VALUE OF SHARING DATA

PATRICK HUMMEL*

FEBRUARY 12, 2018

Abstract. This paper analyzes whether advertisers would be better off using data that would enable them to target users more accurately if the only way they could use this data is by sharing the data with all advertisers. I present a wide range of theoretical results that illustrate general circumstances under which an advertiser can be assured that sharing data will make the advertiser better off. I then empirically analyze how sharing data would affect Google's top 1000 advertisers on mobile apps, and find that over 98% of these advertisers would always be better off sharing data, regardless of how the data affects competitors' bids.

Keywords: Sharing Data; Targeting; Online Auctions; Advertising; Advertiser Welfare

1. Introduction

In some online advertising systems, such as those for Universal App Campaigns at Google, the only way an advertiser would be able to make use of its targeting data in bidding in an auction is by sharing its data with the auctioneer. In these auctions, advertisers are unable to submit separate bids that depend on data that only they have access to, and must instead share the data with the online ad system so that the ad system can optimize the advertiser's bidding on its behalf.

An effect of this is that when an advertiser shares its data with the online ad system, the online ad system will then be able to use this data to help other advertisers bid more accurately in auctions as well. Since sharing data will then enable an advertiser's competitors to bid more accurately, an advertiser might wonder whether the advertiser would be better off sharing its data with the online ad system. Will an advertiser be better off sharing its

*Google Inc., 1600 Amphitheatre Parkway, Mountain View, CA 94043, USA.

data with an online ad system when one considers that this could also enable the advertiser's competitors to bid more accurately?

This paper presents both theoretical and empirical analysis that suggests that this will be the case. I begin by theoretically analyzing the circumstances under which an advertiser will be better off sharing data if there is exactly one advertising opportunity available in each auction. When there is only one advertising opportunity available, I find that there are several circumstances under which there is a theoretical guarantee that an advertiser will be better off if the advertiser shares its data. An advertiser is guaranteed to be better off sharing its data if (i) the advertiser was initially losing the auction, (ii) the advertiser's strongest competitor was initially bidding significantly more than the next-strongest competitor, or (iii) the competing advertisers would adjust their bids in the same way as a result of the advertiser sharing its data. Furthermore, in cases where an advertiser was winning an auction and facing a strong competitor, sharing data will frequently enable this advertiser to identify advertising opportunities that it would be better off not winning.

I then consider a setting in which there are multiple positions for advertising opportunities available in each auction. As in the case in which only a single advertising opportunity is available, I find that if an advertiser was initially losing the auction, then sharing data can only be beneficial for the advertiser. I further find that if sharing data would induce advertisers who are initially ranked higher than the advertiser who shares its data to change their bids, that this can also only help the advertiser who shares its data. Finally, if sharing data would induce the lower-ranked advertisers to adjust their bids by a common multiplier or the highest losing advertiser is initially bidding quite a bit more than the next-highest losing advertiser, then sharing data can again only help the advertiser who shares its data.

Given the wide range of circumstances under which an advertiser will be better off sharing its data, stringent conditions must be met in order for sharing data to make an advertiser worse off. In particular, in order for sharing data to make an advertiser worse off in a particular auction, it is typically necessary that (i) the advertiser would have its ad shown in the absence of sharing its data, (ii) the lower-ranked advertisers were initially bidding considerably less than the advertiser who shared its data, (iii) the top losing advertisers were initially making similar bids to one another, and (iv) the losing advertisers would adjust their bids in a relatively uncorrelated manner as a result of being given this data. In the rare cases that these conditions are all simultaneously met, it is theoretically possible that sharing data might make an advertiser worse off, but otherwise sharing data should make an advertiser better off.

I then analyze whether the conditions that need to be satisfied in order for an advertiser to be better off sharing its data with an online ad system seem to be met empirically. I analyze data from Google auctions on both mobile apps (which display one ad at a time) and Play Search (which sometimes display multiple ads along with organic results). In both of these cases, I analyze how each of the top 1000 advertisers' payoffs would be affected if the advertiser used its data as a strategic input to its campaign when the data could (i) help the advertiser bid more accurately, (ii) help competing advertisers bid more accurately on dimensions in which their values are correlated, and (iii) help competing advertisers bid more accurately on dimensions in which their values are independent of one another.

In particular, for each of the top 1000 app advertisers, I analyze how sharing data would affect advertisers' payoffs under hundreds of combinations of ways that the possibilities (i)-(iii) might affect advertisers' bids. My analysis of each of these possibilities suggests that the vast majority of app advertisers would be better off making use of this data. In particular, I find that over 98% of the top 1000 advertisers on mobile apps would always see improved performance by using their data as strategic inputs into their campaigns regardless of how the data affected other advertisers' bids. On top of this, roughly 65% of the top 1000 advertisers on Play Search would always be better off using this data, and roughly 90% of these advertisers would never suffer a payoff loss greater than 0.1%, while still typically seeing improved performance as a result of sharing data. The combination of these theoretical and empirical results indicates that sharing data is likely to make an advertiser better off, even if this data can be also used to help the advertiser's competitors bid more accurately.

1.1. **Related Work.** While there has been a wide range of work related to the question of how targeting data affects the market for online advertising, little of this work has addressed the question of whether an advertiser would be better off using targeting data in an auction if the only way that advertiser can use the data is by sharing the data with all advertisers. Perhaps the most closely related work to this paper is Bhawalkar *et al.* (2014). Bhawalkar *et al.* (2014) primarily address the question of how the value of targeting data to an advertiser in a decision-theoretic framework depends on various quantities such as the advertiser's utility difference between different realizations of the targeting data, the accuracy of the data, the distribution of competing bids, and the advertiser's budget. This paper also includes some results on the value of data when multiple advertisers are given access to the data, which I build on in this paper.

Other papers analyze how targeting data affects advertiser welfare in less closely related settings. Bergemann and Bonatti (2011) find in a model with many advertising markets and advertisers that an increase in targeting ability results in an increase in the social value of advertising, while decreasing prices for advertising, thereby making advertisers better off. De Cornière and De Nijs (2016) analyze how allowing advertisers to target affects advertiser welfare when this targeting may ultimately affect downstream price competition. And Levin and Milgrom (2010) offer anecdotal evidence that advertiser welfare is likely to be enhanced by improved targeting.

Ghosh *et al.* (2015) address a question for publishers that is related to the question this paper addresses for advertisers. In particular, Ghosh *et al.* (2015) wonder whether publishers would have an incentive to share information about the websites a user has visited when sharing this information may enable advertisers to better target this user on other publishers and thereby possibly hurt the publisher who shares the data. The authors find that when advertisers are homogeneous, then publishers either all benefit or all suffer from sharing information, but when advertisers are heterogeneous, this cookie-matching can help some publishers at the expense of the others.

Other papers on targeting data focus on how improved targeting data affects revenue. Abraham *et al.* (2016) study the effect of *ex ante* information asymmetries on revenue. Board (2009), Fu *et al.* (2012), Hummel and McAfee (2016), Milgrom and Weber (1982), and Palfrey (1983) offer results that shed light on the question of how enhanced information disclosure or improved targeting would affect a seller's revenue. And Bergemann and Välimäki (2006), Bergemann and Wambach (2015), Emek *et al.* (2012), Eső and Szentes (2007), Ganuza (2004), Ganuza and Penalva (2010), Ghosh *et al.* (2007) address questions related to revenue optimization strategies in the context of information disclosure.

While this literature addresses a wide range of questions related to targeting and information disclosure, this work largely fails to address the question of whether an advertiser would be better off using targeting data if the only way that advertiser could use the data was by sharing the data with other advertisers. This paper fills this gap in the literature.

