Output Bidding: A New Search Advertising Model Complementary to Keyword Bidding

[Position Statement]

Ali Dasdan[†] Santanu Kolay[†] Panagiotis Papadimitriou[‡] Hector Garcia-Molina[‡]

[†]Yahoo! Inc.
Sunnyvale, CA, USA
{dasdan,santanuk}@yahoo-inc.com

[‡]Stanford University
Stanford, CA, USA
{ppapadim@,hector@cs.}stanford.edu

ABSTRACT

A search engine can be modeled as a mapping that takes in user keywords as input and produces search results as output. Currently, the most dominant form of search advertising is *input bidding* where advertisers bid to associate their advertisements with keywords. We propose *output bidding* where advertisers bid to associate their advertisements with search results without changing or replacing any search result. We argue how the two forms of bidding complement each other as well as how output bidding also ties in contextual advertising to search advertising. We propose three basic variations with examples: associating advertisements with output from the same site, using output to make more expressive input bidding, and associating advertisements with output in general. We experimentally show the potential of output bidding using real data.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: Search Process; H.4.m [**Information Systems Applications**]: Auction Theory

General Terms

Advertising, economics, algorithm

Keywords

Keyword advertising, keyword bidding, search auctions, search advertising, sponsored search

1. INTRODUCTION

Search engines offer their service free of charge because of the revenue they earn from search engine advertising [11]. The most dominant form of search engine advertising is keyword advertising [11] in which search engines run a continuous auction to sell advertisement (ad) space on the search engine results page (SERP). In this form of auction, advertisers bid on keywords in user queries and search engines determine the winning bids using a combination of the bid amounts and the clickability of the associated ads. In the end, search engines earn money when users click on the ads and advertisers benefit when these clicks result in a purchase.

AdAuctions2009, July 6, 2009, Stanford, CA.

To describe output bidding [7], we look at a search engine as a mapping from *input* user queries (keywords) to *output* search results. The output data contains organic and sponsored results on the SERP, including ads, as well as various metadata associated with results and input query. For example, URLs in results, listed site categories, or the facets of input query are part of the output data. See Figure 1 for a screen shot of a SERP with some of its sections identified.

We refer to the current keyword bidding as *input bidding* since advertisers bid on terms in the *input*. Analogously, with *output bidding* advertisers bid on terms anywhere in the *output*. For example, an advertiser may wish to have his ad appear whenever a particular URL appears in the SERP, regardless of the input query, and hence bids on that URL. Similarly, he may bid on particular site categories like business sites. Output bidding can be used in conjunction with input bidding. For instance, an advertiser may bid on some keywords, with the condition that the first three organic results do not contain any of a set of prohibited URLs.

Intent. We note that advertisers actually want to bid on *user intent*, especially purchasing intent; keywords are but one signal for understanding this intent. We then claim that output is a far richer and superior signal towards the same goal. The richness is obvious from the amount of the output data and metadata compared to what a few keywords in input can provide. The superiority is there because the output consists of all the information that a search engine, with its massive investment in its infrastructure, generates as a response to its input. Note that output does include the metadata of input too.

Association. We note that for input bidding, since advertisers have no control over input, they can only bid to *associate* their ads with input. We then claim that for output bidding, since advertisers have no control over output, they can only bid to associate their ads with output. Moreover, the organicity of output is beneficial to all parties involved, including users, advertisers, search engines, and content publishers; for example, search engines have a strong incentive not to game the output for short-term benefit.

2. BASIC VARIATIONS

We consider three basic variations of output bidding with increasing complexity and possibly benefit. We again emphasize that none of these variations imply any change to the organicity of the search results. More details and examples can be found in [7].

1. Paid association. Advertisers are allowed to associate advertisements with search results coming from their own sites. For example, a home improvement company may want to announce the new discounts on "air conditioners" around an organic search result from its site. This ad may or may not be query dependent. Note that this variation is different from paid inclusion [11] because the organic result stays organic in this variation.

One way to consider about paid association is for content providers to take advantage of the popularity of their content. In other words, if a provider invests in resources to make their content increasingly more relevant to a search engine, the provider can get extra monetary benefit through paid association on top of the referrals they get from search engines for their organic content.

