Microtargeting and Information Asymmetry in Online Advertising

Wee Chaimanowong, Francisco Cisternas, Stylianos Despotakis*

November 11, 2024

Abstract

In online advertising auctions, advertisers bid on ad impressions by using consumer data to target users effectively. However, disparities in data access and types among advertisers create information asymmetries that influence auction outcomes and publishers' revenues. This paper studies the impact of such asymmetries by developing a theoretical model that incorporates three key elements: (1) differentiation between types of consumer data—recognizing that data varies by source and characteristics; (2) information asymmetry among advertisers—acknowledging that not all advertisers have equal access to all consumer data; and (3) possible correlations in advertisers' valuations—understanding that certain data can affect advertisers' valuations in correlated ways.

Our findings reveal that under specific conditions—when advertisers' valuations are positively correlated based on certain consumer data and when information asymmetry exists among them—publishers can improve both their revenue and the auction's conversion rates by limiting data access and disabling microtargeting. Additionally, we show that when advertisers' valuations are independent, information asymmetry can be advantageous for publishers, suggesting that selectively allowing microtargeting can be beneficial. Interestingly, both informed and uninformed advertisers may, in some cases, gain when their competitors acquire more information.

Keywords: Online advertising; Targeting; Information asymmetry; Conversion rate; Ad revenue; Auction theory.

*The authors are listed in alphabetical order. Wee Chaimanowong: weechaimanowong@cuhk.edu.hk (CUHK Business School, The Chinese University of Hong Kong), Francisco Cisternas: fcisternas@cuhk.edu.hk (CUHK Business School, The Chinese University of Hong Kong), Stylianos Despotakis (corresponding author): sdespota@cityu.edu.hk (City University of Hong Kong). The authors thank Amin Sayedi, Chenxi Liao, and various seminar participants for their helpful feedback and comments. Despotakis acknowledges financial support from the Research Grants Council of Hong Kong (Project No. CityU 21509121).

1 Introduction

Online advertising is one of the most important revenue generation mechanisms for publishers and a vital channel for advertisers seeking to reach potential consumers. Central to the success of online advertising is the ability to target advertisements to users based on rich consumer data. This targeting is mainly done through real-time bidding auctions, where advertisers bid for ad impressions by using available information about the user behind each impression. Improvements in data collection technologies have allowed advertisers to gather large amounts of consumer data, allowing for highly personalized advertising experiences. However, this field is characterized by significant heterogeneity in the types of data available and asymmetries in data access among advertisers, which importantly affect market outcomes.

Consumer data in online advertising is complex, made up of various types such as behavioral data—capturing a user's past online activities, preferences, and purchasing history—and contextual data—relating to the content being viewed or the device being used. These data types originate from different sources, including first-party data collected directly by the publisher, third-party data aggregated by data brokers, and proprietary data held by individual advertisers. Each type of data offers distinct insights and has unique characteristics that influence its utility in targeting and valuation.

Despite the apparent wealth of data, not all advertisers have equal access to all consumer information. Large advertisers or those integrated with major ad exchanges may have advanced tools and special access to detailed user data, enabling them to engage in microtargeting. In contrast, smaller advertisers or new entrants might rely solely on publicly available or basic contextual data. This difference creates an information asymmetry in the advertising auction, where informed advertisers can adjust their bids more effectively than their less-informed counterparts. Such asymmetry can lead to inefficiencies in the auction outcomes, affecting both the competitiveness of the bidding process and the overall welfare of the market participants.

From the publisher's perspective, the control over data sharing presents strategic choices with significant implications. Publishers must decide the extent to which they share their proprietary data with advertisers. By enabling microtargeting through the sharing of detailed user data, publishers might enhance the value of their ad impressions, potentially increasing revenue from advertisers willing to pay premiums for precise targeting. However, excessive sharing of data can also widen asymmetries and might lead to reduced competition among advertisers, ultimately harming the publisher's revenue in the long run. Conversely, by limiting data access—such as by disabling third-party cookies or restricting data sharing—publishers might create a more level playing field among advertisers, which could enhance competition and lead to different revenue dynamics.