2. The Model

I consider a setting in which there is a set of n advertisers who are competing in an auction. Each advertiser i has some initial estimated value for a click, which I denote by v_i, and I assume without loss of generality that $v_1 \geq v_2 \geq \ldots \geq v_n$.

If an advertiser shares its data, then each advertiser i will learn that its true value for the advertising opportunity is $\tilde{v}_i = m_i v_i$, where m_i is a non-negative multiplier that is a random draw from some distribution with mean 1. Throughout I also allow for the possibility that the values of m_i may be correlated for the different advertisers.

I consider two possible auction formats in which the advertisers might compete. I first consider a standard second-price auction, in which there is one advertising opportunity available for sale, and the bidder who submits the highest bid wins the advertising opportunity and pays a cost-per-click equal to the second-highest bid.

The second possibility I consider is a position auction in which there are s positions for advertising opportunities available for sale. In such an auction, the advertiser who submits the j^{th}-highest bid is placed in position j and obtains a total of x_j clicks, where x_j is decreasing in j for all $j \leq s$, and $x_j = 0$ for all $j > s$. Prices are then set using the Vickrey-Clarke-Groves (VCG) mechanism, so an advertiser who is placed in position k will pay a total cost of $\sum_{j=k}^{s}(x_j - x_{j+1})b_{j+1}$ for its x_k clicks, where b_{j+1} denotes the $j+1^{th}$-highest bid in the auction.[1] This translates into a cost-per-click of $\frac{1}{x_k}\sum_{j=k}^{s}(x_j - x_{j+1})b_{j+1}$.

Regardless of whether there are multiple positions, if an advertiser with a value v per click obtains a total of x clicks and pays a cost-per-click of c, then this advertiser obtains a payoff of $x(v - c)$. The mechanisms described have the property that each advertiser will have an incentive to make a bid in the auction that equals the advertiser's expected value per click. I seek to compare the advertiser's profit from advertising when the advertiser does not share its data with the advertiser's profit when the advertiser shares its data.

3. Auctions with One Position

I first consider auctions for a single advertising opportunity. Under such auctions, there are two basic situations where sharing data might affect an advertiser. The first situation is a setting in which the advertiser would have lost the auction in the absence of sharing its data. And the second situation is a setting in which the advertiser would have won the auction in the absence of sharing its data. I consider both of the settings in turn.

3.1. **Losing Advertisers.** First note that if an advertiser would have lost the auction if the advertiser did not share its data, then sharing data can only make the advertiser better off. The advertiser will not obtain any net profit if the advertiser does not share its data, but if the advertiser shares its data, there is a chance that the advertiser will reveal information that enables the advertiser to win the auction and achieve a positive profit. Thus sharing data can only make the advertiser better off in this case.

3.2. **Winning Advertisers.** The more interesting case is the case where the advertiser would have won the auction in the absence of sharing its data, as it is theoretically possible

[1] Edeleman *et al.* (2007) and Varian (2007) have shown that under a generalized second-price (GSP) auction, there is an equilibrium which results in the same auction outcomes that would result under the dominant strategy equilibrium of the VCG mechanism. Thus the results I derive for VCG immediately extend to a GSP auction.

that sharing data will make an advertiser who would have won the auction without sharing its data worse off. Nonetheless, even in this setting, there are numerous important ways that sharing data could make the advertiser better off.

First note that sharing data could enable the advertiser to identify some advertising opportunities that it is currently winning that are not valuable to the advertiser, and avoid winning these auctions. This could happen, for instance, in the special case where $m_1 < 1$ and $m_j = 1$ for all $j > 1$, in which case the advertiser with the highest initial estimate of its value learns that the advertising opportunity is not as valuable to this advertiser as the advertiser previously thought, while the other advertisers' estimates of their values are unaffected by the data. In this case, sharing data could enable the advertiser to spend less money on advertising opportunities that are not inherently valuable to the advertiser, thereby making the advertiser better off.

But even if sharing data does not help the winning advertiser bid more accurately, this advertiser might still benefit from sharing its data. It turns out that even if the data is only useful for helping competitors bid more accurately that this could in turn make the winning advertiser better off. To understand why this is the case, I first consider what would happen if the winning advertiser only faces one competing bidder in the auction.

3.2.1. *One Competitor.* When the winning advertiser only faces one competitor, there are two possibilities that might occur when the winning advertiser shares its data. The first possibility is that sharing data will never change which advertiser wins the auction. If this is the case, then sharing data will have no net effect on auction outcomes because the price the winning advertiser will pay will remain the same in expectation. Regardless of whether the winning advertiser shares its data, the competing advertiser has an incentive to make a bid equal to its expected value for an advertising opportunity given the information that it has. This average value of advertising opportunities does not change as a result of sharing data, so the average competing bid, and thus the average price that the winning advertiser pays, will be unaffected by sharing data. Thus sharing data will have no net effect on auction outcomes in this case.

The second possibility is that sharing data will sometimes change which advertiser wins the auction because it will sometimes induce the competing advertiser to raise its bid above the original winning advertiser's bid. In this case, sharing data will make the advertiser strictly better off. The reason for this is that the competitor will continue to make the same average bid as a result of the data sharing, so if the competitor raises its bid above the original winner's bid in some auctions, it must also lower its bid in other auctions. The benefits to the original winner from this bid lowering will always be at least as large as any losses resulting from bid raising, so sharing data will be beneficial. This is illustrated in the following example:

Example 1. *Suppose* $v_1 = 1$ *and* $v_2 = 0.8$*, so that initially advertiser* 1 *is making a profit of* $1 - 0.8 = 0.2$ *in each auction. Now suppose that after sharing its data with advertiser* 2 *that advertiser* 2 *learns that* $\tilde{v}_2 = 1.1$ *with probability* $\frac{1}{2}$ *and* $\tilde{v}_2 = 0.5$ *with probability* $\frac{1}{2}$*. Then advertiser* 1*'s expected profit from an auction changes to* $\frac{1}{2}(1 - 0.5) = 0.25 > 0.2$*, meaning advertiser* 1 *becomes better off as a result of sharing its data.*

The insight from this example is more general than this. The key insight is that once a competitor raises its bid above the original winner's bid, then further bid raising has no adverse effect on the original winner because the original winner is now losing the auction. However, the original winner always benefits from additional bid lowering because this reduces the price the winner must pay in the auction. This is illustrated in the following theorem, taken from Bhawalkar *et al.* (2014):

Theorem 1. [Bhawalkar *et al.* (2014)] *When there are two advertisers, each advertiser prefers to share any targeting data it has with the other advertiser.*

Theorem 1 thus provides a theoretical guarantee that sharing data will make an advertiser better off if the advertiser only has one competitor.