2. More expressiveness in input bidding. In one form, advertisers can bid only on input but can use output to condition their bid. For example, a company selling accessories for a cell phone brand X may bid on the query "cell phone" but may not want its ads to be shown when a SERP contains no mention of X. This variation requires improvements to the bidding language to allow the specification of conditions on bids by advertisers.

In another form, input is expanded with keywords from output and advertisers bid as usual on input. Various information retrieval techniques can be used to decide on which keywords to extract from output. Another consideration can be given to keywords that have had some affinity to ads or bid terms in the past. Note that this variation is similar to contextual matching with output context to match. Also, this can help bypass the restrictions on brand name bidding.

3. Bidding on output. Advertisers are allowed to actually bid for output directly. For example, two companies selling a cell phone brand X can bid to associate their ads with the appearance of a search result on X, independent of input. They can even tie their bidding to the brand or the site of the search result if brand name bidding is allowed. In another form, it is also possible to factor in the brand owner to have a say in the auction for this search result, e.g., through revenue sharing. We leave the details of this last form to a future publication.

3. USER INTERFACE ISSUES

For all three variations, one basic issue is where to show the new ads if they need to be associated with output directly. This naturally requires studies with users and advertisers to finalize. For now, one possibility is to use the unused space in the sponsored section of a SERP. Another possibility is to actually insert a clearly marked sponsored link close to the associated search result. The marking can involve different background or foreground colors and indentation. For an example of how to combine shopping and business results with organic results, see Fig. 1. Finally, we again emphasize that we do not advocate the use of paid inclusion or preferred listings [11] in which advertisers would pay to have their advertisements appear among the organic search results.

4. AUCTION MODELING ISSUES

Although separate auctions can possibly be run for organic and sponsored sections, interesting questions arise with output bidding: What happens if an advertiser wants to bid on more than one item (a bundle) or multiple advertisers can be shown for the same item? What if the same advertiser wants to bid on both input and output simultaneously?

Full generality in answering these questions seems to require techniques from combinatorial auctions [6] but with appropriate restrictions on the decomposition of bids and bundle sizes, the current auction models [2, 3, 10, 21, 22] may be used. For bundles with at most one organic result and at most one sponsored result, one heuristic solution is proposed in [8]. More research is needed in this area. However, in the short term, we can think of two potential solutions that can take advantage of the current auction models.

In one solution, which fits well with the second variation involving the input expansion with data from output, the current keyword auction models can be used "as is". This is possible because all terms, irrespective of where they come from: input or output, can be treated as regular keywords. Note that even though the same auction models can be used, the implementation may have some challenges as we will discuss next.

In another solution, which fits well with the third variation involving direct output bidding, the current auction models for contextual matching can be used. This is possible because output in a SERP can simply be considered as a context to match ads against.

5. IMPLEMENTATION ISSUES

In input bidding, a search engine can deal with organic and sponsored results separately. Output bidding creates a dependence between them. Although a naive implementation may serialize these parts, a careful implementation can hide most of the latency using offline analysis and caching. Note that search engines already benefit significantly from caching for serving organic results and ads; therefore, we expect that caching will also help for this integration of organic and sponsored search.

Offline analysis is needed for figuring out the best candidates for output bidding. For example, it can be used for extracting keywords from search result pages as bidding candidates [4]. These keywords are always valid for their source page. The implication for serialization appears when the analysis results need to be merged and summarized over all search results in a SERP. In this case, the merging and summarization implies serialization because the SERP needs to be built first. Since the ads processing part usually takes less time than the organic results processing part, the se-