Moreover, the impact of consumer data on advertisers' valuations is not uniform and can be correlated across advertisers in complex ways. For instance, information indicating a user's high income level may simultaneously increase the valuations of luxury goods advertisers but might be less relevant to discount retailers. Similarly, a user's intent to purchase a car is highly valuable to car advertisers but holds little significance for unrelated industries. These correlations in valuations based on consumer data influence how advertisers perceive the value of an impression and, consequently, how they bid in auctions.

Understanding the interaction between the types of consumer data, the asymmetry in data access among advertisers, and the correlations in valuations is important for analyzing online advertising auctions. Despite the practical significance, there is a lack of theoretical frameworks that integrate all these elements to evaluate their collective impact on market outcomes. Addressing this gap is essential for publishers aiming to form data-sharing strategies that optimize their objectives, whether it be maximizing revenue or increasing the conversion rates of auctions, and for policymakers concerned with promoting fair competition in digital markets.

This paper investigates the effects of information asymmetry arising from differential access to diverse types of consumer data in online advertising auctions. We develop and study a model that incorporates three key elements: (1) differentiation between types of data, (2) information asymmetry between advertisers, and (3) correlation among advertisers' valuations.

Differentiation between types of data. Recognizing that consumer data is heterogeneous, we model multiple dimensions of data—such as behavioral and contextual data—each contributing differently to advertisers' valuations. This approach differs from traditional models that consider consumer data as a single homogeneous variable, allowing for a nuanced analysis of how distinct data types influence auction outcomes.

Information asymmetry between advertisers. Our model incorporates scenarios where advertisers have varying levels of access to consumer data. By distinguishing between informed advertisers, who possess comprehensive data, and uninformed advertisers, with limited data access, we analyze how this asymmetry affects bidding strategies and market efficiency.

Correlation among advertisers' valuations. We consider the possibility that certain consumer data impacts multiple advertisers' valuations in correlated ways. This aspect acknowledges that some data attributes may simultaneously increase the attractiveness of an impression to several advertisers, introducing complex dynamics into the bidding process.

Research Questions

The value of microtargeting in online advertising has been extensively debated within both the advertising industry and the academic literature. Proponents argue that microtargeting improves the precision of advertising campaigns, allowing advertisers to reach consumers who are most likely to be interested in their products or services. This precision is believed to increase the value of ad impressions, thereby increasing publishers' revenues. For example, Google researchers observed a 52% reduction in publishers' revenue when third-party cookies were disabled, underscoring the importance of microtargeting (Ravichandran and Korula, 2019).

Critics, however, argue that microtargeting can lead to negative market outcomes for publishers. One critique centers on the role of intermediaries like ad exchanges, which often charge a significant portion of the resulting revenue for their microtargeting services (Hsiao, 2020). Without the need for microtargeting, publishers could potentially reduce costs by eliminating these intermediaries. A second critique is that enhanced information among advertisers can lead to reduced competition and lower prices due to diminished market thickness (Levin and Milgrom, 2010). By withholding information and limiting microtargeting, publishers might strengthen competition and achieve higher revenues.

While these arguments focus on the revenue aspect, another critical metric for evaluating the effectiveness of advertising is the conversion rate. The prevailing view is that microtargeting should improve conversion rates by serving more relevant ads to consumers. However, there is anecdotal evidence suggesting this may not always be the case. For example, in an experiment with various advertisers, the Dutch public broadcaster NPO found that conversion rates did not decline and sometimes even improved when ads were targeted using only partial data (Snelders et al., 2020). And for this to happen, revenue does not have to be sacrificed since NPO also saw a significant increase in advertising revenue after discontinuing the use of tracking cookies (Edelman, 2020). Similarly, The New York Times experienced analogous results when they stopped offering behavioral targeting on their pages (Davies, 2019).