3.2.2. *Multiple Competitors.* The theoretical guarantee in Theorem 1 does not fully extend to a setting where an advertiser faces multiple competitors. Instead it is now theoretically possible to construct contrived examples under which the original winner of the auction might become worse off as a result of sharing its data:

Example 2. *Suppose there are* $n = 3$ *bidders,* $v_1 = 1$*, and* $v_2 = v_3 = 0.5$*. Also suppose that after sharing its data with the competing bidders, advertisers* $j \in \{2, 3\}$ *learn that* $\tilde{v}_j = 0.55$ *with probability* $\frac{1}{2}$ *and* $\tilde{v}_j = 0.45$ *with probability* $\frac{1}{2}$*, where the values of* $\tilde{v}_2$ *and* $\tilde{v}_3$ *are independent of one another. Then advertiser* 1 *will obtain a greater profit by not sharing its data than by sharing data.*

Proof. Since $\tilde{v}_j$ assumes the value 0.55 with probability $\frac{1}{2}$ and $\tilde{v}_j = 0.45$ with probability $\frac{1}{2}$ for each $j \in \{2, 3\}$ and the values of $\tilde{v}_2$ and $\tilde{v}_3$ are independent of one another, then with probability $\frac{3}{4}$, at least one of the values of $\tilde{v}_j$ will equal 0.55, and with probability $\frac{1}{4}$, both of these values will equal 0.45. Thus if the advertiser shares its data, then with probability $\frac{3}{4}$, the highest competing bid will be 0.55, and with probability $\frac{1}{4}$, the highest competing bid will be 0.45, meaning the expected value of the second-highest bid will be 0.525. However, if the advertiser does not share its data, then the highest competing bid will be 0.5. Thus if the advertiser shares its data, then the advertiser will have to pay a higher price in expectation, so sharing its data will make the advertiser worse off. □

Example 2 illustrates that it is theoretically possible that sharing data will make an advertiser worse off if the advertiser faces multiple competitors. However, the construction of the above example relies on several assumptions that are all critical in order to obtain this result and may be unlikely to hold in practice.

First, it is necessary that the two highest competing advertisers would both make similar bids to one another if the advertiser did not share its data. If the third-highest bid is initially much lower than the second-highest bid, then the third-highest bidder is unlikely to move up to the top two as a result of sharing its data, and the situation is not much different from what would happen if the winning advertiser only faces one competitor. Since we know the winning advertiser is better off sharing its data if the winning advertiser only faces one competitor, it follows that this advertiser will also be better off sharing its data if the third-highest bid is initially much lower than the second-highest bid. Formally, we have the following result (this theorem and all subsequent theorems are proven in the appendix):

Theorem 2. *Suppose the values of m_j are drawn from distributions such that $\tilde{v}_2 \geq \tilde{v}_j$ for all $j > 2$ with probability 1. Then advertiser 1 will obtain at least as large a payoff by sharing its data than by not sharing its data.*

In addition to the above, it is also necessary for the two highest competing advertisers to adjust their bids differently from one another. If instead of this, any time one competing advertiser adjusted its bid by a factor of m the other competing advertiser also adjusted its bid by a factor of m, then the outcome of the auction will be the same for the original winner as it would have been if there were only one competing advertiser in the auction. This in turn implies that the winning advertiser would be better off sharing its data if the competing advertisers' bid adjustments were highly correlated with one another. In particular, we have the following result:

Corollary 1. *Suppose the values of m_j are drawn from a distribution such that the values of m_j for $j \geq 2$ are all equal with probability 1. Then advertiser 1 will obtain at least as large a payoff by sharing its data than by not sharing its data.*

These results thus indicate that there are far more circumstances under which sharing data will be beneficial to an advertiser than there are ways that this can make the advertiser worse off. Sharing data will help the advertiser in auctions the advertiser would have lost in the absence of sharing its data, it will help the advertiser avoid winning auctions that the advertiser would have won otherwise but are not valuable to the advertiser, it will make the advertiser better off when the advertiser only faces one strong competitor, and it will make the advertiser better off when the advertisers' competitors would adjust their bids in similar ways as a result of being given access to the data.

Furthermore, in cases where the highest competing advertiser is only bidding slightly less than the original winner, it is very likely that sharing data will enable the advertiser who was originally winning the auction to selectively win only the advertising opportunities that are most valuable to this advertiser, and thereby make this advertiser better off. Combining all these insights suggests that sharing data will typically only make an advertiser worse off if (i) the advertiser would have won the auction in the absence of sharing data, (ii) all of the competing advertisers are initially bidding considerably less than the winning advertiser, (iii) the top two competing advertisers are making similar bids, and (iv) the competing advertisers would adjust their bids in a manner uncorrelated with one another as a result of being given this data. Thus stringent conditions are typically needed in order for sharing data to make an advertiser worse off.

4. Position Auctions

I now extend my analysis in the previous section to auctions for multiple positions. As in the case where there is one position in Section 3.1, if an advertiser would not have its ad shown if the advertiser did not share its data, then sharing data can only make the advertiser better off. If the advertiser does not share its data, then the advertiser will necessarily obtain zero profit from advertising. However, if the advertiser does share its data, then it is possible that the advertiser may be able to profitably win an advertising opportunity that the advertiser would not have won in the absence of sharing its data. Thus I focus on the case in which an advertiser would have had its ad shown in the absence of sharing data in this section.

Even if the advertiser would have its ad shown if the advertiser did not share its data, it is still possible that sharing data could make the advertiser better off. As in the case with one position, there is the possibility that sharing data would enable the advertiser to bid more accurately, which could in turn increase the advertiser's payoff by either enabling the advertiser to increase the number of clicks that it purchases when its value is higher than expected or enabling the advertiser to decrease the number of clicks it purchases when its value is lower than expected. This corresponds to the case in which $m_i \neq 1$ for the advertiser i that shares its data but $m_j = 1$ for all $j \neq i$.

On top of this, sharing data can benefit an advertiser even if the data can only be used to help competitors bid more accurately. Suppose, for instance, the data is primarily used to help the higher-ranked advertisers bid more accurately. In particular, suppose that if advertiser i shares its data, then each advertiser $j > i$ learns that $m_j = 1$ and thus does not adjust its bid as a result of advertiser i sharing data. In this case, sharing data is again beneficial for advertiser i:

Theorem 3. *Suppose sharing data does not affect the bids of the lower-ranked advertisers in the sense that if advertiser i shares its data, then each advertiser $j > i$ learns that $m_j = 1$. Then sharing data will make advertiser i at least as well off.*

The intuition for Theorem 3 is that if as a result of sharing data, the higher-ranked advertisers adjust their bids in such a manner that these advertisers would continue to be ranked higher than advertiser i, then sharing data does not affect advertiser i's payoff. However, if one or more of these advertisers lowers its bid below advertiser i's bid, then advertiser i will be able to move up a position and obtain a greater payoff from advertising. Thus sharing data can only be beneficial in this case.

Sharing data can also benefit an advertiser if the data is primarily used to help the lower-ranked advertisers bid more accurately. In particular, suppose that the only way that sharing data affects bids is by inducing the advertisers $j > i$ to adjust their bids by some common multiplier m. In this setting I am able to prove the following result:

Theorem 4. *Suppose that if advertiser i shares its data, then each advertiser $j > i$ will adjust its bid by some common multiplier m that is a random draw from some distribution with mean 1. Then sharing data will make advertiser i at least as well off.*

The reason Theorem 4 holds is that if sharing data never induces any of the lower-ranked advertisers to raise their bids above advertiser i's original bid, then advertiser i will obtain the same expected payoff as before because the fact that these advertisers are adjusting their bids by a common component with mean 1 means that advertiser i's expected price for these clicks will not be affected. However, if sharing data sometimes induces the lower-ranked advertisers to raise their bids above advertiser i's original bid, then advertiser i will necessarily benefit as a result of sharing data. The reason for this is that if an advertiser

$j > i$ raises its bid above advertiser i's bid in some auctions, then advertiser j will also lower its bid by a correspondingly large amount in other auctions. However, while additional bid raising above advertiser i's bid has no effect on advertiser i's payoff, additional bid lowering always reduces advertiser i's costs and thus benefits advertiser i. Because of this, the net effect of the lower-ranked advertisers' bid changes can only benefit advertiser i in the setting considered in Theorem 4.