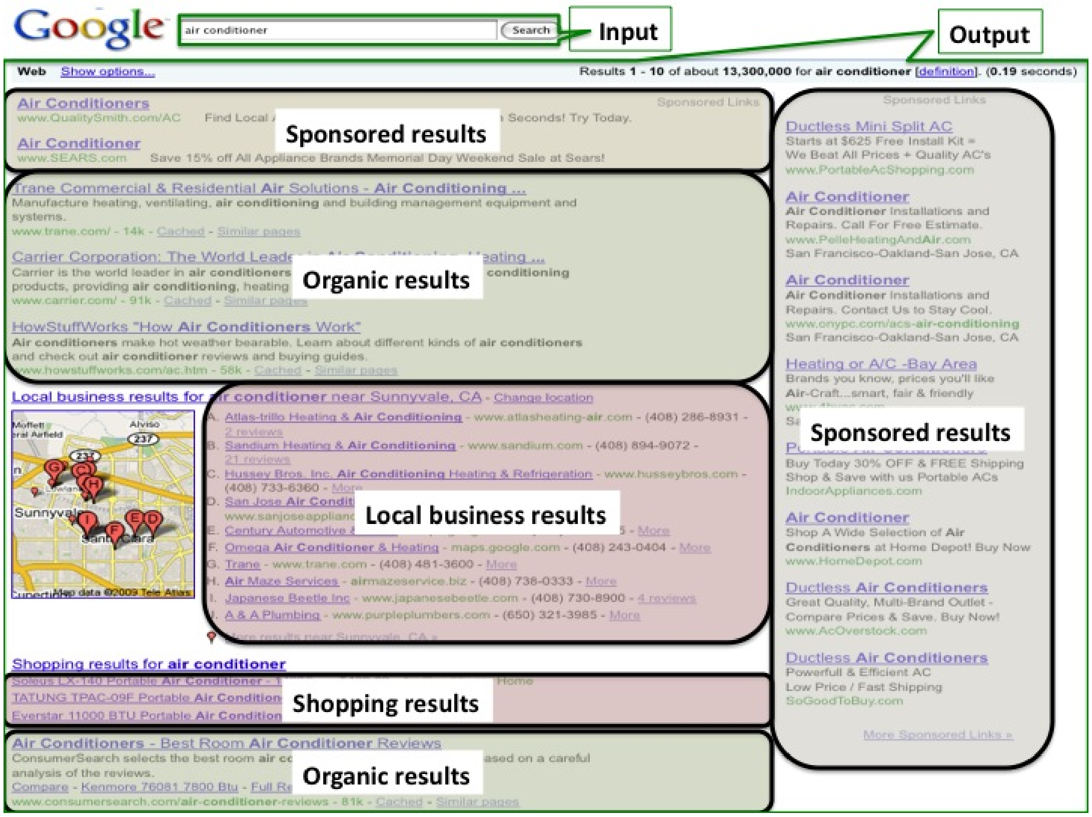

Figure 1: The structure of a search engine results page with input query "air conditioner" as input and search results as output. Search results usually contain organic and sponsored search results. However, note the sections "local business results" and "shopping results" among the organic results.

rialization of the processing tasks may still be acceptable. However, to completely hide the additional latency, caching comes to the rescue because if the analysis results are also cached with the SERP, then the validity of the SERP cache implies the validity of its metadata too, including the output for bidding.

We finally remark that more research and experience are needed to optimize the implementation of output bidding. For example, if cached SERP data can still be useful even if some of the cached entries are invalid. In this case, future research can explore how to remove the corresponding ads and update the auction results accordingly with possibly similar figures of merit.

6. BENEFITS

At a minimum, the first variation of output bidding allows content providers to take advantage of their investment for improving the quality of their organic content. In other words, if their organic content is relevant enough to be viewed and clicked by users through SERPs, this variation opens up a way for these providers to utilize this success. Note that this use is in addition to what these providers already do with the referrals from search engines for their organic and ad content.

This variation may also benefit the search engine by improving the search results freshness. For example, say that a user searches to buy an air conditioner and the search engine returns in the SERP an organic result from an advertiser's web site. The advertiser could advertise next to the search result the new discount on its air conditioners because there was no discount at the crawl time.

Similarly, at a minimum, the second variation of output bidding can be used to enhance input bidding significantly without any changes on the current marketplace parameters for input bidding. It allows not only to expand the bidding keywords but it also allows to control the targeting of ads to maximize the benefit from their impressions. Targeting is expected to help as it is probably true that advertisers and content publishers know. or at least can focus on, their own advertising contexts better than a search engine, which needs to deal with balancing the interests of millions of customers.

We believe with all its variations, output bidding allows more uses for the richness of the output context, more control on ad targeting, and some relief from the effort that goes into figuring out what keywords to bid on. The last point is especially true for advertisers advertising globally in different markets and languages. We finally note that in § 8, we provide some experimental evidence for these benefits.

