Who Welcomes Behavioral Targeting: An Economic Analysis

Jianqing Chen  
University of Calgary  
jianqing.chen@haskayne.ucalgary.ca

Jan Stallaert  
University of Connecticut  
jan.stallaert@business.uconn.edu

September 15, 2010

Abstract

Recently there has been an increased interest in using targeted advertising online: a user is presented with advertisements that are more suitable, based on her past browsing and search behavior and other available information (e.g., her gender and age information registered on a website). This technique is known as behavioral targeting and besides receiving scrutiny from user advocacy groups as being potentially intrusive to online privacy, has been hailed as the new Holy Grail in online advertising. In this paper, we study the economic implications when an online publisher engages in behavioral targeting and selects an advertisement based on the best fit with a user. The online publisher auctions off the advertising slot and is being paid on pay-per-click basis. Using a horizontal differentiation model, we show that revenue for the online publisher may in some circumstances double when using behavioral targeting. In addition, we demonstrate that almost all advertisers are better off under this scheme and prove how these results are strongly affected by the number of advertisers and valuations of click-throughs.
1 Introduction

Advances in information technology have radically changed online advertising, most notably in the measurability of advertising outcomes and in tailoring online advertisements to an end-user’s interests. Information technology can now easily monitor clicks on a specific advertisement—which to some degree is regarded as the effectiveness of advertising—and subsequently pay-per-click has become a new standard pricing practice for online advertising. Meanwhile, technology enables the delivery of more targeted advertisements to consumers, for example, based on the keyword a consumer enters with a search engine or the location of consumers inferred from their computers’ IP addresses. One radical innovation in targeted advertising recently is behavioral targeting, a technology aimed to increase the effectiveness of advertising by online publishers. Behavioral targeting uses information collected from an individual’s web-browsing behavior—such as the pages they have visited or the searches they have made—in the selection of advertisements to display. Whereas previously the data recorded about user behavior on the Web was mainly restricted to the use of cookies recording a user’s Web visits, recent technology is much more sophisticated and detailed about a user’s actions and online behavior. A recent study by the Wall Street Journal (Angwin, 2010) found that the nation’s top 50 websites install on average 64 pieces of tracking technology, usually with no warning. Advertising using behavioral targeting may become a sizeable industry: eMarketer estimates that online advertisers will spend over $1.1 billion in targeted advertising in 2010, rising to over $2.6 billion in 2014 (Hallerman, 2010). The move to behavioral targeting can be explained by its highly increased effectiveness. Experiments have shown that click-through rates can be increased as much as 670% using behavioral targeting (Yan et al., 2009).

Despite such dramatic potential improvements for the advertiser and the online publisher, some users and user advocacy groups have expressed concerns over the privacy issues raised by behavioral targeting (Clifford, 2009). To this date, the Federal Trade Commission (FTC) has attempted to let advertisers and publishers self-regulate: the Commission has established...
a set of principles that ISPs and other collectors of user behavioral data should heed (FTC Staff Report, 2009). One such principle is that the data collector should receive “affirmative express consent [from the user] to the Use of Sensitive Data.” Online publishers such as Google demand the user to explicitly opt in before collecting any sensitive data and allows the user to select what information can be gathered and what it can be used for.

Increasingly, the Web is being used for high-bandwidth video. Youtube and ESPN3 are solely devoted to transmitting video, but more traditional Web sites such as CNN and the New York Times increasingly incorporate video streams. It is common to have a sponsored video advertisement (which can be tailored to the user, if wanted) displayed before the video starts. Ultimately, such targeted advertisements could be sent to TVs using a technology called “addressable television.” This much touted technology where different advertisements can be sent to different TV sets was until recently thought to be “vaporware” rather than a practical reality. However, new advances have been made in the technology as evidenced by Google’s recent investment in Invidi in May 2010 (Kafka, 2010). The possibility of addressable television together with the increasing use of unicast advertising for sponsored high-bandwidth video brings new possibilities for advertisers as well as advertising publishers.

In this paper, we assume that one advertising slot is for sale. This slot may be used for unicast advertising (i.e., behavioral targeting) in which different advertisement are displayed based on the fit with a user as described above, or for more traditional advertising in which the same advertisement is displayed for all users. Advertisers derive a benefit from users clicking on the advertisements, indicating possible consumer interest for the good or service advertised. Users (or potential customers) have different preferences for the advertisements, which translates into a different likelihood of clicking the advertisement. We assume that there is no universal ranking of the preferences over the user population and that every advertisement has at least some user(s) who prefer this advertisement over all others. In other words, we rule out the existence of superior advertisements that are best liked by the whole user population, or inferior advertisements which all users would rank last. Hence, we
assume that advertisements are not vertically but horizontally differentiated. The publisher puts the slot up for sale using an auction, and chooses the winner(s) according to the products of users’ probabilities of click-through and the advertisers bids for a click-through, similar to the weighted unit-price score (Liu et al., 2010). The advertisers pay according to the second weighted unit price rule when he wins the auction, which is the payment scheme most commonly used in online advertising auctions at present.

The question we attempt to answer is who benefits (and what are the conditions required) from behavioral targeting as compared to traditional advertising? Would the online publisher benefit from targeting the advertisements? Since targeting advertisements is similar to product differentiation, it will give rise to relaxed competition (Tirole, 1988) between the advertisers and hence it is possible that advertisers need not bid as aggressively for their advertisements to be displayed under behavioral targeting. That is, by focusing on a specific user segment, an advertiser’s advertisement may be selected with a relatively low bid on this segment whereas under traditional advertising his advertisement would never have been selected. This competitive effect will depress the online publisher’s income by realizing a lower revenue per click-through.

On the other hand, the negative effect of relaxed competition for the online publisher may be offset by a positive propensity effect. By targeting advertisements, the probability of a click-through is increased resulting in a higher volume of click-throughs. At the outset, it is not clear whether the competitive effect or the propensity effect will dominate, or which conditions drive the net effect. In addition, we study what the payoff implications are for the online advertisers and whether they would prefer to join in behavioral targeting, as well as the effect on social welfare of publisher and advertisers.

We find that when the number of advertisers is large enough and the advertisers have similar valuations, behavioral targeting is better for both the online publisher and (most of) the advertisers. We find that if the users are sufficiently heterogeneous and the number of advertisers grows infinitely large, the expected income for the publisher can double as
compared to traditional advertising. We also demonstrate that the impact of behavioral targeting on the publisher’s revenue is mainly driven by the number of advertisers and by the degree of disparity among the advertisers in the valuation of click-throughs.

To our knowledge, this research is among the first to study the economic impact of behaviorally targeted advertising when advertising slots are auctioned off. Our analysis indicates that if online publishers are careful in addressing users’ privacy concerns—so that they opt in the targeted advertising—they can substantially increase their expected revenue without harming the advertisers or users.

The rest of the paper is organized as follows. In the next section, we discuss the related literature. We set forth our model in section 3. In section 4 we provide an equilibrium analysis and study the effect of behavioral targeting on the publisher, advertisers, and social welfare. Section 5 concludes the paper.

2 Literature Review

Behavioral targeting is a relatively new phenomenon and there has been no much research done in addressing the similar questions raised in this paper. One related stream of literature is the studies on traditional targeted advertising. Gal-Or et al. (2006) study how an advertiser should allocate his resources to increase the quality of his targeting. They show how improved accuracy and recognition of consumers affects price competition and profitability for this advertiser. Iyer et al. (2005) look at the situation with multiple advertisers and show how the firms should target their advertisements to the consumers with high preference for their product and away from comparison shoppers. They derive conditions under which targeted advertising will be profitable and which advertiser would adopt the targeted advertising strategy. Esteban and Hernandez (2007) show that targeted advertising and targeted pricing can lead to fragmentation into local monopolies. To our knowledge the literature in targeted advertising has not addressed the situation where the publisher allocates the advertising slot
based on individual users’ preferences and advertisers compete for advertising slots via an auction.

Our research is loosely related to the literature on bundling (e.g., Palfrey (1983)), as behavioral targeting can be seen as unbundling customers and allocating different segments to different advertising firms. In Ghosh et al. (2007) auctions for sponsored search are studied and the authors find cases in which it is better for the online publisher to bundle different contexts so as to maximize his revenue. Our research contributes to this stream of literature by integrating an auction within a horizontal differentiation model.

3 Baseline Model

We consider a case in which \( n \) advertisers compete for an advertising slot offered by a publisher in a specific context. A group of online users with a measure of one unit may view the advertising slot. The advertisers fit each user’s interest or need to different degrees, and the users have different preferences across the advertisers and thus have different probabilities to click on the advertisements from different advertisers. In particular, we use a circular model (Salop, 1979) to represent users’ preference or probabilities of clicking on the advertisements from different advertisers.