It is also worth noting that the setting in Theorem 4 is not the only setting under which sharing data can benefit an advertiser, even if the lower-ranked advertisers would change their bids in response. Even if the lower-ranked advertisers adjust their bids differently from one another, we sometimes have a guarantee that the advertiser will benefit from sharing data if there is a sufficiently small amount of competition in the auction. In particular, I prove the following result:

Theorem 5. *Suppose $x_j - x_{j+1}$ is increasing in j for all $j \leq s$ and there are $n = s + 1$ bidders in the auction. Then sharing data will make advertiser i at least as well off.*

To understand why Theorem 5 holds, note that we can think of the total cost a bidder in position i must pay as the sum of the costs for the incremental clicks the bidder must pay for each of the positions $j \leq i$, where the incremental clicks a bidder achieves for moving up to position j from position $j+1$ is equal to $x_j - x_{j+1}$, and the price for each of these incremental clicks is set by the $j + 1^{th}$-highest bid. If a competing advertiser adjusts its bid as a result of sharing data, then this advertiser will sometimes raise bids and sometimes lower bids. In cases where the advertiser raises its bid, the advertiser will make the incremental clicks in a higher position more expensive, but if the advertiser lowers its bid enough to move down a position, the advertiser will make the incremental clicks in a lower position cheaper. Since the condition that $x_j - x_{j+1}$ is increasing in j implies that there are more incremental clicks in the lower positions, the cost reductions from the cases in which this bidder lowers its bid will be at least as large as any losses resulting from the case in which the bidder raises its bid.

An important application of the setting in Theorem 5 is the case in which there are no more than $s = 2$ positions available for sale, as is typically the case in advertising auctions on Google Play Search. In that case, the technical condition that $x_j - x_{j+1}$ is increasing in j for all $j \leq s$ is equivalent to $x_2 - x_3 > x_1 - x_2$, which holds if and only if $2x_2 > x_1 \Leftrightarrow x_2 > \frac{x_1}{2}$. Thus if an advertiser could achieve at least half as many clicks in second position as the advertiser can achieve in top position, Theorem 5 would guarantee that an advertiser would always be better off sharing its data if there are exactly $s = 3$ bidders. Since an advertiser will typically be able to achieve at least half as many clicks in second position as the advertiser could obtain in top position, Theorem 5 indicates that we can expect an advertiser to be better off sharing its data in any auctions on Play Search with no more than three bidders.

The previous results indicate that there are a wide range of circumstances under which there is a theoretical guarantee that sharing data will make an advertiser better off. However, if these conditions are not satisfied, it is at least theoretically possible that sharing data will make an advertiser worse off. I illustrate this in the following example:

Example 3. *Suppose there are $n = 4$ bidders, $s = 2$ positions, $v_1 = 4$, $v_2 = 2$, $v_3 = v_4 = 1$, and if advertiser 2 shares its data, then neither advertisers 1 or 2 will adjust their bids, but each of the advertisers 3 and 4 will adjust their bids independently of one another by bidding 1.1 in a random half of the auctions and 0.9 in the remaining auctions. Then advertiser 2 will obtain a greater profit by not sharing its data than by sharing its data.*

Proof. If each of the advertisers 3 and 4 adjusts its bids independently of one another in the manner described above, then with probability $\frac{3}{4}$, the third-highest bid in the auction will be 1.1, and with probability $\frac{1}{4}$, the third-highest bid in the auction will be 0.9, meaning the expected value of the third-highest bid will be 1.05. However, if the advertiser does not share its data, then the third-highest bid will be 1. Thus if the advertiser shares its data, then the advertiser will have to pay a higher price in expectation, so sharing its data will make the advertiser worse off. $\square$

While Example 3 illustrates that an advertiser need not be better off sharing its data, this example relies on a fairly peculiar construction. In order for this example to work, it is necessary that the advertiser who shares its data is facing two losing advertisers who are both initially bidding about the same amount, the advertisers both adjust their bids differently from one another, and neither the advertiser nor the higher-ranked advertisers would adjust their bids in a manner that leads to re-ranking as a result of sharing data. If any of these conditions is not met, then sharing data would typically make the advertiser better off.

To see why these conditions are typically needed, note that if advertiser 4 were initially bidding significantly less than advertiser 3, then regardless of how the lower-ranked advertisers adjusted their bids as a result of sharing data, it is very likely that advertiser 4 would continue to be ranked outside the top three advertisers. Thus the effect that sharing data would have on the advertisers' payoffs in this case would not be much different than the effect of sharing data when there are only three advertisers in the auction, where we know from Theorem 5 that sharing data would typically be beneficial. In particular, we have the following result:

Theorem 6. *Suppose the values of m_j are drawn from distributions such that $\tilde{v}_{s+1} \geq \tilde{v}_j$ for all $j > s+1$ with probability 1. Also suppose that $x_j - x_{j+1}$ is increasing in j for all $j \leq s$. Then sharing data will make each advertiser i at least as well off.*

Theorem 6 thus indicates that it is typically necessary for the two losing advertisers to both initially be bidding similarly to one another in order for sharing data to be harmful. If advertiser $s+1$ is bidding considerably more than advertiser $s+2$, then $\tilde{v}_{s+1} \geq \tilde{v}_j$ is very likely to hold for all $j > s+1$, and Theorem 6 implies that each advertiser will be at least as well off if this advertiser shares its data.

It is also necessary that the losing advertisers adjust their bids differently from one another. If these advertisers adjust their bids in a manner that is perfectly correlated with one another, then we know from Theorem 4 that sharing data will necessarily make advertiser 2 better off. Thus this condition is also necessary.

Finally, it is also typically necessary that neither the advertiser who shares its data nor any of the higher-ranked advertisers would adjust their bids in a manner that leads to re-ranking. If the higher-ranked advertisers adjust their bids as a result of sharing data, then we know from Theorem 3 that if this leads to re-ranking, then this will necessarily make the advertiser who shares its data better off. And the advertiser can certainly only be made better off if, as a result of being able to bid more accurately, the advertiser finds that it now wants to bid in such a way that it is placed in a different position. Thus this last condition is also typically needed.

In conclusion, in order for an advertiser to become worse off as a result of sharing data in a particular auction, it is typically necessary for (i) the advertiser to have its ad shown in the absence of sharing data, (ii) the top two losing bidders in the auction to initially be bidding similarly to one another, (iii) the top losing bidders to adjust their bids in an uncorrelated manner, and (iv) for the other bidders to either be bidding considerably less or considerably more than the bidder who shares its data so that neither the advertiser who shares its data nor the higher-ranked advertisers will want to adjust their bids in such a way that the advertiser who shares its data will be re-ranked. Thus stringent conditions must be satisfied in order for sharing data to make an advertiser worse off.

5. Empirical Analysis

To obtain a better sense of whether the types of conditions that would need to be satisfied in order for sharing data to make an advertiser better off are met empirically, I conduct some empirical analysis of Google's ad auctions. In particular, I seek to measure how payoffs for the top 1000 advertisers on mobile apps and Play Search would be affected if these advertisers shared their data. These settings differ in that only one ad is displayed on mobile apps, whereas two ads are frequently displayed on Play Search.

We know from Section 3.1 that if an advertiser in an auction for one advertising opportunity did not win the auction, then this advertiser can only become better off if it shares its data. And we also know from Theorem 3 that when there are multiple positions, if higher-ranked advertisers change their bids as a result of an advertiser sharing its data, then this can only benefit the advertiser in question. Because of this, the case where it is least likely that sharing data will be beneficial to an advertiser is the case in which there are no advertisers that are higher-ranked, and the advertiser would have been placed in top position even if the advertiser had not shared its data. Thus I focus on this case throughout my analysis.

In my analysis I consider three possible ways that sharing the data could affect an advertiser. First I allow for the possibility that the data could be used to help the advertiser who shared the data bid more accurately in the auction. This possibility can clearly only make the advertiser better off.