7. LIMITATIONS

We can see two main limitations of output bidding. One main limitation may result from the use of organic content for bidding. A worst case scenario is that a search engine knowing the bids on an output item, say, a site, may modify its algorithms to show this site more frequently to earn more revenue. A similar scenario also exists today if a search engine uses excessive paid inclusion to show ads in the organic results section or it shows too many ads in the sponsored results section. Fortunately, for all these scenarios, the checks and balances side also already exists: the whole page relevance and user trust. Since too much commercial content reduces relevance or if users realize their explicit or disguised use, the search engine may experience significant losses in their market share. Because of these two sides balancing one another, we think a search engine is unlikely to abuse this power.

The other main limitation involves the implementation challenges, that we discussed in § 5.

We also note that it may take some time for all parties involved, namely, advertisers, users, search engines, and publishers, to get used to the output bidding idea and optimize its use. Since this learning curve is expected of every new idea, we do not consider it as a limitation of output bidding.

8. EXPERIMENTS AND RESULTS

A full design, implementation, and deployment of output bidding is a sizable undertaking, and requires the cooperation of all parties involved in search advertising. What did instead is to show the potential of output bidding by computing some simple lower bounds from offline experiments on real search engine data.

For our first experiment, we verified a hypothesis suggested by Preston McAfee of CalTech and Yahoo Labs: an ad appearing as a sponsored result in a SERP gets higher click-through rate (CTR) if it is somehow correlated with an organic result appearing in the same SERP. We used three days of Yahoo! search logs, and for each query produced the resulting SERP (including the ads displayed). We found that CTR was an order of magnitude higher when the site name appearing in an ad also appeared in the organic results, as compared to cases when site names did not match. For non-navigational queries, CTR was three times higher. This result suggests that output bidding could be useful for advertisers. That is, they could bid on keywords as usual, but ask that their ads only be displayed when the sites mentioned in their ad also appear in the organic results. The resulting CTR for the displayed ads would be significantly higher (and the advertiser would not have to pay for ads with lower expected CTR).

For our second experiment, we simulated the potential for getting new advertisements through keywords extracted from output, i.e., the second variation in § 2. As our base set, we took a sample of 100 queries for which Yahoo! could show no ads or very few ads. (For these queries there were very few or no matching ads.) Figure 2 shows, in the "before plot," the distribution of ads using only the original keywords. As one can see, the vast majority of the queries yielded zero, one or two ads, and no queries yielded more than 5 ads.

The "after" curve in Figure 2 shows the distribution of matching ads when we enhance the input keywords with 10 keywords obtained from the output SERP. The selected extra keywords have the highest TF-IDF scores [5] in the output. We generated the extra ads by submitting these extra

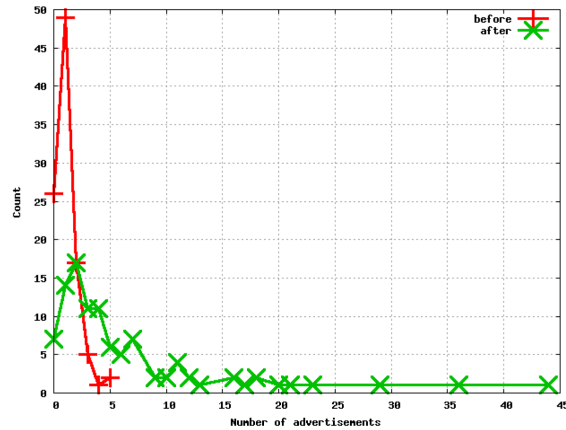

Figure 2: Increase in the number of advertisements due to output bidding. The before plot shows the histogram of the number of ads for 100 user queries. The after plot shows the the same histogram with additional queries created out of the titles and abstracts appearing in the SERPs of the original user queries.

keywords to the Yahoo! search engine. One can see in Figure 2 that the number of generated ads is significantly larger. Overall, there was an increase of more than five times in the total number of ads. Our results suggest that the output SERP contains valuable information that can be exploited by advertisers. In this case we simply extracted keywords form the output, but in general, an advertiser could extract this useful information in other ways and use it to formulate his bidding.

Although more experiments, e.g., using simulations as in [13], are obviously needed to evaluate the full potential of output bidding, we believe these simple experiments with real data point to very encouraging results in the use of output bidding for enhancing search engine advertising. We think that using all the benefits of output bidding to their fullest potential can only produce far more benefit than what we have observed.

9. RELATED WORK

For search advertising, a short review is given in [11], a longer one in [16], and one with more algorithmic focus in [12]. However, for some of the original references, e.g., the introduction of keyword bidding in [9], that are missing from these reviews, Wikipedia is a good source. Also see [18] for research issues regarding keyword advertising.