We assume that the advertisers, indexed by \( i = 1, \ldots, n \), are symmetrically distributed along a circle clockwise in the order of 1, 2, \ldots, and \( n \). The perimeter of the circle is 1 and thus the shortest distance between any two adjacent advertisers along the circle is \( 1/n \). Unless indicated otherwise, we use the term distance to refer to the shortest distance along the circle. Each user is represented by a point on the circle and users are uniformly distributed along the circle. The distance between a user and an advertiser reflects the degree of matching between the user and the advertiser: the longer the distance is, the lower the degree of the matching is. Therefore, the most favorable user for an advertiser is the user that is located at the same spot as the advertiser. For advertiser \( i \), we denote \( p_i \) as the probability of his most
favorable user clicking on its advertisement if displayed. The probabilities of other users clicking his advertisement are decayed to different degrees depending on their distances to the advertiser. If the distance between advertiser $i$ and user $j$ is $x_{ij}$, the decay $q_{ij}$ is

$$q_{ij} = 1 - \gamma x_{ij}$$

such that the probability of user $j$ clicking on advertiser $i$ is $p_i q_{ij}$, where $x_{ij} \in [0, 1/2]$ and $\gamma$ is the decay factor ($0 < \gamma \leq 2$). Notice that $\gamma$ also measures the heterogeneity of users’ preferences and a high $\gamma$ means that users have very different probabilities of clicking on an advertisement. We also call the probability $p_i q_{ij}$ advertiser $i$’s click-through rate facing user $j$. Figure 1 illustrates the basic elements of the model, with each small circle representing an advertiser (e.g., advertiser $i$) and any point (e.g., user $j$) on the big circle representing a user.

The publisher knows the advertisers’ types, characterized by the locations on the circle. The publisher can learn a user’s preference (e.g., by monitoring its browsing history and the registered information) if the user opts in the behavioral advertising and hence is able to determine the location of each user on the circle.

Following the common practice in online advertising, we assume that the publisher uses unit-price auctions to sell the advertising slot. Advertisers bid on per-click unit price $b$ for the
slot and the winner is chosen based on weighted unit prices, that is, the product of their bids and the (expected) click-through rates measured by \( p_i q_{ij} \). Under behavioral targeting, the publisher identifies each user’s preference/click-through rates across advertisers and uses this information to choose the advertiser being displayed to a specific user. As the users differ in their preference/click-through rates, different advertisements may be presented to different users. Under traditional advertising, the publisher does not learn or does not use users’ preference information and one auction is used for the whole group of users. The winner is chosen based on his bid and expected click-through rate across all users, and the same advertisement is displayed to each user. Notice that the per-click bid times the (expected) click-through rate is the expected payment that an advertiser bids. Therefore, the winner of an auction is essentially determined by advertisers’ expected payments for the user(s). The advertiser with the highest proposed expected payment wins the auction, and pays the unit price that matches the second highest proposed expected payment (i.e., second-score rule in Liu et al. 2010, similar to second-price auctions). As the winning unit price times its click-through rate is its expected payment, the winner is equivalently paying the second highest proposed expected payment for the user(s).

We denote the unit value that advertiser \( i \) derives from each click by \( v_i \), and thus the value of user \( j \) to advertiser \( i \) is \( v_i p_i q_{ij} \). We let \( z_i \equiv v_i p_i \) and call \( z_i \) advertiser \( i \)'s reference value, which is the value that advertiser \( i \) derives from its most favorable user. Then the value of user \( j \) to advertiser \( i \) is advertiser \( i \)'s reference value \( z_i \) discounted by the decay \( q_{ij} \). Notice that we can always rank \( z_i \) in non-increasing order and denote the \( i \)-th highest value as \( z(i) \) such that \( z(1) \) is the highest value among all \( z_i \)'s, and \( z(2) \) is the second highest, and so on. For ease of exposition, we here assume that \( z(n) \geq z(1)(1 - \frac{2}{n}) \); that is, even if the advertiser with the lowest value happens to be adjacent to the one with the highest value, it is still socially efficient to present the lowest-value advertiser to its most favorable user. In other words, the existence of any advertiser, even the advertiser with the lowest value, brings in additional social welfare (which can also be seen as a situation where there are no
strictly dominated advertisers).

4 Equilibrium Analysis

We first provide a general analysis of advertisers’ equilibrium bidding and the publisher’s equilibrium payoff under traditional advertising and behavioral targeting. We study the conditions under which the publisher prefers behavioral targeting over traditional advertising and the conditions under which some advertisers are better off under behavioral targeting. We also study the social welfare under the two advertising strategies. We then focus on two special cases for discussion: one where all advertisers have the same value and one where there exists an advertiser having a reference value significantly higher than the others.

4.1 General Analysis

Under traditional advertising, the publisher does not track a specific user’s behavior and does not learn their preference. The winner determination and winning price are based on advertisers’ overall expected click-through rates. The expected decay across all users for any advertiser is

$$E(q) = 2 \int_{0}^{1/2} (1 - \gamma x)dx = 1 - \frac{\gamma}{4}$$

and hence the overall expected click-through rate for advertiser $i$ is $p_i(1 - \gamma/4)$. Following the standard argument for the equilibrium bidding under second-price auctions, we have the following lemma.

**Lemma 1.** Under traditional advertising, bidding their true per-click unit value is advertisers’ (weakly) dominant strategy.

**Proof.** All proofs are in the appendix unless indicated otherwise. \qed

Basically, if advertisers bid lower than their true unit value, they may lose some auctions which they could have won by bidding the true unit value. If they bid higher than their true
unit value, advertisers may risk earning negative payoff from winning auctions with price higher than the true value to them. Therefore, bidding the true unit value is advertisers’ best action and equilibrium strategy.

As a result, in equilibrium, the advertiser with the highest expected payment who wins the advertising slot is the one with the highest expected value \( \langle v_i, p_i E(q) \rangle \). In other words, the advertiser with the highest reference value wins the advertising slot (by noting \( z_i = v_i p_i \)).

The publisher’s revenue is the second highest expected payment and is thus the second highest expected value

\[
\pi_T = \left[ b_i p_i \left(1 - \frac{\gamma}{4}\right) \right]_{(2)} = \left[ v_i p_i \left(1 - \frac{\gamma}{4}\right) \right]_{(2)} = (1 - \frac{\gamma}{4}) z_{(2)}
\]

where \([\cdot]_{(2)}\) is the operator for the second highest value in the bracket among all \(i\)'s.

Under behavioral targeting, the publisher knows a user’s preference and allocates the advertising slot based on advertisers’ bids and how well an advertiser fits the user’s preference. Specifically, the advertiser with the highest \( b_i p_i q_{ij} \) will be presented in the advertising slot for user \( j \). So, different advertisers may win the advertising slot for different users because users generally differ in their preference represented by \( q_{ij} \). Along a similar argument, we can show that bidding true unit value is advertisers’ equilibrium strategy.

**Lemma 2.** Under behavioral advertising, bidding their true per-click unit value is advertisers’ (weakly) dominant strategy.

Provided that advertisers bid their true unit value in equilibrium, an advertiser’s proposed expected payment for a user is also its value of the user. In auctioning the advertising slot for user \( j \), the publisher assigns the slot to the advertiser with the highest value \( v_i p_i q_{ij} = z_i q_{ij} \); that is, when user \( j \) visits the web, the advertiser with the highest value wins the auction and is presented in the slot to the user.

We next show that any user \( j \) located between advertisers \( i \) and \( i + 1 \) must be assigned to either advertiser \( i \) or advertiser \( i + 1 \) in equilibrium. For any advertiser \( k \) different from
Advertisers $i$ and $i + 1$, we suppose the shortest path from user $j$ to advertiser $k$ (with distance $x_{kj}$) passes advertiser $i$. If advertiser $i$ has a higher reference value than advertiser $k$ does (i.e., $z_i \geq z_k$), advertiser $i$ certainly has higher value of user $j$ as user $j$ has less decay for advertiser $i$ (i.e., $q_{ij} > q_{kj}$ since $x_{ij} < x_{kj}$). Otherwise (i.e., if $z_i < z_k$), we notice that for advertiser $i$’s most favorable user advertiser $i$ has a higher value than advertiser $k$ does under the assumption on value relationship $z(n) \geq z(1)(1 - \frac{2}{n})$. User $j$ has the same additional decay from advertiser $i$’s most favorable user to both advertisers, but the decay has greater negative effect on advertiser $k$’s value as it has a higher reference value. Therefore, advertiser $i$ has higher value of user $j$ as well. Technically, this conclusion can also be seen by noticing (assuming $k < i$)

$$z_k(1 - x_{kj}\gamma) = z_k \left[1 - \left(\frac{i-k}{n} + x_{ij}\right)\gamma\right] < z_i - z_k x_{ij}\gamma < z_i (1 - x_{ij}\gamma)$$

where the first inequality follows from the assumption on value relationship and the second inequality is because of $z_i < z_k$. All together, we can see that either advertiser $i$ or $i + 1$ has the highest value of any user $j$ located between them so the user is assigned to one of the advertisers.