The second possibility I consider is the possibility that the data would enable the competing advertisers to refine their bids by a common component that reflects the common value of the user to the competing advertisers. This is also likely to be a common use of the data, but the data is likely to be less useful for this purpose than it is for the purpose described in the previous paragraph because the data will typically be more strongly correlated with the outcomes that the advertiser cares about than it is with the outcomes that the other advertisers care about.

The final use of the data that I consider is that the data could be used to help competing advertisers refine their bids in a manner that is independent of how the other advertisers refine their bids by identifying users that are valuable to one advertiser but not to the other advertisers. While the data may be used in this way, this use of the data should be less common than the use described in the previous paragraph because if a user is more valuable to one competing advertiser, the user will also typically be more valuable to the other competing advertisers.

To model these uses of the data, I assume that the winning advertiser's value would be modified by some multiplier m_w that is a random draw from some cumulative distribution function $F(\cdot)$ with expected value 1. I also assume that the competing advertisers' values will be modified by some common multiplier m that is a random draw from some cumulative distribution function $G(\cdot)$ with expected value 1. Finally, I assume that each competing advertiser j will have its estimated value modified by some additional multiplier m_j that is an independent draw from a cumulative distribution function $H(\cdot)$ with expected value 1. Thus if the winning advertiser shares its data, then each competing advertiser j will learn that its true value for the advertising opportunity is $\tilde{v}_j = mm_jv_j$, where v_j denotes advertiser j's original estimate of its value.

To match empirical evidence on the distributions of advertisers' bids (Lahaie and McAfee 2011; Ostrovksy and Schwarz 2016; Sun *et al.* 2014), I assume that $F(\cdot)$, $G(\cdot)$, and $H(\cdot)$ are all lognormal distributions with mean 1, and I allow the standard deviations in the lognormal distributions σ_F, σ_G, and σ_H to vary in my different simulations. I focus on a setting in which $\sigma_F \geq \sigma_G \geq \sigma_H$ since I expect the winning advertiser's data to do at least as much to affect the winning advertiser's estimate of its value as it does to affect the competitors' estimates of their values, and I also expect the data to do at least as much to affect the common component of competing advertisers' estimates of their values as it does to affect the idiosyncratic components of competing advertisers' estimates of their values.

In particular, for every tuple of values $(\sigma_F, \sigma_G, \sigma_H)$ such that σ_F, σ_G, and σ_H are all integral multiples of 0.05 satisfying $0 \leq \sigma_H \leq \sigma_G \leq \sigma_F \leq 0.5$, I simulated the auction outcomes that would arise if the values of m_w, m, and m_j in each auction were independent and identically distributed draws from the lognormal distribution with mean 1 and standard deviations σ_F, σ_G, and σ_H respectively. After simulating these auction outcomes, I then calculated the difference between the profit the advertiser would obtain with and without sharing data.

5.1. **Mobile Apps.** I begin by presenting the results for advertising opportunities on mobile apps. Since the auctions on mobile apps are standard second-price auctions for a single advertising opportunity, throughout I assume that each advertiser is making a bid that equals the advertiser's true value for an advertising opportunity. For each of the top 1000 app advertisers, I conducted numerical simulations to measure how this advertiser's payoff would be affected in the subset of auctions in which the advertiser won the auction under each of the tuples of values $(\sigma_F, \sigma_G, \sigma_H)$ presented above.

For 974 of the top 1000 advertisers, I found that the advertiser experienced a statistically significant payoff increase by sharing its data under each of the hundreds of tuples of values $(\sigma_F, \sigma_G, \sigma_H)$ considered. Moreover, 988 of the top 1000 advertisers either experienced a statistically significant payoff increase or no statistically significant change in payoff under each of the hundreds of tuples of values $(\sigma_F, \sigma_G, \sigma_H)$ considered. Only four of the top 1000 advertisers ever experienced a loss in payoff greater than 0.2% under at least one of the hundreds of tuples of values $(\sigma_F, \sigma_G, \sigma_H)$ considered, and only one of the top 1000 advertisers ever experienced a loss in payoff greater than 0.4% under at least one of the hundreds of tuples of values $(\sigma_F, \sigma_G, \sigma_H)$ considered.

Moreover, of the few advertisers who sometimes experienced a statistically significant payoff loss as a result of sharing their data under at least one of the hundreds of tuples of values $(\sigma_F, \sigma_G, \sigma_H)$ considered, only two advertisers experienced a statistically significant payoff decrease in at least 1% of the hundreds of tuples of values considered, and only one of these advertisers experienced a statistically significant payoff decrease in at least 3% of the hundreds of tuples of values considered. These results thus suggest that virtually all of the top 1000 advertisers would be better off sharing their data for auctions on mobile apps.

5.2. **Play Search.** Next I present the results for Play Search. Play Search has a mix of auctions in which only one advertisement is displayed as well as other auctions in which two advertisements are displayed. The auctions with only one advertising opportunity are equivalent to standard second-price auctions, so in these auctions I assume that each advertiser is making a bid that equals the advertiser's true value for an advertising opportunity.

The auctions with two positions are GSP auctions in which each bidder has a bidder-specific reserve price. In these auctions, if each bidder i makes a bid of b_i and has a bidder-specific reserve price of r_i, then bidders are ranked on the basis of the scores $b_i - r_i$, and each bidder pays a cost-per-click equal to the minimum bid the bidder needed to make in order to maintain its position. In particular, if bidder i is placed in position k, then this bidder will pay a cost-per-click equal to $b_{(k+1)} - r_{(k+1)} + r_i$, where $b_{(k+1)}$ denotes the bid of the advertiser in position $k+1$, and $r_{(k+1)}$ denotes this bidder's bidder-specific reserve price.

In these auctions, equilibrium bidding behavior is governed by strategies analogous to those in Edelman *et al.* (2007) and Varian (2007). In particular, if advertiser i does not have one of the two largest values of $v_i - r_i$, where v_i denotes advertiser i's value per click, then this advertiser will have an incentive to make a bid of $b_i = v_i$ in equilibrium. And if advertiser i does have one of the two largest values of $v_i - r_i$, then this advertiser will obtain a payoff of $x_2(v_i - (b_3 - r_3 + r_i))$ if the advertiser is placed in position two and a payoff of $x_1(v_i - (b_2 - r_2 + r_i))$ if the advertiser is placed in position one, where b_2 and r_2 denote the bid and reserve price of the bidder k with the highest competing value of $b_k - r_k$.

Thus advertiser i prefers to be placed in position one rather than position two if and only if $x_1(v_i - (b_2 - r_2 + r_i)) \geq x_2(v_i - (b_3 - r_3 + r_i)) \Leftrightarrow v_i - (b_2 - r_2 + r_i) \geq \frac{x_2}{x_1}(v_i - (b_3 - r_3 + r_i)) \Leftrightarrow$ $b_2 - r_2 \leq (1 - \frac{x_2}{x_1})(v_i - r_i) + \frac{x_2}{x_1}(b_3 - r_3)$. This in turn implies that advertiser i has a dominant strategy of making a bid b_i that satisfies $b_i - r_i = (1 - \frac{x_2}{x_1})(v_i - r_i) + \frac{x_2}{x_1}(b_3 - r_3)$ so that advertiser i will be placed in position 1 if and only if $b_2 - r_2 \leq (1 - \frac{x_2}{x_1})(v_i - r_i) + \frac{x_2}{x_1}(b_3 - r_3)$. Thus if one of the top two advertisers i is making a bid b_i, I assume that this advertiser's value v_i is a solution to the equation $b_i - r_i = (1 - \frac{x_2}{x_1})(v_i - r_i) + \frac{x_2}{x_1}(b_3 - r_3)$, which in turn holds if and only if $x_1(b_i - r_i) = (x_1 - x_2)(v_i - r_i) + x_2(b_3 - r_3) \Leftrightarrow v_i = r_i + \frac{x_1(b_i - r_i) - x_2(b_3 - r_3)}{x_1 - x_2}$.