The divergence in the different perspectives highlights several key research questions. **Under what conditions does microtargeting improve or diminish publisher revenues and conversion rates?** While microtargeting can increase the value of certain impressions, its overall impact is not clearly understood, especially when considering the heterogeneity of data types and advertisers' access to information. **How does information asymmetry among advertisers affect auction outcomes?** The role of asymmetric information in shaping bidding strategies and market efficiency requires further examination, particularly in contexts where data access varies significantly among advertisers. Lastly, **what are the implications of correlations in advertisers' valuations based on consumer data?** Understanding how different types of consumer data influence correlated valuations among advertisers is essential for assessing the impact of data-sharing policies on auction dynamics.

Contributions

This paper addresses these questions by developing a theoretical model that integrates the differentiation of consumer data types, information asymmetry among advertisers, and the correlation of advertisers' valuations. Our contributions are threefold:

- **We establish conditions under which limiting microtargeting can simultaneously improve publisher revenue and auction conversion rates.** Specifically, we demonstrate that when there is a positive correlation among advertisers' valuations linked to certain consumer data and when not all advertisers have access to this data, disabling microtargeting can lead to more competitive bidding.[1] This increased competition enhances both the publisher's

[1]This result is in contrast to the *linkage principle* (Milgrom and Weber, 1982) that suggests that less-informed advertisers will underbid in the auction for the impression to avoid a phenomenon similar to the winner's curse (McAfee and McMillan, 1987), and therefore the publisher should commit to reveal all available information to increase the bids and the revenue. Our result differs due to the data asymmetry that exists among advertisers.

revenue and the efficiency of the auction process.

- **We analyze scenarios where information asymmetry benefits the publisher.** Our findings reveal that in some cases, allowing only certain advertisers to engage in microtargeting (instead of everyone or no one) can result in higher revenues for the publisher.

- **We explore the counterintuitive outcomes where advertisers may benefit from their competitors gaining more information.** In situations where advertisers' valuations are independent, an increase in competitors' information can benefit all advertisers through improved allocation efficiency.

By addressing these aspects, our study offers new insights into the strategic role of consumer data in online advertising auctions. We provide a nuanced understanding of how publishers can decide on data-sharing policies to optimize revenue and market outcomes. Additionally, our analysis contributes to the broader discussion on data privacy and competition policy, highlighting the complex interactions between data access, market efficiency, and welfare.

2 Related Literature

This research contributes to the growing literature on targeted advertising and online advertising auctions. Below, we discuss in more detail some related theory papers on targeted advertising.

From the advertisers' perspective, improving firms' ability to target consumers typically improves revenues, but in some cases it can have negative effects. For example, Chen et al. (2001) study the effects of imperfect targetability on prices for different segments of consumers. Interestingly, they found that improving the targetability of a firm can sometimes benefit both the firm and its competitor. Iyer et al. (2005) describe a model of competing firms who can target different segments of consumers with advertising and show that targeted advertising will improve the firms' profits and, moreover, it can sometimes be more valuable than targeted pricing. Bergemann and Bonatti (2011) show that better targeting causes an increase in the number of consumer-product matches, but prices of ads change non-monotonically in the targeting capacity. Brahim et al. (2011) study a model with two firms competing in prices and targeted advertising. They show that firms' profits can be lower with targeted relative to random advertising. Esteves and Resende (2016) study how targeted advertising can be used by competing firms to price discriminate different segments of consumers. Zhang and Katona (2012) study how contextual advertising affects product market competition. Johnson (2013) considers targeted advertising in combination with advertising avoidance technology. He shows that targeting will increase firms' profits, but it can make consumers worse off. Hummel and McAfee (2016) study how the number of bidders in an auction affects a seller's revenue under two different settings (bundling vs. targeting). A difference in our model is that bidders' valuations need not be independent. Despotakis and Yu (2022) study a multidimensional targeting model and show that sometimes the use of multiple dimensions of data to target consumers can have negative effects for a firm.