Within the user segment between advertisers $i$ and $i + 1$ or user segment $i|(i + 1)$, the users who have a strong preference for advertiser $i$ are assigned to advertiser $i$ and the others are assigned to advertiser $i + 1$. A marginal user who has the same value to both advertisers exists in segment $i|(i + 1)$; that is, the marginal user determined by

$$(1 - \gamma x_i)z_i = \left[1 - \gamma\left(\frac{1}{n} - x_i\right)\right] z_{i+1}$$

where $x_i$ is the marginal user’s distance from advertiser $i$ as illustrated in Figure 2. We can derive

$$x_i = \frac{(z_i - z_{i+1}) + \frac{1}{n}\gamma z_{i+1}}{\gamma(z_i + z_{i+1})}$$

(1)
It is worth noting that the assumption on value relationship ensures that $x_i \in [0, 1/n]$ such that each advertiser has some market coverage.

So, for the above user segment $i|(i+1)$, advertiser $i$ wins the users within distance $x_i$ and advertiser $i+1$ wins the rest. Now we compute advertiser $i$’s payment for each user $j$ within distance $x_i$. Advertiser $i$ pays the second highest proposed expected payment for any user within distance $x_i$. In equilibrium, the second highest proposed expected payment for user $j$ is the highest value of user $j$ to the remaining advertisers. This value must be from either advertiser $i+1$ or advertiser $i-1$. The reason is similar to that for above finding that either advertiser $i$ or $i+1$ has the highest value of any user between them. For instance, if the shortest path from an advertiser to user $j$ passes advertiser $i+1$, we can similarly conclude that the advertiser’s value of the user is not as high as advertiser $i+1$’s. Again, if advertiser $i+1$ has a higher reference value than the other advertiser, advertiser $i+1$ has a higher value of user $j$ because of the less decay. Otherwise, we notice for advertiser $i+1$’s most favorable user advertiser $i+1$ has a higher value than the other advertiser under the assumption on value relationship. User $j$ has the same additional decay from advertiser $i+1$’s most favorable user to both advertisers, but the decay has greater negative effect on the other advertiser’s value as it has a higher reference value. Therefore, advertiser $i+1$ has a higher value of user $j$ as well. We next distinguish two cases: $v_{i+1} \geq v_{i-1}$ and $v_{i+1} < v_{i-1}$,
and examine the expected payment in each case.

When $v_{i+1} \geq v_{i-1}$, advertiser $i+1$’s value of any user that advertiser $i$ wins in the segment $i|(i+1)$ is higher than advertiser $i-1$’s (see Figure 2a). Therefore, the expected payment from advertiser $i$ for the users that the advertiser wins within this segment is

$$z_{i+1} \int_0^{x_i} \left[ 1 - \gamma \left( \frac{1}{n} - x \right) \right] dx = z_{i+1} x_i \left( 1 - \frac{1}{n} \gamma + \frac{x_i}{2} \gamma \right)$$

(2)

which is the size of the light gray area in Figure 2a.

When $v_{i+1} < v_{i-1}$, there are some users (close to advertiser $i$) within segment $i|(i+1)$ for which advertiser $i-1$’s value is higher than advertiser $i+1$’s (see Figure 2b). In this case, the price for those users close to advertiser $i$ will be determined by the value of advertiser $(i-1)$ (the second highest value among all advertisers). For some user in segment $i|(i+1)$ located at distance $y_i$ from advertiser $i$, advertisers $i-1$ and $i+1$ have the same value. We call this user the marginal user for payment and the location of the marginal user for payment (i.e., the value of $y_i$) is determined by the equality: $[1 - \gamma (\frac{1}{n} + y_i)] z_{i-1} = [1 - \gamma (\frac{1}{n} - y_i)] z_{i+1}$.

Hence the distance from the marginal user for payment to advertiser $i$ in this case is given by:

$$y_i = \frac{(1 - \frac{2}{n})(z_{i-1} - z_{i+1})}{\gamma (z_{i-1} + z_{i+1})}$$

The price that advertiser $i$ pays for the users located within distance $y_i$ is the value of advertiser $i-1$ and the price that advertiser $i$ pays for the rest of the users is the value of advertiser $i+1$. So, when $v_{i+1} < v_{i-1}$, the expected payment for all users in segment $i|(i+1)$ allocated to advertiser $i$ is

$$z_{i-1} \int_0^{y_i} \left[ 1 - \gamma \left( \frac{1}{n} + x \right) \right] dx + z_{i+1} \int_{y_i}^{x_i} \left[ 1 - \gamma \left( \frac{1}{n} - x \right) \right] dx = z_{i+1} x_i \left( 1 - \frac{1}{n} \gamma + \frac{x_i}{2} \gamma \right) + (z_{i-1} - z_{i+1}) \left( 1 - \frac{\gamma}{n} \right) \frac{y_i}{2}$$

(3)

which is the size of the light gray and dark gray areas in Figure 2b.
The difference in the expected payments in Eq. (2) and Eq. (3) lies in the fact that the latter has an extra term \((z_i - 1 - z_i + 1)(1 - \gamma n)\frac{y}{2}\), which is the size of the dark gray area in Figure 2b. The extra term or additional expected payment/revenue to the publisher is because advertiser \(i\)'s payments in both segments are affected by his neighbor with the higher reference value: the neighbor with the higher reference value not only raises the payments for advertiser \(i\) in the segment between them (i.e., segment \((i - 1)|i\) in Figure 2b), but also raises the price for some users in the segment at the other side of advertiser \(i\) (i.e., the users in the dark gray area in Figure 2b). We call such extra term the cross-border effect. We call the term in Eq. (2) or the first term in Eq. (3) the base payment.

The expected payment from advertiser \(i\) for its users within user segment \((i - 1)|i\) can be derived similarly. Depending on the relative competitiveness in advertisers \(i - 1\) and \(i + 1\), advertiser \(i\) pays the cross-border effect in one segment or the other. In either way, the cross-border effect takes the same form \(\frac{(1 - \frac{1}{n})^2(z_{i-1} - z_{i+1})^2}{2\gamma(z_{i-1} + z_{i+1})}\). Considering the expected payment from all \(n\) advertisers, we can formulate the expected revenue for the publisher as

\[
\pi_B = \sum_{i=1}^{n} z_{i+1} x_i \left[1 - \frac{1}{n} \gamma + \frac{\gamma x_i}{2}\right] + z_{i-1} \left(\frac{1}{n} - x_{i-1}\right) \left[1 - \frac{1}{n} \gamma + \frac{\gamma}{2} \left(\frac{1}{n} - x_{i-1}\right)\right] + \frac{(1 - \frac{1}{n})^2(z_{i-1} - z_{i+1})^2}{2\gamma(z_{i-1} + z_{i+1})}
\]

where we let \(z_0 \equiv z_n\) and \(z_{n+1} \equiv z_1\) to simplify the notation.

We first examine how the revenue under behavioral targeting changes with each advertiser's reference value. According to the auction rule, the expected revenue generated from each user is the second highest proposed expected payment for the user, or the second highest value of the user to the advertisers. When an advertiser's reference value is increased and so is his value of the user, the second highest value of the user to the advertisers is (weakly) increased. Under the value relationship assumption, the advertiser has some market coverage under his original reference value and captures some additional users from his neighbor when increasing his reference value. For those users, the second highest value of each user after the increase is the highest value of each user before the increase. Therefore, the expected
revenue from those users is strictly increasing and so is the total expected revenue.

**Lemma 3.** The revenue under behavioral advertising is increasing in $z_i$, $i = 1, 2, \ldots, n$.

The revenue under behavioral targeting not only depends on the values of the advertisers but may also depend on how those values are structured along the circle (e.g., whether the advertiser with the highest reference value is adjacent to the one with the second highest reference value). Notice that the second highest reference value determines the total revenue for the publisher under traditional advertising. To make a revenue comparison under the two different advertising strategies later, we fix the second highest reference value and compute the maximum and minimum revenue for the publisher under different distributions of advertiser’s values when using behavioral targeting.

