sum_{C^i_{j,k} \in \mathcal{C}^i_{j,k}} \prod_{m \in C^i_{j,k} \setminus \{k\}} \mathbf{1} \left\{ b_m \overline{s}_m \varepsilon_m > b_i \overline{s}_i \varepsilon_i \right\} \prod_{m \in C^i \backslash C^i_{j,k}} \mathbf{1} \left\{ b_m \overline{s}_m \varepsilon_m < b_k \overline{s}_k \varepsilon_k \right\} \mathbf{1} \left\{ b_i \overline{s}_i \varepsilon_i > b_k \overline{s}_k \varepsilon_k \right\}.$$



Use empirical distribution of:

- scores to approx uncertainty in scores.
- bidder configurations to approx uncertainty in bidder configurations.

Use assumption 2 ($E[log(\epsilon_{it})] = 0$) to get $E[log(s_{it})] = log(\bar{s_{it}})$), so you can estimate the mean score from the observed realizations of the scores $\hat{s_i}$. Form the sample of estimated shocks to the scores by $\hat{\epsilon}_{it} = \frac{s_{it}}{\hat{s_i}}$ and compute the empirical distribution of shocks. Both are consistent estimators.

- 1. Take draws from the empirical distribution of configurations and shocks to the scores.
- 2. Compute the rank of the bidder.
- 3. Compute the price paid.
- 4. Estimate the total expenditure $(T\hat{E}(\cdot))$ function as (simulated) sample average across draws.
- 5. Estimate the expected quantity of clicks $(\hat{Q(\cdot)})$ as (simulated) sample average across draws.
- 6. Compute the empirical numerical derivative.
- 7. Obtain \hat{v}_i

Theorem 4

The estimators are consistent. Under the sufficient conditions of Theorem 1 and Assumption 2, the estimates of valuations are asymptotically normal.

Data - General Description

- ▶ 2 high-value search phrases.
- 3-month period between 2006 and 2009.
- ▶ 7500 searched per week between both phrases.
- 8ads displayed.
- data used only from one week at a time.
- details preserved for confidentiality purposes.

Data - Observables for each user query

- Advertiser and specific ad.
- position of the ad.
- per-click bid and system-assigned score.
- per-click-prices.
- clicks received by each advertiser.

NU-EOS overestimates value per click compared to SEU.

NU-EFLB underestimates value per click compared to SEU.

Table 1: Mean of valuations for different models across search phrases

	All bidders		Top bidders in >50% queries excluded		
Model; Search phrase	#1	#2	#1	#2	
NU-EFLB	0.047	0.217	0.045	0.189	
NU-EOS	0.136	0.451	0.070	0.276	
SEU	0.059	0.292	0.057	0.261	

We report the means of recovered valuations across search phrases and bidders for both of the models, where for the NU models we use the modified (monotone) ICC curves. The values are normalized by the highest observed bid for search phrase #1.

NU-EOS overestimates per query profit per click compared to SEU.

NU-EFLB underestimates perquery profit per click compared to SEU.

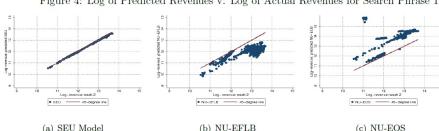
Table 2: Lerner Index Avg. (COC) (advertisement is unit of analysis)

All bidders			Top bidders					
					in >	50% que	ries excl	uded
Model	Mean	25%	50%	75%	Mean	25%	50%	75%
	Sear	rch phra	se #1, a	dvertisers	equally we	ighted		
NU-EFLB	0.143	0.020	0.090	0.219	0.134	0.019	0.086	0.217
NU-EOS	4.365	0.197	0.379	0.666	0.632	0.195	0.371	0.647
SEU	0.420	0.259	0.352	0.503	0.400	0.259	0.347	0.485
	Search ph	rase #1	, advert	isers weigh	ted by exp	ected cl	icks	
NU-EFLB	0.681	0.141	0.157	1.536	0.153	0.141	0.141	0.141
NU-EOS	10.921	0.686	4.456	20.753	2.652	0.314	4.456	4.456
SEU	1.810	0.468	1.566	3.111	1.018	0.289	1.566	1.566
	Sear	rch phra	se #2, s	dvertisers	equally we	ighted		
NU-EFLB	0.289	0.076	0.176	0.449	0.266	0.071	0.161	0.384
NU-EOS	1.715	0.259	0.514	1.035	0.810	0.244	0.488	0.871
SEU	0.735	0.319	0.506	0.879	0.709	0.312	0.483	0.829
	Search ph	rase #2	, advert	isers weigh	ted by exp	ected cl	icks	
NU-EFLB	0.644	0.559	0.718	0.817	0.563	0.232	0.711	0.730
NU-EOS	14.310	1.108	1.603	13.876	1.245	0.831	1.436	1.603
SEU	1.173	0.679	1.066	1.731	1.227	0.776	1.184	1.731

Counterfactual experiments

Taking valuations from each model using one week of data, and predicting revenue for the following week.