De Corniere and De Nijs (2016) show that when a platform chooses to reveal the information it has about a consumer to advertisers, the advertisers will set higher prices in anticipation of a better matching. This will benefit the advertisers and the platform. Our setting differs in that revealing information can actually worsen the matching between the advertisers and the consumer, resulting in a lower social welfare. This is because in addition to the full disclosure or non-disclosure of information, we also consider the case where not all advertisers have access to the same information about the consumer, and this asymmetry plays an important role in our model. Shen and Miguel Villas-Boas (2018) study advertising based on the past purchase behavior of consumers and examine how it affects product prices for a monopolist. Rafieian and Yoganarasimhan (2021) show that the revenues of ad-networks can increase when they allow users to preserve their privacy. This is because more precise targeting can thin out the market and soften competition, in a similar fashion to Levin and Milgrom (2010). However, when this happens, the targeting becomes less efficient. In our model, we can replicate this effect for the case of independent valuations between the advertisers, but for the case of dependent valuations, we see that revenue and efficiency can increase or decrease together.

Ada et al. (2022) study the impact of providing ad context information in ad exchanges and find that in most cases ad exchanges can boost publishers' revenues by sharing context information with ad buyers. (Shin and Shin, 2022) demonstrate that irrelevant advertising can stem from strategic decisions within the ad agency-advertiser relationship, rather than simply technological imperfections. The study also explores how contractual restrictions can lead to inefficiencies in ad delivery, and suggests that the prevalence of irrelevant ads may decrease, but not disappear, as the number of impressions available in the market increases. Choi and Sayedi (2023b) investigate the effects of private exchanges on the display advertising market, finding that while private exchanges offer higher quality impressions compared to open exchanges, they can also create information asymmetry among advertisers, which can hurt publisher's revenue. Choi and Sayedi (2023a) examine the effects of ad agencies on the online advertising market, revealing that publishers face a trade-off when deciding whether to withhold targeting information from agencies, which can either mitigate "bid rotation" and attract direct advertisers or reduce the efficiency of allocation for agency-using advertisers.

This paper contributes to the targeted advertising literature by examining the role of information asymmetry in targeted advertising. In the presence of this asymmetry, we show that both publisher revenue and conversion rates can increase simultaneously when the publisher disables microtargeting, which is not the case for symmetric advertisers. It also contributes to the literature on online advertising auctions. In our model we examine the interaction between different conditions on advertisers' valuations (correlated and independent), which consist of two different components (contextual and behavioral), and different information settings (where different advertisers have access to different information), providing a comprehensive look at how information asymmetries can affect the equilibrium market outcomes.

3 Model

Two[2] advertisers are competing for a single ad impression, provided by a publisher.

Data Differentiation. Advertiser i's valuation, v_i, for the impression consists of two components:

$$v_i = \kappa \cdot b_i + (1 - \kappa) \cdot c_i,$$

where $b_i, c_i \in [0, 1]$ are random variables matching two different types of consumer data to each advertiser. For example, we can interpret c_i as a function from the consumer's contextual data to a matching value in $[0, 1]$ for advertiser i. Similarly, we can interpret b_i as a function from the

[2]We start with two advertisers for simplicity, but the results hold for any number of advertisers. Because two advertisers are sufficient to showcase some of the core insights and intuitions, we start with two advertisers in Section 4. In Sections 5 and 6, we generalize the model and results to any number of advertisers ($n \geq 2$).

consumer's behavioral data to a matching value in $[0, 1]$ for advertiser i. The parameter $\kappa \in [0, 1]$ is a weight parameter that allows for different contributions of the two types of data to the valuations (i.e., the larger κ is, the more important b_i is in comparison to c_i in a particular setting).

The valuation $v_i \in [0, 1]$ is the amount of money advertiser i is willing to pay to display an ad to this specific consumer with matching values b_i and c_i. If we assume that the advertiser receives a normalized profit of 1 if they manage to convert the consumer, then we can also think of v_i as the probability that an impression for advertiser i will result in a conversion. As we will see next, advertiser i does not necessarily know their v_i, because they might not have access to all the data (but they have some expectations about it).