**Proposition 1.** Under behavioral advertising, given the value of $z(2)$, (a) the value structure with $z(1) = z(2)/(1 - \frac{2}{n})$ and $z(3) = z(4) = \ldots = z(n) = z(2)$ generates the highest revenue among all possible value structures; (b) the value structure with $z(1) = z(2)$, $z(3) = z(4) = \ldots = z(n) = (1 - \frac{2}{n})z(2)$, and the two highest-value advertisers being $\frac{2}{n}$ distant to each other (if possible) generates the lowest revenue among all possible value structures.

Based on the monotonicity between the revenue and advertisers’ reference value in Lemma 3, part (a) is intuitive. Given the second highest reference value, the revenue is increasing in those lower reference values. The maximum can be reached only if those lower values reach as high as the second highest value (within the order constraint). The revenue is also increasing in the highest reference value, which is constrained at the value specified in the proposition because of the value relationship assumption.

The argument for part (b) is in the same spirit as the argument for part (a): given the second highest value, all the other values should be as low as possible so the highest value is equal to the second highest. One additional issue beyond part (a) is the relative position of the two advertisers with the highest reference value. The intuition for the result regarding the relative position is as follows.
First, the revenue from the base payment is lower when two highest-value advertisers are not adjacent than when they are. This is because when the two highest-value advertisers are adjacent to each other, each of them faces a competitor very similar to himself and the resulting face-to-face competition between them significantly raises the prices for the users they are competing for. In contrast, when the two highest-value advertisers are not adjacent, each of them directly competes with a low-value advertiser and pays a low price for the users they win. Second, when the two highest-value advertisers are \(\frac{2}{n}\) distant to each other, there are two cross-border effects (one from each user segment between two low-value advertisers with one neighbored by the highest-value advertiser), which is minimum among all value structures. To see that, if the two highest-value advertisers are adjacent, there are four cross-border effects (one from each user segment between the highest-value advertiser and low-value advertisers and one from each user segment between two low-value advertisers with one advertiser neighbored by the highest-value advertiser). Similarly, we can show that if the two highest-value advertisers are further than \(\frac{2}{n}\) away from each other, the number of cross-border effects will be four. Notice the total revenue is composed of the base payments and cross-border effects. Therefore, the revenue is the lowest one among all when the two highest-value advertisers are \(\frac{2}{n}\) distant to each other.

Facing the two advertising strategies, the publisher may choose the one generating higher revenue, which involves comparing \(\pi_T\) and \(\pi_B\). Such comparison can be done numerically, given all the related parameters. In general, either one may lead to a higher revenue than the other, depending on the heterogeneity of users’ preference, number of advertisers, and the structure of the advertiser’s value distributions. Next we derive conditions under which one advertising strategy is superior to the other for the publisher.

**Proposition 2.** (a) If and only if the number of advertisers is small \((n = 2)\), the publisher is (weakly) better off by using traditional advertising regardless of advertiser’s value structure. (b) If and only if the number of advertisers is large enough \((n \geq 6)\), the publisher is better off by using behavioral targeting regardless of advertiser’s value structure.
When only two advertisers compete for the advertising slot, we assume that advertiser 1 has a higher reference value than advertiser 2 without loss of generality. Under traditional advertising, advertiser 1 wins the auction and pays advertiser 2’s expected value of the users. The gray area in Figure 3a illustrates the expected payment for one user segment. The other user segment is symmetric. Under behavioral targeting, advertiser 1 wins users within distance $x_1$ and pays advertiser 2’s expected value of that user group, and advertiser 2 wins the rest and pays advertiser 1’s expected value of those users. The gray area in Figure 3b illustrates the expected payment. The main difference in the expected payments under the two advertising strategies lies in the expected payment for the users located at $[x_1, 1/2]$ from advertiser 1. We can see that for those users the expected payment under traditional advertising is greater than that under behavioral targeting. Therefore, the revenue under traditional advertising is generally higher that that under behavioral advertising if both advertisers have some positive market shares. Such an observation is consistent with the bundling literature: “bundling benefits a seller when it reduces valuation heterogeneity” (Geng et al., 2006). Traditional advertising is like “bundling” all users together, which reduces the advertisers’ valuation heterogeneity and increases the competition and thus the revenue. When one advertiser has no market share (i.e., $x_1 = 1/2$), the revenue under the two advertising strategies is the same.

When the number of advertisers is greater than 2, the advertiser with the highest reference value continues to win the auction and pays the expected value of the advertiser with the
second highest reference value for all users. For the users who are far from the second-highest advertiser, because of the decay, the expected value of the second-highest advertiser will be significantly lower than expected values of multiple advertisers who fit those users better (i.e., who are closer to such users on the circle). Under traditional advertising, the expected payment for those users is the expected value of those users to the second-highest advertiser. Under behavioral targeting, the expected payment for those users is the second highest expected value of those users to the advertisers. Therefore, the expected payment for those users under behavioral targeting would be higher, since the competitive effect dominates. As a result, we cannot extend the argument for the case with two advertisers to the case with more advertisers. At the same time, we can easily find cases in which the revenue under behavioral targeting is greater than that under traditional advertising when there are more than two advertisers. The two specific cases discussed later in this section serve as an illustration of such cases. All together, we can see that if and only if there are only two advertisers, the publisher is (weakly) better off by using traditional advertising rather than behavioral targeting.

One general insight that emerged from the above discussion is the efficiency loss associated with the allocation under traditional advertising. When he wins the auction under traditional advertising, the advertiser with the highest reference value wins all the users and pays the second-highest advertiser’s value of each user. In the presence of a number of other advertisers, assigning those users who are distant from the winner or have low probability to click on his advertisement might be inefficient in a sense that other advertisers may value those users more and result in higher revenue. We call this the propensity effect which positively affects the revenue under behavioral targeting. Once the number of advertisers becomes considerably large, the efficiency loss that results from this propensity effect becomes severe. As a result, the efficiency loss effect dominates the bundling or competition effect, and the publisher is better off when using behavioral targeting.

When the number of advertisers is neither large or small, either advertising strategy could
generate higher revenue than the other depending on the advertiser’s value structure and heterogeneity of users’ preference. The minimum and maximum revenue under behavioral targeting with value structures specified in the proposition, are cases where the superiority of one strategy over the other can be proven indisputably.

Next, we examine advertisers’ payoffs under different advertising strategies. Under traditional advertising all advertisers earn zero payoff since they have zero market share, except the advertiser with the highest reference value. Under behavioral targeting, in contrast, all advertisers may make positive payoffs unless an advertiser is totally dominated by its competitor (even over the advertiser’s most favorable user) such that the advertiser has zero market share. Therefore, all advertisers are (weakly) better off under behavioral advertising except the one with the highest reference value.

For the advertiser with the highest reference value (say advertiser $k$), the payoff under traditional advertising is its expected value of all users ($z_{(1)}E(q)$) net the expected payment (the second highest expected value of all users $z_{(2)}E(q)$); that is, $(z_{(1)} - z_{(2)})(1 - \frac{\gamma}{4})$. Under behavioral targeting, his payoff is the expected value of the users that the advertiser wins net the expected payment that the advertiser needs to pay (specified in Eq. (2) and Eq. (3)). His payoff in one user segment is the size of the dotted area of each subfigure in Figure 2. The total payoff comes from the two user segments that the winner belongs to and can be formulated as:

$$
\frac{1}{2} \left[ z_{(1)} - z_{k+1}(1 - \frac{1}{n}\gamma) \right] x_k + \frac{1}{2} \left[ z_{(1)} - z_{k-1}(1 - \frac{1}{n}\gamma) \right] \left( \frac{1}{n} - x_{k-1} \right) - \frac{(1 - \frac{\gamma}{n})^2(z_{k+1} - z_{k-1})^2}{2\gamma(z_{k+1} + z_{k-1})}
$$

(4)

where $x_k$ and $x_{k-1}$ are defined as before. In general, whether the highest value advertiser can be better off depends on the parameters such as the value of its neighbors and the number of advertisers. For some special cases, we can clearly see that it is possible for the highest-value advertiser to be better off under either advertising strategy. For example, if the highest-value advertiser has a comparable competitor (i.e., $z_{(1)} = z_{(2)}$, as in the special
case with the same value discussed later in this section), the advertiser earns zero payoff under traditional advertising but has positive payoff under behavioral targeting. Therefore, advertiser $k$ is better off under behavioral targeting. In another special case when there is one dominant player (discussed later in this section), we shall show that the highest-value advertiser could be worse off under behavioral targeting.

**Proposition 3.** (1) All advertisers are (weakly) better off under behavioral targeting except the advertiser with the highest reference value. (2) The advertiser with the highest reference value could be better off or worse off depending on the competitive situation (e.g., the second highest reference value, direct neighbors’ reference value, the number of advertisers, and $\gamma$). Specifically, if Eq. (4) \( (z_{(1)} - z_{(2)})(1 - \frac{\gamma}{4}) \), the advertiser is better off.