In [12], the authors discuss sponsored search from the view points of three parties: advertisers, auctioneers (who are search engines), and search users. With its stronger focus on content in SERP, output bidding also factors in content owners or publishers as another first-class party to sponsored search. Future research can show how to formalize this new relationship with and benefit to content publishers.

Output bidding in its general form (with many variations including those in § 2) is originally proposed in [7] (based on the basic idea conceived in early 2006). In [8], the authors develop an extension of output bidding and propose an auction model for bidding for bundles of search results from output. Since the problem of handling arbitrary bundles has high complexity [6], the authors reduce the problem complexity by decomposing bundles from advertisers into smaller bundles that can contain at most one organic result and at most one sponsored result. In [15], the authors propose algorithms for computing optimal bundles for input bidding. Their proposal can also be useful for output bidding. They also cite references on bundling. We expect that more general auction models for bundles are likely to use techniques from combinatorial auctions, e.g., see [6] for an excellent overview of the area.

A variation of output bidding for associating ads with site names in SERP is independently proposed in [19]. The authors state that the site owner is not required to participate in the auction and that the ad matching to a site in SERP may also be independent of the input query that generated the SERP. Similar proposals are also discussed in [7].

Apart from these references, we are unaware of any other work on output bidding. However, there are some related work on exploring the relationship between organic and sponsored results and on using parts of output for input bidding, which are presented next.

In [23][1], an interplay between organic and sponsored results is considered, given the asymmetric competition between strong and weak advertisers and sites. More related work exploring this interplay can be found in [17, 14].

In [23], the authors set up a game-theoretic model to answer whether or not advertisers placed at the top organic result positions should bid actively for sponsored positions, whether or not organic listing benefits or harms search engine revenue, social welfare, and overall sales diversity (e.g., to many small advertisers in the long tail), and whether or not organic ranking can be improved as a result of this interplay. This study finds that organic listing can hurt search engine revenue but help increase social welfare and sales diversity. To optimize for all three metrics together, this study also proposes a new mixed organic listing scheme where the ranks of less popular sites, or "the weak", are boosted to more prominent positions to give them a higher chance of visibility to users.

Allowing bids on more parameters for input bidding are discussed in [1, 20] where the authors of [1] allow advertisers to specify bids and a preference for positions in the list of ads and the author of [20] allows advertisers to impact the number of ads shown. In both references, the authors show how to generalize the current auction models to handle these new constraints. We consider these references as related to output bidding because the new parameters for bids refer to parameters of sponsored search section, hence, those of output.

In [4], the authors study query expansion using web search results to find relevant ads to a query. They submit a query to a search engine and then use the top-scoring web pages to gather additional knowledge about the query. They then use this knowledge to create an augmented query, which is evaluated against the ad corpus to retrieve relevant ads for the original Web query. This approach takes advantage of the search results and is related to the second variation of output bidding. However, the authors view the use of the search results as a means to eliminate the need for advertiser bidding rather than a opportunity to augment the bidding language and give more control to the advertiser over the ad targeting.

[1]We thank Prof. Whinston for emailing us a copy of this paper on 05/29/2009.

10. CONCLUSIONS AND FUTURE WORK

Output bidding uses the output of a search engine either to enhance input bidding, which is the current search engine advertising model, or to create a new form of search engine advertising model. Like input bidding, output bidding also preserves the organicity of search results from advertiser influence. Two experiments performed using real data have shown that the potential gain from output bidding can be significant.

Future work includes work on interesting questions raised by output bidding. Most of these questions have been presented in the "issues" sections.

Acknowledgments

We greatly appreciate the help and encouragement we have received from many of our colleagues, whose names are listed below in alphabetical order: Arnab Bhattacharjee, Chi-Chao Chang, Ozgur Cetin, Marc Davis, Joseph Deck, Kamran Gholamy, Rica Gonen, Frank Filippini, Dz-Mou Jung, Prabhakar Krishnamurthy, Tina Krueger, Sanjay Kshetramade, Jean-Marc Langlois, Tuoc Luong, Preston McAfee, Mark Morrissey, Nagesh Pobbathi, Kartik Ramakrishnan, Ben Shahshahani, Kerem Tomak, Andrew Tomkins, Caroline Tsay, Emre Velipasaoglu, and Sharad Verma.