Under these assumptions about advertisers' values, for each of the top 1000 advertisers, I conducted numerical simulations to measure how this advertiser's payoff would be affected in the subset of auctions in which the advertiser was in top position under each of the tuples of values $(\sigma_F, \sigma_G, \sigma_H)$ presented above. The results were as follows:

For 367 of the top 1000 advertisers, I found that the advertiser experienced a statistically significant payoff increase by sharing its data under each of the hundreds of tuples of values $(\sigma_F, \sigma_G, \sigma_H)$ considered. Moreover, 643 of the top 1000 advertisers either experienced a statistically significant payoff increase or no statistically significant change in payoff under each of the hundreds of tuples of values $(\sigma_F, \sigma_G, \sigma_H)$ considered. Only 102 of the top 1000 advertisers ever experienced a loss in payoff greater than 0.1% under at least one of the hundreds of tuples of values $(\sigma_F, \sigma_G, \sigma_H)$ considered, only 53 of the top 1000 advertisers ever experienced a loss in payoff greater than 0.2% under at least one of the hundreds of tuples of values $(\sigma_F, \sigma_G, \sigma_H)$ considered, and only 21 of the top 1000 advertisers ever experienced a loss in payoff greater than 0.4% under at least one of the hundreds of tuples of values $(\sigma_F, \sigma_G, \sigma_H)$ considered.

Moreover, of the minority of advertisers who sometimes experienced a statistically significant payoff loss as a result of sharing their data under at least one of the hundreds of tuples of values $(\sigma_F, \sigma_G, \sigma_H)$ considered, the majority of these advertisers only experienced a statistically significant payoff loss under a very small fraction of the hundreds of tuples of values $(\sigma_F, \sigma_G, \sigma_H)$ considered. For example, only 97 of the top 1000 advertisers experienced a statistically significant payoff loss in at least 3% of the hundreds of tuples of values considered, and only 68 of the top 1000 advertisers experienced a statistically significant payoff loss in at least 5% of the hundreds of tuples of values considered.

While these results are less favorable than the results on mobile apps, the results nonetheless indicate that vast majority of the top 1000 advertisers would be better off sharing their data for auctions on Play Search. The differences in the results stem from the fact that there tends to be a larger gap between the highest bid and the second-highest bid in the auctions on Play Search than in auctions on mobile apps. Because of this, it is less likely that sharing data will enable the winning advertiser to successfully identify advertising opportunities that the advertiser did not want to win on Play Search than it is on mobile apps. Thus, on average, sharing data is less beneficial to the the winning app advertisers on Play Search than it is on mobile apps.

Finally, it is worth remembering that these results are for the subset of auctions in which the advertiser in question initially had the highest bid in the auction. Since the auctions in which an advertiser had the highest bid in the auction are the subset of auctions where an advertiser is least likely to benefit from sharing its targeting data, it follows that the fraction of advertisers that would benefit from sharing targeting data would likely be even larger than what is suggested by this analysis.

6. Conclusion

This paper has analyzed whether advertisers would be better off using data that would enable them to target users more accurately if the only way they could use this data is by sharing the data with all advertisers and thereby enabling their competitors to also bid more accurately. I have presented both theoretical and empirical results that suggest that advertisers would typically be better off using their data in this case, even if this requires them to share their data and thereby enable their competitors to bid more accurately.

Theoretically I have shown that relatively stringent conditions will need to be met in order for an advertiser to become worse off as a result of sharing its data. In particular, it is typically necessary for (i) the competing advertisers to initially be bidding considerably less than the advertiser who shares its data, (ii) the strongest competing advertisers to initially be making similar bids to one another, and (iii) the competing advertisers to adjust their bids in an uncorrelated manner in order for sharing data to make an advertiser worse off. Under other circumstances, an advertiser will typically be better off sharing its data.

Empirically I have analyzed how sharing data would affect Google's top 1000 advertisers in auctions on mobile apps and Play Search when sharing data may (i) help the advertiser bid more accurately, (ii) help the competing advertisers bid more accurately on dimensions in which their values are correlated, and (iii) help the competing advertisers bid more accurately on dimensions where their values are independent. After analyzing hundreds of combinations of ways that the possibilities (i)-(iii) might affect advertisers' values for advertising opportunities, I found that the vast majority of the top 1000 advertisers would be better off sharing their data under each of these hundreds of possibilities. In particular, 98% of the top 1000 advertisers on mobile apps would always be better off using data as strategic inputs to their campaigns. Moreover, this held even though I restricted attention to auctions in which the advertiser initially had the highest bid in the auction, the subset of auctions in which the advertiser is least likely to benefit from sharing data.

Acknowledgments

I thank Jeffrey Colen, Florin Constantin, Evan Ettinger, Cameron Grace, Martin Handwerker, Chris Harris, Eric Hayashi, Mike Hochberg, Samuel Ieong, Davis King, Belinda Langer, Aranyak Mehta, David Mitby, Hal Varian, Sergei Vassilvitskii, Alana Vieira, and Di Wang for helpful comments and discussions.

Appendix

Proof of Theorem 2: If the values of m_j are drawn from distributions such that $\tilde{v}_2 \geq \tilde{v}_j$ for all $j > 2$ with probability 1, then the outcome of the auction will be the same as the outcome that would arise if the bidders $j > 2$ did not participate in the auction: Advertiser 1 will only lose the auction if advertiser 2 bids more than advertiser 1, and if advertiser 1 wins the auction then this advertiser will pay a cost equal to advertiser 2's bid.

But we know from Theorem 1 that if the bidders $j > 2$ do not participate in the auction, then advertiser 1 prefers to share its data with its competitors. From this it follows that advertiser 1 will obtain at least as large a payoff by sharing its data than by not sharing in this setting. □

Proof of Corollary 1: If the values of m_j are drawn from a distribution such that the values of m_j for $j \geq 2$ are all equal with probability 1, then $\tilde{v}_2 \geq \tilde{v}_j$ will hold for all $j > 2$ with probability 1 since $v_2 \geq v_j$ implies $mv_2 \geq mv_j$ for all m and thus $\tilde{v}_2 \geq \tilde{v}_j$ whenever $m_2 = m_j = m$. But since $\tilde{v}_2 \geq \tilde{v}_j$ holds for all $j > 2$ with probability 1, it follows from Theorem 2 that advertiser 1 will obtain at least as large a payoff by sharing its data than by not sharing in this setting. □

Lemma 1. *If an advertiser with value v is placed in position k, then this advertiser obtains a payoff of $x_k v - \sum_{j=k}^{s}(x_j - x_{j+1})b_{j+1} = \sum_{j=k}^{s}(x_j - x_{j+1})(v - b_{j+1})$.*

Proof. Since the advertiser in position k obtains x_k clicks and pays a cost-per-click of $c = \frac{1}{x_k}\sum_{j=k}^{s}(x_j - x_{j+1})b_{j+1}$, we know that this advertiser obtains a payoff of $x_k(v - c) = x_k(v -$

$\frac{1}{x_k}\sum_{j=k}^{s}(x_j - x_{j+1})b_{j+1})$. This then simplifies to $x_k v - \sum_{j=k}^{s}(x_j - x_{j+1})b_{j+1} = \sum_{j=k}^{s}(x_j - x_{j+1})(v - b_{j+1})$. □

Proof of Theorem 3: It suffices to prove this result for the case in which advertiser i would not change its bid as a result of sharing its data. Note that if advertiser i does not share its data, then the advertiser will be placed in position i and obtain a payoff of $\sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b_{j+1})$ (by Lemma 1). If the advertiser does share its data, then there are two possibilities:

The first possibility is that the advertiser will continue to be placed in position i after sharing its data. In that case, all of the advertisers in positions $j > i$ will still be making the same bids as before, so advertiser i will continue to obtain a payoff of $\sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b_{j+1})$.