Whether the advertiser with the highest reference value is better off under behavioral targeting involves many factors and is complicated to derive. In the special case with one dominant player discussed later in this section, we shall demonstrate how the number of advertisers and the heterogeneity among user preferences affect the highest-value advertiser’s tradeoff. The highest-value advertiser’s neighbors also directly affect the competition facing the advertiser under behavioral targeting, and thus affect the advertiser’s payoff. In general, the higher the neighbors’ values are, the less payoff the highest-value advertiser can obtain (which is reflected in Figure 2, as the dotted areas shrink when the lines ending at $z_{i+1}$ move upward). Therefore, when neither neighbor is the one with the second highest value (so the payoff under traditional advertising remains the same), the highest-value advertiser is more likely to be better off under traditional advertising. When one neighbor is the advertiser with the second highest value, the change in its value affect the highest-value advertiser’s payoff under traditional advertising as well and its effect is two-edged.

We next compare the social welfare created under the two different advertising strategies. Social welfare is defined as the sum of the publisher’s payoff and advertisers’ payoffs, or the value created by users through advertising. In a sense, social welfare concerns the total “pie” created from the advertising slot, which would be an important consideration for the
publisher for a long-run business and/or when facing competition for advertisers.

Recall that in equilibrium advertisers bid their true value and that each user is allocated to the advertiser with the highest value under behavioral targeting. Such realized highest value of each user is the realized social welfare from each user, which is divided between the advertiser and the publisher by the auction rule. Therefore, the equilibrium allocation under the behavioral advertising is the way that creates the maximum social welfare. In contrast, the equilibrium allocation under traditional advertising cannot generate the maximum social welfare. To see that, for example, the users located at \( [x_1, 1/2] \) from advertiser 1 in Figure 3 are assigned to advertiser 1 under traditional advertising, although advertiser 2 values them more. So, the social welfare created under behavioral targeting is higher than that created under traditional advertising.

**Proposition 4.** The social welfare is maximized under behavioral targeting. The allocation under behavioral targeting generates more social welfare than that under traditional advertising.

We next explore two special cases to illustrate some results derived above. The special cases we consider resemble scenarios frequently observed in practice, and hence the results of the analysis also yield useful practical implications. We will focus on the benefits from behavioral targeting from the perspectives of the publisher and the highest-value advertiser, since the benefit of behavioral targeting in terms of social welfare has already been established.

### 4.2 Symmetric Advertisers

One special case is when all advertisers are symmetric, that is, they all have the same unit value (say 1). Such a case corresponds to a symmetric oligopoly competition case in the regular price competition setting. This case would be appropriate when competing advertisers have a comparable competitive advantage or similar marginal benefit from each click.
In this situation, when a is user located between advertiser $i$ and $i+1$, advertiser $i$ wins the advertising slot for this user if the user is more likely to click on the advertisement from advertiser $i$ than on the advertisement from advertiser $i+1$. When advertiser $i$ wins those users, the expected price that the advertiser pays for each user is

$$\frac{1}{2} \left[ (1 - \frac{1}{2n} \gamma) + (1 - \frac{1}{n} \gamma) \right] = 1 - \frac{3}{4n} \gamma$$

Since the expected price for each user is the same across advertisers, the above expression is also the total revenue for the publisher. Comparing the expected revenue under the traditional advertising ($\pi_T = (1 - \frac{1}{4})$) and under behavioral targeting, we can conclude the following.

**Proposition 5.** When there are only two advertisers, the publisher is better off under traditional advertising. When the number is three, the publisher is indifferent between traditional advertising and behavioral advertising. When the number of advertisers is greater than 3, the publisher is better off under behavioral targeting.

The intuition is again the balance between the competitive effect and propensity effect. When the number of advertisers is low, the auction under traditional advertising can leverage the competitive effect (from bundling the users) at the cost of efficiency because of the lower propensity of users to click on the advertisement. Notice that in this case, the competition is intensively fierce under traditional advertising such that the winner barely earns a profit. In contrast, under behavioral advertising, advertisers are differentiated from each other because of users’ heterogenous preferences and all of them earn some payoff.

**Proposition 6.** All advertisers are better off under behavioral advertising when they have the same unit value.
4.3 One Dominant Advertiser

Another special case we consider here is when one dominant advertiser has a significant competitive advantage in the market place while the other advertisers are of similar competitive advantage. In particular, we let \( z_1 > z_2 = z_3 = \ldots = z_n = 1 \). This class of value structure contains the one that leads to the highest revenue under behavioral targeting as discussed in Proposition 1.

The revenue under traditional advertising is the same as in the case with the same value (\( \pi_T = (1 - \gamma/4) \)). The revenue under behavioral targeting \( \pi_B \) can be formulated as

\[
2x_1(1 - \frac{1}{n}\gamma + \frac{x_1}{2}\gamma) + 2z_1(\frac{1}{n} - x_1) \left[ 1 - \frac{1}{n}\gamma + \frac{1}{2}\left(\frac{1}{n} - x_1\right)\gamma \right] + \frac{n - 2}{n}(1 - \frac{3}{4n}\gamma) + \frac{(1 - \frac{2}{n})^2(z_1 - 1)^2}{\gamma(z_1 + 1)^2}
\]

(5)

where \( x_1 \) is defined as (by Eq. (1))

\[
x_1 = \frac{(z_1 - 1) + \frac{1}{n}\gamma}{\gamma(z_1 + 1)}
\]

According to Lemma 3, the above revenue under behavioral advertising in this special case is greater than in the case with the same value (since advertiser 1’s reference value is higher in the former than in the latter while the other reference values are the same). Therefore, based on Proposition 5, we can conclude that when the number of advertisers is more than two, the publisher prefers behavioral targeting over traditional advertising. Meanwhile, according to Proposition 2, when there are only two advertisers, traditional advertising always generates higher revenue for the publisher compared to behavioral targeting.

**Proposition 7.** When there are only two advertisers, the publisher is better off under traditional advertising. When the number of advertisers is greater than 2, the publisher is better off under behavioral targeting.

Next, we examine advertisers’ payoff under the two advertising strategies. Under traditional targeting, advertiser 1’s payoff is \((z_1 - 1)(1 - \frac{\gamma}{4})\). Under behavioral advertising, by
Eq. (4), the advertiser’s payoff is

\[
\left[ z_1 - \left( 1 - \frac{\gamma}{n} \right) \right] x_1 = \frac{\left[ z_1 - \left( 1 - \frac{\gamma}{n} \right) \right]^2}{\gamma(z_1 + 1)}
\]

by noting that the advertisers’ neighbors have the same reference value so the last term in Eq. (4) is zero. Because of the symmetry, the payoff is simply twice as the size of the dotted area in Figure 2a.

**Proposition 8.** In the case with a dominant advertiser, the dominant advertiser (with the highest value) is (weakly) better off under behavioral advertising when there are only two advertisers regardless of its per-click value or \( \gamma \). When \( n \geq 3 \), if

\[
z_1 \leq \frac{(1 - \frac{z_1}{n}) - \sqrt{\gamma(1 - \frac{z_1}{4})[(1 - \frac{z_1}{4})^2 - (1 - \frac{\gamma}{2})^2]}}{(1 - \frac{\gamma}{2})^2}
\]

the advertiser is better off under behavioral targeting; otherwise, the advertiser is worse off.

The intuition is as follows. When the dominant advertiser has a low valuation (close to the other advertisers’), the face-to-face competition under traditional advertising leaves the dominant advertiser little profit margin. In contrast, under behavioral targeting with the relaxed competition, the dominant advertiser can reap its benefit from its favorable users. Therefore, the dominant advertiser is better off under behavioral targeting. When he has a high valuation, the dominant advertiser can grab the whole group of users under traditional advertising while maintaining a considerable profit margin. Under behavioral targeting, the dominant advertiser wins significantly less users as other advertisers have advantage over their favorable users. Therefore, the dominant advertiser is worse off under behavioral targeting.

Notice that the number of advertisers plays an important role in mediating the aforementioned tradeoff. When the number increases, the dominant advertiser’s payoff remains the same under traditional advertising, while his payoff decreases under behavioral targeting.
as more advertisers split the users. As a result, everything else being equal, the dominant advertiser is more likely to be better off under behavioral targeting when the number of advertisers is low. In the extreme case with only two advertisers, the dominant advertiser is always (weakly) better off under behavioral targeting. The result under the case with two advertisers can also be seen in Figure 3. The payoff under traditional advertising is the size of $A$ minus the size of $B$ in Figure 3a, whereas the payoff under behavioral targeting is the size of $A$ in Figure 3b. So the dominant advertiser is better off under behavioral targeting as long as the other advertiser has some market share. When the other advertiser has no market share under behavioral targeting, the dominant advertiser’s payoff is the same under both advertising strategies.