Figure 4: Log of Predicted Revenues v. Log of Actual Revenues for Search Phrase 1



(c) NU-EOS

Table 3: Mean squared deviation of the predicted revenues per query from the true revenues (norm by true mean revenue per query)

Model for equilibrium (Model for values)

Moon 25% 50% 75%

Model for equilibrium (Model for values)	Mean	25%	50%	75%	
Search phrase #1					
SEU (SEU)	0.00011	0.00000	0.00001	0.00002	
NU-EFLB (SEU)	0.34054	0.26660	0.38197	0.44344	
NU-EOS (SEU)	0.01593	0.00229	0.00446	0.00875	
NU-EFLB (NU-EFLB)	0.15941	0.10394	0.18185	0.21074	
NU-EOS (NU-EOS)	0.24965	0.13932	0.22225	0.28404	
NU-EFLB (NU-EOS)	5.89371	5.29678	5.51929	5.75895	
NU-EOS (NU-EFLB)	0.45432	0.37673	0.51380	0.57772	
Search ph	rase #2				
SEU (SEU)	0.00095	0.00003	0.00022	0.00108	
NU-EFLB (SEU)	0.02870	0.00214	0.01135	0.03830	
NU-EOS (SEU)	0.15713	0.09096	0.14865	0.21837	
NU-EFLB (NU-EFLB)	0.02073	0.00262	0.00972	0.02126	
NU-EOS (NU-EOS)	0.16492	0.00151	0.01211	0.05811	
NU-EFLB (NU-EOS)	37.47682	0.14496	2.63420	23.88841	

 $0.12617 \quad 0.04462$

0.09855

0.17046

NU-EOS (NU-EFLB)

Table 5: Predicted revenues and welfare for the SEU generalized second price auction model versus the Vickrey auctionl, using SEU values and actual bidder configurations: with decomposition by bidders in different positions

	Positions						
Model (values)	All	1	2-5	6-8			
Search phrase #1							
Revenue SEU (SEU)	2.490	1.583	0.886	0.021			
Revenue Vickrey=NU-EOS(SEU)	1.947***	1.032***	0.890**	0.026			
Welfare SEU (SEU)	8.228	6.358	1.838	0.032			
Welfare Vickrey=NU-EOS (SEU)	8.231	6.365	1.834	0.032			
Search phrase #2							
Revenue SEU (SEU)	3.114	1.846	1.203	0.066			
Revenue Vickrey= NU - $EOS(SEU)$	3.153***	1.910***	1.152***	0.091			
Welfare SEU (SEU)	6.513	3.433	2.957	0.122			
Welfare Vickrey=NU-EOS (SEU)	6.531	3.459	2.950	0.122			

This table represents the expected total per query revenue of the auction platform and the expected per query social welfare. For all counterfactuals, we used SEU estimated values. To compute the expected clicks we used our estimated position clickthrough rates and the scores of the advertisers. Values and cost per click are normalized by the maximum bid for the search phrase # 1. To group the SEU welfare and SEU revenue by positions we used the positions based on SEU equilibrium bids. To group the EOS welfare and EOS revenue by positions we used the positions based on EOS equilibrium bids. The table also represents the results of the hypothesis test for the difference in expected total per query revenue of the auction platform and the expected per query social welfare in NU-EOS model versus SEU model, using asymptotic standard errors. * - corresponds to significance at 10% level, ** - significance at 5% level, and *** - significance at 15% level, and *** - significance at 15% level,

- ► Squashing: 1st phrase -2%revenue, -.5%efficiency.
- ► Squashing: 2nd phrase +9%revenue, -4.5%efficiency.
- ► Coarsening: 1st phrase +2%revenue, -1%efficiency.
- ► Coarsening: 2nd phrase +18%revenue, -2%efficiency.
- ► Coarsening: trade-off between efficiency and competition. More targeted ads, have less competition

Conclusions

- Model closer to real-world application.
- Neater identification (equilbrium uniqueness).
- Bidder valuations and profits lower than recognized by prev. literature.
- Better predicting power out-of-sample.
- Allows comparisons between auction formats.
- Allows assessment of different scoring methods.