The second possibility is that the advertiser will be placed in some position $k < i$ after sharing its data. In this case, advertiser i will obtain a payoff of $\sum_{j=k}^{s}(x_j - x_{j+1})(v_i - b'_{j+1})$, where b'_{j+1} denotes the new $j+1^{th}$-highest bid in the auction. Now $\sum_{j=k}^{s}(x_j - x_{j+1})(v_i - b'_{j+1}) \geq \sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b'_{j+1})$ since $k < i$ and $(x_j - x_{j+1})(v_i - b'_{j+1}) \geq 0$ for all $j \geq k$. And $\sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b'_{j+1}) \geq \sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b_{j+1})$ since $b'_{j+1} \leq b_{j+1}$ for all $j \geq i$. Thus $\sum_{j=k}^{s}(x_j - x_{j+1})(v_i - b'_{j+1}) \geq \sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b_{j+1})$, meaning the payoff the advertiser achieves by sharing its data will be at least as high as the payoff the advertiser would achieve if the advertiser did not share its data. The result then follows. □

Proof of Theorem 4: Note that if advertiser i does not share its data, then advertiser i will be placed in position i and obtain a payoff of $\sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b_{j+1})$ (by Lemma 1). To prove the result, I must then show that advertiser i's expected payoff will be at least this large if the advertiser shares its data. First I prove this for the case in which advertisers in positions $j \leq i$ do not change their bids as a result of advertiser i sharing its data.

Note that if advertiser i shares its data, then each advertiser $j > i$ will have its bid multiplied by some common multiplier m, and advertiser i will then be placed in the highest position k satisfying $v_i > mb_{k+1}$ and obtain a payoff $\sum_{j=k}^{s}(x_j - x_{j+1})(v_i - mb_{j+1})$. This payoff can then alternatively be expressed as $\sum_{j=k}^{s}(x_j - x_{j+1})(v_i - mb_{j+1}) = \sum_{j=i}^{s}(x_j - x_{j+1})(\max\{0, v_i - mb_{j+1}\})$.

Now since $E[m] = 1$ and $\max\{0, v_i - mb_{j+1}\}$ is a convex function of m, we know from Jensen's inequality that $E[\max\{0, v_i - mb_{j+1}\}] \geq \max\{0, v_i - b_{j+1}\}$. And since $\max\{0, v_i - b_{j+1}\} = v_i - b_{j+1}$ for all $j \geq i$, it then follows that $E[\max\{0, v_i - mb_{j+1}\}] \geq v_i - b_{j+1}$ for all $j \geq i$. Thus advertiser i's expected payoff from sharing data, $E[\sum_{j=i}^{s}(x_j - x_{j+1})(\max\{0, v_i - mb_{j+1}\})]$, satisfies $E[\sum_{j=i}^{s}(x_j - x_{j+1})(\max\{0, v_i - mb_{j+1}\})] \geq \sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b_{j+1})$, meaning this advertiser would achieve at least as large an expected payoff from sharing data as from not sharing data.

The above result was for the case in which advertisers in positions $j \leq i$ do not change their bids as a result of advertiser i sharing its data. But we know from the result in Theorem 3 that advertiser i can only benefit if sharing data induces advertisers in positions $j \leq i$ to change their bids as a result. Thus the conclusion in the previous paragraph also extends to settings in which all advertisers may change their bids as a result of data sharing. The result then follows. $\square$

Proof of Theorem 5: To prove this result it suffices to prove that if an arbitrary individual advertiser changes its bid as a result of sharing data, then this will make advertiser i at least as well off. We know from Theorem 3 that if an advertiser $j \leq i$ adjusts its bid as a result of advertiser i sharing its data, then this will necessarily make advertiser i at least as well off. Thus it suffices to prove that if some advertiser $j > i$ adjusts its bid as a result of advertiser i sharing its data, then this will make advertiser i at least as well off.

If advertiser i does not share its data, then the advertisers will be ranked in order and advertiser i will obtain a payoff of $\sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b_{j+1})$ (by Lemma 1). However, if advertiser i shares its data, and advertiser $k > i$ adjusts its bid from b_k to b'_k, then advertiser i's payoff will depend on which of the following circumstances the advertiser is in:

The first possibility is that advertiser k will remain in position k even after changing its bid. In this case, advertiser i will obtain a payoff of $\sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b'_{j+1})$, where $b'_{j+1} \equiv b_{j+1}$ if $j + 1 \neq k$.

The second possibility is that advertiser k will move down to some lower position $l > k$ after changing its bid. In this case, advertiser i will obtain a payoff of $x_i v_i - [\sum_{j=i}^{k-2}(x_j - x_{j+1})b_{j+1} + \sum_{j=k-1}^{l-2}(x_j - x_{j+1})b_{j+2} + (x_{l-1} - x_l)b'_k + \sum_{j=l}^{s}(x_j - x_{j+1})b_{j+1}]$. Now since $x_j - x_{j+1}$ is increasing in j for all $j \leq s$ and $b'_k < b_l$, we know that $\sum_{j=k-1}^{l-2}(x_j - x_{j+1})b_{j+2} + (x_{l-1} - x_l)b'_k \leq (x_{k-1} - x_k)b'_k + \sum_{j=k}^{l-1}(x_j - x_{j+1})b_{j+1}$. Thus advertiser i's payoff from sharing data in this case is greater than or equal to $x_i v_i - [\sum_{j=i}^{k-2}(x_j - x_{j+1})b_{j+1} + (x_{k-1} - x_k)b'_k + \sum_{j=k}^{l-1}(x_j - x_{j+1})b_{j+1} + \sum_{j=l}^{s}(x_j - x_{j+1})b_{j+1}] = x_i v_i - \sum_{j=i}^{s}(x_j - x_{j+1})b'_{j+1} = \sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b'_{j+1})$, where $b'_{j+1} \equiv b_{j+1}$ if $j + 1 \neq k$.

The third possibility is that advertiser k will move up to some higher position $l \in (i, k)$ after changing its bid. In this case, advertiser i will obtain a payoff of $x_i v_i - [\sum_{j=i}^{l-2}(x_j - x_{j+1})b_{j+1} + (x_{l-1} - x_l)b'_k + \sum_{j=l}^{k-1}(x_j - x_{j+1})b_j + \sum_{j=k}^{s}(x_j - x_{j+1})b_{j+1}]$. Now since $x_j - x_{j+1}$ is increasing in j for all $j \leq s$ and $b'_k > b_l$, we know that $(x_{l-1} - x_l)b'_k + \sum_{j=l}^{k-1}(x_j - x_{j+1})b_j \leq \sum_{j=l-1}^{k-2}(x_j - x_{j+1})b_j + (x_{k-1} - x_k)b'_k$. Thus advertiser i's payoff from sharing data in this case is greater than or equal to $x_i v_i - [\sum_{j=i}^{l-2}(x_j - x_{j+1})b_{j+1} + \sum_{j=l-1}^{k-2}(x_j - x_{j+1})b_j + (x_{k-1} - x_k)b'_k + \sum_{j=k}^{s}(x_j - x_{j+1})b_{j+1}] = x_i v_i - \sum_{j=i}^{s}(x_j - x_{j+1})b'_{j+1} = \sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b'_{j+1})$, where $b'_{j+1} \equiv b_{j+1}$ if $j + 1 \neq k$.