This special case is relevant not only because it captures many competitive scenarios in reality but also because it encompasses the case that generates the highest revenue possible under behavioral targeting for the publisher, as discussed in Proposition 1. We here use $(\pi_B - \pi_T)/\pi_T$ as the measure of the gain from behavioral targeting, and explore the maximum gain that the publisher could obtain by switching from traditional advertising to behavioral targeting. Notice that if all $z_i$’s are scaled in a same degree (e.g., to $\tau z_i$’s where $\tau$ is a constant), the gain is not affected. Therefore, as the discussion for Proposition 1, we can let $z(2)$ fixed without loss of generality, and the maximum gain can be obtained by solving the following optimization problem.

$$\max_{z(1):z(3)\ldots z(n)} \frac{\pi_B - \pi_T}{\pi_T}$$

which is equivalent to maximizing $\pi_B$ given any $z(2)$. From Proposition 1, we showed that for the value structure generating the highest revenue under behavioral targeting, we must have $z(2) = z(3) = \ldots = z(n)$ and $z(1)$ must be as big as possible.

**Proposition 9.** (a) When the number of bidders is two, the maximum gain is zero. When the number of bidders is greater than two, the maximum gain under the valuation relationship
assumption is \( \frac{\gamma}{4-\gamma} \left[ \frac{(n-2)(n-1)}{n^2} + \left( \frac{n-\gamma}{2n-\gamma} \right)^2 \right] \); The maximum gain without the valuation restriction is \( \frac{\gamma}{4-\gamma} \frac{(n-2)(n+1)}{n^2} \). (b) The maximum gain is increasing in the number of advertisers \( n \) and user heterogeneity \( \gamma \).

It is worth pointing out that the gain from behavioral targeting can be very significant. For example, if there is a high degree of heterogeneity among user preferences (i.e., \( \gamma \to 2 \)), then the gain can be one hundred percent when the number of advertisers becomes large (i.e., \( \frac{(n-2)(n-1)}{n^2} \to 1 \) when \( n \to \infty \)). In other words, the publisher’s revenue can be doubled by switching to behavioral targeting.

5 Conclusion

In this paper we analyzed the economic implications when an online publisher adopts behavioral targeting. We showed that when there are enough advertisers, the profit for the online publisher is higher than under traditional advertising. When there are only two advertisers, the online publisher is better off without behavioral targeting. Most advertisers would voluntarily choose to have their advertisements targeted, since they realize a higher payoff. Only the advertiser with the highest reference value could do worse under behavioral targeting.

Our auction mechanism used the second weighted unit price to compute payments for advertisers. This has the effect that payments are discriminatory, that is, an advertiser would need to pay more for a click-through from a less preferred user. In order to facilitate the implementation of our auction mechanism and eliminate these discriminatory payments, the publisher can compute the average expected payment from an advertiser and use this as a uniform price per click-through.
A Appendix

A.1 Proof of Lemmas 1 and 2

Proof. For any advertiser $i$, its unit value is $v_i$ and we generally denote its click-through rate as $t_i$ (the expected click-through rate under traditional advertising and the click-through rate for a specific user under behavioral targeting). We denote the highest proposed expected payment (per-click bid times click-through rate) from the rest of advertisers as $s_{-i}$. First, it is never optimal for the advertiser to bid higher than its true unit value. If the advertiser can win by bidding the true unit value, bidding higher leads to the same winning price and makes no difference; Otherwise, bidding higher than its true unit value results in two possible results: the advertiser continues to lose the auction and have a zero payoff, or it wins the auction but with a negative payoff since the advertiser have to pay a price above his value (i.e., $v_i < s_{-i}/t_i$). Second, it is never optimal for the advertiser to bid lower than its true unit value. If the advertiser can win with a bid less than $v_i$, it does so as well by bidding $v_i$ without paying more. If it does not win with a bid less than $v_i$, the advertiser will get zero payoff and is (weakly) better off by bidding the true unit value.

A.2 Proof of Proposition 1

Proof. By Lemma 3, the revenue is increasing in $z_i$. Given the order constraint $z_{(i)} \geq z_{(i+1)}$ and the value of $z_{(2)}$, (a) in the structure the highest revenue, we must have $z_{(3)} = z_{(4)} = \ldots = z_{(n)} = z_{(2)}$ and $z_{(1)}$ be as high as possible; (b) in the structure with the lowest revenue, we must have $z_{(1)} = z_{(2)}$, and $z_{(3)} = z_{(4)} = \ldots = z_{(n)}$ being as low as possible. Due to the value relationship $z_{(n)} \geq z_{(1)}(1 - \frac{1}{n})$, when the condition binds, the revenue reaches the maximum and the minimum respectively in (a) and (b).

For (b), we next check how the relative location of the two advertisers with $z_{(2)}$ affects the revenue. When $n = 2$ or $3$, there is a unique relative distribution of those advertisers' values. When $n \geq 4$, we first examine the possible revenue for a user segment $i|\bar{i}+1$. For
a user segment with \( z_i = z(2) \) and \( z_{i+1} = z(2) \) or user segment \( z(2) | z(2) \), the revenue is from a base payment (in Table 1) by Eq. (2) with \( x_i = 1/(2n) \). For user segment \( z(2) | z(n) \), if

\[
\begin{array}{c|c|c|c}
\text{User Segment} & z(2) | z(2) & z(2) | z(n) & z(n) | z(n) \\
\hline
\text{Base Payment} & \frac{1}{n}(1 - \frac{3\gamma}{4n})z(2) & \frac{1}{n}(1 - \frac{2}{n})(1 - \frac{\gamma}{2n})z(2) & \frac{1}{n}(1 - \frac{2}{n})(1 - \frac{3\gamma}{4n})z(2) \\
\end{array}
\]

\( z(2) \)'s other neighbor is of \( z(n) \), the revenue is from a base payment (in Table 1) by Eq. (2) with \( x_i = 1/n \); if \( z(2) \)'s other neighbor is of \( z(2) \), the revenue is the base payment plus the cross-border effect. By Eq. (3), the cross-border effect is

\[
\frac{(1 - \frac{2}{n})^2(z(2) - z(n))^2}{2\gamma(z(2) + z(n))} = \frac{(1 - \frac{2}{n})^2(z(n))^2}{2\gamma(2 - \frac{1}{n})} = \gamma(n - \gamma)^2 \frac{z(2)}{2n^2(2n - \gamma)} \tag{6}
\]

For user segment \( z(n) | z(n) \), if their neighbors are of \( z(n) \), the revenue is from a base payment (in Table 1) by Eq. (2) with \( x_i = 1/(2n) \); if one neighbor is of \( z(2) \) and the other is of \( z(n) \), the revenue is the base payment plus one piece of the cross-border effect defined in Eq. (6). If both neighbors are of \( z(2) \), the revenue is the base payment plus two pieces of the cross-border effect (with one piece at each border).

Notice the base payment from each user segment \( i | (i+1) \) is solely determined by the value \( z_i \) and \( z_{i+1} \) (while the cross-border effect depends on their neighbors’ values). Therefore, the revenue from base payments (excluding the cross-border effects) can differ only when the two highest-value advertisers are adjacent and when they are not. When \( n \geq 4 \), in the former, the revenue consists of that from one \( z(2) | z(2) \) segment, that from two \( z(2) | z(n) \) segments, and that from \((n - 3) \ z(n) | z(n) \) segments; in the latter, the revenue consists of that from four \( z(2) | z(n) \) segments and that from \((n - 4) \ z(n) | z(n) \) segments. The difference in the above revenues is the revenue from one \( z(2) | z(2) \) plus that from one \( z(n) | z(n) \) minus that from two \( z(2) | z(n) \); that
\[
\frac{1}{n}(1 - \frac{3\gamma}{4n})z_{(2)} + \frac{1}{n}(1 - \frac{\gamma}{n})(1 - \frac{3\gamma}{4n})z_{(2)} - \frac{2}{n}(1 - \frac{\gamma}{n})(1 - \frac{\gamma}{2n})z_{(2)}
\]

\[
= \frac{1}{n}(1 - \frac{3\gamma}{4n})z_{(2)} - \frac{1}{n}(1 - \frac{\gamma}{n})(1 - \frac{\gamma}{4n})z_{(2)} = \frac{1}{n}\frac{\gamma}{2n}(1 - \frac{\gamma}{2n})z_{(2)} > 0
\]

which indicates the revenue from base payments is greater when the highest-value advertisers are adjacent.