11. REFERENCES

[1] G. Aggarwal, J. Feldman, and S. Muthukrishnan. Bidding to the top: VCG and equilibria of position-based auctions. In *Proc. Workshop on Approximation and Online Algorithms (WAOA)*, 2006.

[2] G. Aggarwal, A. Goel, and R. Motwani. Truthful auctions for pricing search keywords. In *Proc. Int. Conf. on Electronic Commerce (EC)*. ACM, 2006.

[3] G. Aggarwal, S. Muthukrishnan, D. Pál, and M. Pál. General auction mechanism for search advertising. In *Proc. Int. World Wide Web Conf. (WWW)*, pages 241–50, Apr 2009.

[4] A. Z. Broder, P. Ciccolo, M. Fontoura, E. Gabrilovich, V. Josifovski, and L. Riedel. Search advertising using web relevance feedback. In *Proc. Conf. on Information and Knowledge Management (CIKM)*, pages 1013–22. ACM, 2008.

[5] S. Chakrabarti. *Mining the Web*. Morgan Kaufmann, 2003.

[6] P. Cramton, Y. Shoham, and R. Steinberg, editors. *Combinatorial Auctions*. MIT Press, 2006.

[7] A. Dasdan. System for displaying advertisements associated with search results. US Patent Appl., Apr 2007.

[8] A. Dasdan and R. Gonen. System and method for offering an auction bundle in an online advertising auction. US Patent Appl., Mar 2008.

[9] D. J. Davis, M. Derer, J. Garcia, L. Greco, T. E. Kurt, T. Kwong, J. C. Lee, K. L. Lee, P. Pfarner, and S. Skovran. System and method for influencing a position on a search result list generated by a computer network search engine. US Patent 6269361, Jul 2001.

[10] B. Edelman, M. Ostrovsky, and M. Schwarz. Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. *American Economic Review*, 97(1):242–59, Mar 2007.

[11] D. C. Fain and J. O. Pedersen. Sponsored search: A brief history. In *Proc. Workshop on Sponsored Search Auctions*, Jun 2006.

[12] J. Feldman and S. Muthukrishnan. Algorithmic methods for sponsored search advertising. In *Proc. SIGMETRICS*, pages 91–124, 2008.

[13] J. Feng, H. K. Bhargava, and D. M. Pennock. Implementing sponsored search in web search engines: Computational evaluation of alternative mechanisms. *INFORMS J. on Computing*, 19(1):137–48, 2007.

[14] A. Ghose and S. Yang. Organic and paid search advertising: Complements, substitues or neither? In *Proc. Conf. The Economics of the Software and Internet Industries*. Idei.fr, Jan 2009.

[15] A. Ghosh, H. Nazerzadeh, and M. Sundararajan. Computing optimal bundles for sponsored search. In *Proc. Workshop on Internet and Network Economics (WINE)*, pages 576–83, 2006.

[16] B. J. Jansen and T. Mullen. Sponsored search: An overview of the concept, history, and technology. *Int. J. Electronic Business*, 6(2):114–31, 2008.

[17] Z. Katona and M. Sarvary. The race for sponsored links: A model of competition for paid placement on a search engine. *To appear in Marketing Sci.*, 2009.

[18] D. Liu, J. Chen, and A. B. Whinston. *Business Computing*, volume 3 of *Handbooks in Information Systems*, chapter Current Issues in Keyword Auctions, pages 67–97. Emerald Group Publ. Ltd., 2009.

[19] E. Manavoglu, A. Popescul, B. Dom, and C. Brunk. System for targeting data to sites referenced on a page. US Patent Appl., Jun 2008.

[20] S. Muthukrishnan. Bidding on configurations in internet ad auctions. In *To Appear in Proc. Int. Conf. on Comput. and Combinatorics (COCOON)*, 2009.

[21] H. R. Varian. Position auctions. *Int. J. Industrial Organization*, 25(6):1161–78, 2007.

[22] H. R. Varian. Online ad auctions. *To appear in AER Papers and Proceedings*, May 2009.

[23] L. Xu, J. Chen, and A. B. Whinston. Too organic for organic listing? interplay between organic and sponsored listing in search advertising. Available at SSRN: http://ssrn.com/abstract=1409450, May 2009.