The final possibility is that advertiser k will move up to some higher position $l \geq i$ after changing its bid. In this case, advertiser i will obtain a payoff of $\sum_{j=i+1}^{s}(x_j - x_{j+1})(v_i - b'_{j+1})$, where $b'_{j+1} \equiv b_{j+1}$. Now since $x_j - x_{j+1}$ is increasing in j for all $j \leq s$, we know that $\sum_{j=i+1}^{s}(x_j - x_{j+1})(v_i - b'_{j+1}) \geq \sum_{j=i}^{k-2}(x_j - x_{j+1})(v_i - b'_{j+1}) + \sum_{j=k}^{s}(x_j - x_{j+1})(v_i - b'_{j+1})$. And since $b'_k > v_i$, we also know that $\sum_{j=i}^{k-2}(x_j - x_{j+1})(v_i - b'_{j+1}) + \sum_{j=k}^{s}(x_j - x_{j+1})(v_i - b'_{j+1}) \geq \sum_{j=i}^{k-2}(x_j - x_{j+1})(v_i - b'_{j+1}) + (x_{k-1} - x_k)(v_i - b'_k) + \sum_{j=k}^{s}(x_j - x_{j+1})(v_i - b'_{j+1}) = \sum_{j=i+1}^{s}(x_j - x_{j+1})(v_i - b'_{j+1})$. Thus advertiser i's payoff in this case is again greater than or equal to $\sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b'_{j+1})$.

By combining the results in the previous four paragraphs, we see that regardless of how advertiser k adjusts its bid in response to advertiser i sharing data, advertiser i's payoff will be greater than or equal to $\sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b'_{j+1})$. Since $E[\sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b'_{j+1})] = \sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b_{j+1})$, and advertiser i would obtain a payoff of $\sum_{j=i}^{s}(x_j - x_{j+1})(v_i - b_{j+1})$

if the advertiser did not share its data, it follows that advertiser i will be at least as well off if it shares its data. □

Proof of Theorem 6: If the values of m_j are drawn from distributions such that $\tilde{v}_{s+1} \geq \tilde{v}_j$ for all $j > s+1$ with probability 1, then the outcome of the auction will be the same as the outcome that would arise if the bidders $j > s+1$ did not participate in the auction: Advertiser i will be placed in position k if and only if exactly $k-1$ of the top $s+1$ advertisers bid more than advertiser i. And if advertiser i is placed in position k, the advertiser's cost-per-click will be determined by the bids of the bidders in positions $k+1$ through $s+1$, which are the same as they would be if the bidders $j > s+1$ did not participate in the auction.

But we know from Theorem 5 that if the bidders $j > s+1$ do not participate in the auction, then each advertiser prefers to share its data with its competitors. From this it follows that each advertiser will obtain at least as large a payoff by sharing its data than by not sharing in this setting. □

References

Abraham, Ittai, Susan Athey, Moshe Babaioff, and Michael Grubb. 2016. "Peaches, Lemons, and Cookies: Designing Auction Markets with Dispersed Information." Microsoft Research Typescript.

Bergemann, Dirk and Alessandro Bonatti. 2011. "Targeting in Advertising Markets: Implications for Offline Versus Online Media." *RAND Journal of Economics* 42(3): 417-443.

Bergemann, Dirk and Martin Pesendorfer. 2007. "Information Structures in Optimal Auctions." *Journal of Economic Theory* 137(1): 580-609.

Bergemann, Dirk and Juuso Välimäki. 2006. "Information in Mechanism Design." In: *Advances in Economics and Econometrics*, ed. by Richard Blundell, Whitney K. Newey, and Torsten Persson, pp. 186-221. Cambridge: Cambridge University Press.

Bergemann, Dirk and Achim Wambach. 2015. "Sequential Information Disclosure in Auctions." *Journal of Economic Theory* 159 (Part B): 1074-1095.

Bhawalkar, Kshipra, Patrick Hummel, and Sergei Vassilvitskii. 2014. "Value of Targeting." *Proceedings of the* 7^{th} *International Symposium on Algorithmic Game Theory* (SAGT) 194-205.

Board, Simon. 2009. "Revealing Information in Auctions: The Allocation Effect." *Economic Theory* 38(1): 125-135.

De Cornière, Alexandre and Romain De Nijs. 2016. "Online Advertising and Privacy." *RAND Journal of Economics* 47(1): 48-72.

Edelman, Benjamin, Michael Ostrovsky, and Michael Schwarz. 2007. "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars of Keywords." *American Economic Review* 97(1): 242-259.

Emek, Yuval, Michal Feldman, Iftah Gamzu, Renato Paes Leme, and Moshe Tennenholtz. 2014. "Signaling Schemes for Revenue Maximization." *ACM Transactions on Economics and Computation* 2(2): Article No. 5.

Eső, Péter and Balázs Szentes. 2007. "Optimal Information Disclosure in Auctions and the Handicap Auction." *Review of Economic Studies* 74(3): 705-731.

Fu, Hu, Patrick Jordan, Mohammad Mahdian, Uri Nadav, Inbal Talgam-Cohen, and Sergei Vassilvitskii. 2012. "Ad Auctions with Data." *Proceedings of the* 5^{th} *International Symposium on Algorithmic Game Theory* (SAGT) 184-189.

Ganuza, Juan-José. 2004. "Ignorance Promotes Competition: An Auction Model with Endogenous Private Valuations." *RAND Journal of Economics* 35(3): 583-598.

Ganuza, Juan-José and José S. Penalva. 2010. "Signal Orderings Based on Dispersion and the Supply of Private Information in Auctions." *Econometrica* 78(3): 1007-1030.

Ghosh, Arpita, Mohammad Mahdian, R. Preston McAfee, and Sergei Vassilvitskii. 2015. "To Match or not to Match: Economics of Cookie Matching in Online Advertising." *ACM Transactions of Economics and Computation* 3(2): Article 7.

Ghosh, Arpita, Hamid Nazerzadeh, and Mukund Sundararajan. 2007. "Computing Optimal Bundles for Sponsored Search." *Proceedings of the* 3^{rd} *International Worksho on Internet and Network Economics* (WINE) 576-583.

Hummel, Patrick and R. Preston McAfee. 2016. "When Does Improved Targeting Increase Revenue?" *ACM Transactions of Economics and Computation* 5(1): Article No. 4.

Lahaie, Sebastién and R. Preston McAfee. 2011. "Efficient Ranking in Sponsored Search." *Proceedings of the* 7^{th} *International Workshop on Internet and Network Economics* (WINE) 254-265.

Levin, Jonathan and Paul Milgrom. 2010. "Online Advertising: Heterogeneity and Conflation in Market Design." *American Economic Review: Papers and Proceedings* 100(2): 603-607.

Milgrom, Paul R. and Robert J. Weber. 1982. "A Theory of Auctions and Competitive Bidding." *Econometrica* 50(5): 1089-1122.

Ostrovsky, Michael and Michael Schwarz. 2016. "Reserve Prices in Internet Advertising Auctions: A Field Experiment." Stanford Graduate School of Business Typescript.

Palfrey, Thomas. 1983. "Bundling Decisions by a Multiproduct Monopolist with Incomplete Information." *Econometrica* 51(2): 463-483.

Sun, Yang, Yunhong Zhou, and Xiaotie Deng. 2014. "Optimal Reserve Prices in Weighted GSP Auctions." *Electronic Commerce Research and Applications* 13(3): 178-187.

Varian, Hal. 2007. "Position Auctions." *International Journal of Industrial Organization* 25(6): 1163-1178.