If the two highest-value advertisers are adjacent, the total revenue has four pieces of cross-border effect in addition to the revenue from the base payment. When \( n = 4 \), the four pieces consist of two from the two \( z_{(2)}|z_{(n)} \) segments (because in each segment \( z_{(2)} \) neighbors the other \( z_{(2)} \)) and two from the \( z_{(n)}|z_{(n)} \) segment (because both \( z_{(n)} \) neighbor \( z_{(2)} \)). When \( n \geq 5 \), the four pieces consist of two from the two \( z_{(2)}|z_{(n)} \) segments and two from those two \( z_{(n)}|z_{(n)} \) segments with one \( z_{(n)} \) in each segment neighbored by \( z_{(2)} \).

If the two highest-value advertisers are not adjacent, the total revenue contains no cross-border effect when \( n = 4 \). Therefore, the total revenue (including the revenue from the base payment and the cross-border effect) is lower when the two highest-value advertisers are not adjacent than when they are. When \( n \geq 5 \), if the two highest-value advertisers are \( \frac{2}{n} \) distant to each other (i.e., there is one lowest-value advertiser in between), the total revenue contains two pieces of cross-border effect either from the two \( z_{(n)}|z_{(n)} \) segments with one \( z_{(n)} \) neighbored by \( z_{(2)} \) (if \( n > 5 \)) or from the \( z_{(n)}|z_{(n)} \) segment with both \( z_{(n)} \) neighbored by \( z_{(2)} \) (if \( n = 5 \)). If there are at least two lowest-value advertisers between the two highest-value advertisers, the total revenue contains four pieces of cross-border effect. This is because around each arc connecting the two \( z_{(2)} \) we should have either two \( z_{(n)}|z_{(n)} \) segments with one \( z_{(n)} \) neighbored by \( z_{(2)} \) in each segment or one \( z_{(n)}|z_{(n)} \) segment with both \( z_{(n)} \) neighbored by \( z_{(2)} \). Either case leads to two pieces of cross-border effects with one arc and we have two different arcs. Therefore, when \( n \geq 5 \), the structure with the two highest-value advertisers being \( \frac{2}{n} \) distant to each other generates the least total revenue (with lowest revenue from
the base payment and least number of cross-border effect).

A.3 Proof of Proposition 2

Proof. (a) When \( z_1 = z_2 = ... = z_n \), the revenue under behavioral targeting consists of the base payments from the \( n \) \( z_2 \mid z_2 \) segments, which is \( (1 - \frac{3\gamma}{4n})z_2 \) according to Table 1. Comparing it with \( \pi_T = (1 - \frac{\gamma}{4})z_2 \), we can conclude \( \pi_B \geq \pi_T \) when \( n \geq 3 \). Notice that any value structure with \( z_1 > z_2 \) results in a higher revenue under behavioral targeting (by Lemma 3) and the same revenue under traditional advertising. Therefore, when \( n \geq 3 \), the revenue under traditional advertising may be less than that under behavioral targeting.

When \( n = 2 \), given \( z_2 \), the maximum revenue under behavioral targeting is reached when \( z_1 \) is large enough to win all users. In that case, by Eq.(2) with \( x_i = 1/2 \), we have \( \pi_B = (1 - \frac{\gamma}{4})z_2 \), which is the same as \( \pi_T \). Therefore, the revenue under traditional advertising is generally higher that that under behavioral targeting when \( n = 2 \).

(b) When \( n \geq 5 \), according to the proof of Proposition 1, the least revenue consists of the base payment (from four \( z_2 \mid z_n \) segments and \( n - 4 \) \( z_n \mid z_n \) segments) and two pieces of cross-border effect:

\[
\frac{4}{n}(1 - \frac{\gamma}{n})(1 - \frac{\gamma}{2n})z_2 + \frac{n - 4}{n}(1 - \frac{\gamma}{n})(1 - \frac{3\gamma}{4n})z_2 + \frac{\gamma(n - \gamma)^2}{n^3(2n - \gamma)}z_2
\]

The difference between the above \( \pi_B \) and \( \pi_T = (1 - \frac{\gamma}{4})z_2 \) is

\[
\left[ (1 - \frac{\gamma}{n})(1 - \frac{3n - 4}{4n^2} - \gamma) + \frac{\gamma(n - \gamma)^2}{n^3(2n - \gamma)} - (1 - \frac{\gamma}{4}) \right]z_2
\]

\[
= \left[ n^2(n - 4) - (3n - 4)(n - \gamma) + \frac{4(n - \gamma)^2}{2n - \gamma} \right] \frac{\gamma z_2(2)}{4n^3}
\]

The term in the above square bracket is increasing in \( \gamma \) by noticing the first-order derivative

\[
3n - 4 - \frac{8(n - \gamma)}{2n - \gamma} + \frac{4(n - \gamma)^2}{(2n - \gamma)^2} > 3n - 4 - \frac{8(n - \gamma)}{2n - \gamma} = 3n - 8 + \frac{4\gamma}{2n - \gamma} > 0
\]
and thus is greater than its value at zero  \( n^2(n - 4) - (3n - 4)n + 2n = n(n - 6)(n - 1) \) (which is nonnegative if \( n \geq 6 \)). Therefore, the above difference is positive and behavioral advertising generates higher revenue if \( n \geq 6 \). If \( n = 5 \), the above difference is negative for any \( \gamma \) by noticing the value of the term in the square bracket at \( \gamma = 2 \) is \(-3.5\).

When \( n = 4 \), the least revenue under behavioral advertising is \((1 - \frac{\gamma}{4})(1 - \frac{\gamma}{3})z_2\) (from four \( z_2 \mid z(n) \) segments), which is less than \( \pi_T = (1 - \frac{\gamma}{4})z_2 \) regardless of \( \gamma \). When \( n = 3 \), the least revenue consists of the base payment (from two \( z_2 \mid z(n) \) segments and one \( z_2 \mid z(2) \) segment) and two pieces of cross-border effect:

\[
\frac{1}{3}(1 - \frac{\gamma}{4})z_2 + \frac{2}{3}(1 - \frac{\gamma}{3})(1 - \frac{\gamma}{6})z_2 + \frac{\gamma(3 - \gamma)^2}{27(6 - \gamma)}z_2 = (1 - \frac{\gamma}{4})z_2 + \left[ -\frac{1}{6} + \frac{\gamma}{27} + \frac{(1 - \gamma/3)^2}{3(6 - \gamma)} \right] \gamma z_2
\]

which is less than \( \pi_T \) regardless of \( \gamma \) because the term in the square bracket is negative. When \( n = 2 \), we have shown in part (a) that the revenue under behavioral targeting is always lower than that under traditional advertising.

All together, we can conclude that if \( n < 6 \), the revenue under behavioral targeting may be less than that under traditional advertising. If \( n \geq 6 \), the least revenue under behavioral targeting is greater than the revenue under traditional advertising. Therefore, if and only if \( n \geq 6 \), the publisher is better off by using behavioral targeting. \( \square \)

### A.4 Proof of Proposition 8

*Proof.* The difference in the payoff under behavioral targeting and under traditional advertising is

\[
\Delta(z_1) = \frac{[z_1 - (1 - \frac{\gamma}{4})^2]}{\gamma(z_1 + 1)} - (z_1 - 1)(1 - \frac{\gamma}{4})
\]

\[
= \frac{1}{\gamma(z_1 + 1)} \left[ (1 - \frac{\gamma}{2})^2 z_1^2 - 2(1 - \frac{\gamma}{n})z_1 + (1 - \frac{\gamma}{n})^2 + \gamma(1 - \frac{\gamma}{4}) \right]
\]

\[
\equiv \frac{1}{\gamma(z_1 + 1)} \left[ A z_1^2 + B z_1 + C \right]
\]
We notice that

\[ B^2 - 4AC = 4(1 - \frac{\gamma}{n})^2 - 4(1 - \frac{\gamma}{2})^2 \left[ (1 - \frac{\gamma}{n})^2 + \gamma(1 - \frac{\gamma}{4}) \right] \]

\[ = 4\gamma(1 - \frac{\gamma}{4}) \left[ (1 - \frac{\gamma}{n})^2 - (1 - \frac{\gamma}{2})^2 \right] \]

When \( n = 2 \), \( B^2 - 4AC = 0 \). Therefore, \( \Delta(z_1) \geq 0 \).

When \( n \geq 3 \), \( B^2 - 4AC > 0 \) and \( \Delta(z_1) = 0 \) has two roots:

\[ z' = \frac{(1 - \frac{\gamma}{n}) - \sqrt{\gamma(1 - \frac{\gamma}{4})[(1 - \frac{\gamma}{n})^2 - (1 - \frac{\gamma}{2})^2]}}{(1 - \frac{\gamma}{2})^2} \]

and

\[ z'' = \frac{(1 - \frac{\gamma}{n}) + \sqrt{\gamma(1 - \frac{\gamma}{4})[(1 - \frac{\gamma}{n})^2 - (1 - \frac{\gamma}{2})^2]}}{(1 - \frac{\gamma}{2})^2} \]

Recall that \( \Delta(1) > 0 \) according the result under the special case with the same value. Therefore, we have either \( z'_1 > 1 \) or \( z''_1 < 1 \). Since

\[ z'_1 + z''_1 = \frac{-B}{A} = \frac{2(1 - \frac{\gamma}{n})}{(1 - \frac{\gamma}{2})^2} > 2 \]

it must be that \( z'_1 > 1 \). Under the value relationship assumption, we have \( z_1 \leq 1/(1 - \frac{\gamma}{n}) \). Since \( 1/(1 - \frac{\gamma}{n}) < (1 - \frac{\gamma}{n})/(1 - \frac{\gamma}{2})^2 < z''_1 \), \( z_1 < z''_1 \). Therefore, if \( z_1 < z'_1 \), \( \Delta(z_1) > 0 \) and thus the advertiser is better off under behavioral targeting; otherwise, it is worse off. Notice that it is possible \( z_1 > z'_1 \) when \( z_1 \) is large enough (e.g., close to \( 1/(1 - \frac{\gamma}{n}) \)), because

\[ \frac{1}{1 - \frac{\gamma}{n}} - z'_1 = \frac{1}{(1 - \frac{\gamma}{n})(1 - \frac{\gamma}{2})^2} \left[ (1 - \frac{\gamma}{2})^2 - (1 - \frac{\gamma}{n})^2 + (1 - \frac{\gamma}{n}) \sqrt{\gamma(1 - \frac{\gamma}{4})[(1 - \frac{\gamma}{n})^2 - (1 - \frac{\gamma}{2})^2]} \right] \]

\[ = \frac{\sqrt{(1 - \frac{\gamma}{n})^2 - (1 - \frac{\gamma}{2})^2}}{(1 - \frac{\gamma}{n})(1 - \frac{\gamma}{2})^2} \left[ (1 - \frac{\gamma}{n}) \sqrt{\gamma(1 - \frac{\gamma}{4})} - \sqrt{(1 - \frac{\gamma}{n})^2 - (1 - \frac{\gamma}{2})^2} \right] \]

\[ = \frac{\sqrt{(1 - \frac{\gamma}{n})^2 - (1 - \frac{\gamma}{2})^2}}{(1 - \frac{\gamma}{n})(1 - \frac{\gamma}{2})^2} \frac{(1 - \frac{\gamma}{2})^2 [1 - (1 - \frac{\gamma}{n})^2]}{(1 - \frac{\gamma}{n}) \sqrt{\gamma(1 - \frac{\gamma}{4})} + \sqrt{(1 - \frac{\gamma}{n})^2 - (1 - \frac{\gamma}{2})^2}} > 0 \]
A.5 Proof of Proposition 9

Proof. (a) When $n = 2$, according to Proposition 2, the publisher is (weakly) better off under traditional advertising. Meanwhile, when the lower-value advertiser gets no market share under behavioral targeting, two advertising strategies could lead to the same revenue for the publisher. Therefore, if $n = 2$, the maximum gain is zero.

When $n > 2$, under the value relationship assumption, $\pi_B$ is maximized when $z(2) = z(3) = \ldots = z(n)$ and $z(1) = z(2)/(1 - \frac{\gamma}{n})$ by Proposition 1. Without loss of generality, we let $z_1 = z(1)$ and normalize $z(2) = 1$. By substituting $z(1)$ and $x_1 = 1/n$ into Eq. (5), we can obtain the maximum $\pi_B$ and thus calculate the maximum gain as

$$\frac{\pi_B}{\pi_T} - 1 = \frac{\frac{n-2}{n} (1 - \frac{3\gamma}{4n}) + \frac{2}{n} (1 - \frac{\gamma}{2n}) + \left(\frac{n-\gamma}{(2n-\gamma)n}\right)^2 \gamma}{1 - \frac{\gamma}{4}} - 1 = \frac{\frac{\gamma}{4} - \frac{2}{n} (\frac{n-2}{4} + \frac{1}{n}) + \left(\frac{n-\gamma}{(2n-\gamma)n}\right)^2 \gamma}{1 - \frac{\gamma}{4}}$$

$$= \frac{(1 - \frac{3n-2}{n^2}) \gamma + \left(\frac{n-\gamma}{(2n-\gamma)n}\right)^2 \gamma}{4 - \gamma} = \frac{\gamma}{4 - \gamma} \left[\frac{(n-2)(n-1)}{n^2} + \left(\frac{n-\gamma}{(2n-\gamma)n}\right)^2\right]$$

Without the restriction on the value, by Lemma 3, $\pi_B$ is maximized when $z(2) = z(3) = \ldots = z(n)$ and $z(1)$ is large enough to win the whole users. We again let $z_1 = z(1)$ and normalize $z(2) = 1$. For each of the two segments with one end at $z(1)$ (i.e., segments $n|1$ and $1|2$), the expected revenue is the lower-value advertiser’s expected value of all the users in the segment, which is $\frac{1}{n}(1 - \frac{\gamma}{2n})$. For each of other $n - 2$ segments (e.g., $i|i + 1$), the expected revenue is advertiser $i$’s expected value of the half of the users who are closer to it than to advertiser $i$ and advertiser $i + 1$’s expected value of the other half, which is $\frac{1}{n}(1 - \frac{\gamma}{4n})$. We can thus obtain the maximum $\pi_B$ and calculate the maximum gain as

$$\frac{\pi_B}{\pi_T} - 1 = \frac{\frac{n-2}{n} (1 - \frac{\gamma}{4n}) + \frac{2}{n} (1 - \frac{\gamma}{2n})}{1 - \frac{\gamma}{4}} - 1 = \frac{\frac{\gamma}{4} - \frac{2}{n} (\frac{n-2}{4} + \frac{1}{n})}{1 - \frac{\gamma}{4}}$$

$$= \frac{\gamma (1 - \frac{n+2}{n^2})}{4 - \gamma} = \frac{\gamma (n-2)(n+1)}{n^2}$$

(b) The maximum gain without the restriction on the value $(\frac{\gamma}{4-\gamma} \frac{(n-2)(n+1)}{n^2})$ is clearly
increasing in $\gamma$. The gain is also increasing in $n$ as its first-order derivative with respect to $n$ is positive; that is, $\frac{\gamma}{4-\gamma}(\frac{1}{n^2} + \frac{4}{n^3}) > 0$.

We next show the monotonicity in $n$ and $\gamma$ of the maximum gain under the value assumption. We check the first-order derivative of the term in the square bracket in the right-hand side of Eq. (7) with respect to $n$:

$$\frac{3n - 4}{n^3} - \frac{2(n - \gamma)}{(2n - \gamma)n} \frac{(2n^2 - 4n\gamma + \gamma^2)}{(2n - \gamma)^2 n^2} > \frac{3n - 4}{n^3} - \frac{2(2n^2 - 4n\gamma + \gamma^2)}{(2n - \gamma)^2 n^2} = \frac{(2n - 4)(2n - \gamma)^2 + (4n^2\gamma - n\gamma^2)}{(2n - \gamma)^2 n^3} > 0$$

which indicates the maximum gain is increasing in $n$.

To see that the maximum gain is increasing in $\gamma$, we next show that its first-order derivative with respect to $\gamma$ is positive:

$$\frac{4}{(4-\gamma)^2} \left[ \frac{(n-2)(n-1)}{n^2} + \frac{(n-\gamma)^2}{(2n-\gamma)n} \right] - 2 \frac{\gamma}{4-\gamma} \frac{n-\gamma}{(2n-\gamma)n} \frac{n^2}{(2n-\gamma)^2 n^2} = \frac{2(n-\gamma)}{(4-\gamma)^2(2n-\gamma)^3 n^2} \left[ \frac{2(n-2)(n-1)(2n-\gamma)^3}{n-\gamma} + 2(n-\gamma)(2n-\gamma) - (4-\gamma)n\gamma \right] > \frac{2(n-\gamma)}{(4-\gamma)^2(2n-\gamma)^3 n^2} \left[ 4(2n-\gamma)^2 + 2(n-\gamma)(2n-\gamma) - (4-\gamma)n\gamma \right]$$

Notice that the term in the above square bracket can be simplified into $20n^2 - 26n\gamma + 6\gamma^2 + n\gamma^2$, which is positive. Therefore, the above first-order derivative is positive.

References