For ease of interpretability, throughout the paper, we will often refer to b_i as the behavioral variable (corresponding to behavioral data) and to c_i as the contextual variable (corresponding to contextual data). However, the model itself is agnostic to the specific types of data these variables correspond to, so different interpretations are valid (i.e., demographic data can be included, or behavioral and contextual data can be switched, etc.). The important thing is the properties these types of data satisfy, which we describe next.

Asymmetry of Information. So far, b_i and c_i are interchangeable. What differentiates them in the model is that we assume that c_i is always known by advertiser i, but b_i is not necessarily known.[3] More specifically, we consider three different information settings with regards to b_i:

- **Full Information (FI)**: All advertisers know their b_i.
- **Information Asymmetry (IA)**: Advertiser 1 knows b_1, while Advertiser 2 does not know b_2. We refer to Advertiser 1 as the *exchange advertiser*, while we refer to Advertiser 2 as the *direct advertiser*.[4]
- **Contextual Targeting (CT)**: None of the advertisers know their b_i.

Based on the data that is available to each advertiser, they form an expectation of their v_i, that they use for their decisions. When an advertiser knows and uses both b_i and c_i, we refer to this as

[3] This assumption also fits with the interpretation that c_i corresponds to contextual data and b_i to behavioral data. Usually contextual information is easily accessible, but behavioral information for a consumer is not always available.

[4] An interpretation for the data asymmetry is that Advertiser 1 participates through an ad exchange that gives them access to the extra information, while Advertiser 2 does not use an intermediary ad exchange, hence the names.

microtargeting. When only c_i is known, we refer to this as *contextual targeting*.

Advertiser Correlation. The random variables b_i of different advertisers can potentially be correlated with each other. We consider the following two extreme cases for the b_i's:

- **Common-value case (CV)**: $b_1 = b_2 =: b$, where b is drawn from a distribution with CDF F.
- **Independent-values case (IV)**: b_1, b_2 are i.i.d. draws from a distribution with CDF F.

For parsimony and analytical tractability, we model the distribution F as a Bernoulli distribution in $\{0, 1\}$, with $\Pr[b_i = 1] = p$ for some probability $p \in [0, 1]$.

We can also consider different cases regarding the correlation of the variables c_i. However, since c_i is always known by advertisers, the common-value case is not that interesting, so we omit it here for brevity, and we consider the following case for the c_i's:[5]

- The random variables c_i are i.i.d. draws from a distribution with CDF G.

We model the distribution G as a uniform distribution in $[0, 1]$.[6]

The Auction. The impression is sold using a second-price auction run by the publisher. The advertisers bid based on the expectations they have formed about their v_i, and the advertiser with the highest bid wins and pays the second-highest bid to the publisher.[7]

Notation. For each behavioral case $\sigma \in \{\mathrm{CV}, \mathrm{IV}\}$ (common-value, independent-values) and each information setting $\tau \in \{\mathrm{FI}, \mathrm{IA}, \mathrm{CT}\}$ we denote by W_τ^σ and V_τ^σ the publisher's expected revenue and the expected conversion rate, respectively. Similarly, we denote by E_τ^σ and D_τ^σ the exchange

[5]For completeness, we consider the common-value case for c_i's separately in Appendix C.

[6]This is again for simplicity. In Sections 5 and 6, we consider an arbitrary distribution G instead of a uniform and show that the results remain robust.

[7]In practice, more elaborate selling mechanisms are possible. For example, an ad exchange can run its own auction among its advertisers and submit the clearing price to the publisher; the publisher can then run an additional auction with all the received bids to determine the winner. Each auction can also be of different formats, e.g. second-price or first-price. Here, we abstract away from the complications of the selling mechanism itself by using a standard second-price auction for selling the impression. Even though alternative mechanisms can complicate the analysis significantly, the intuition for the results described in the paper still holds. To demonstrate this, in Lemma 4 we show that the result that the conversion rate can increase by disabling microtargeting, holds for a wide variety of selling mechanisms, including e.g. a single or multiple first-price auctions. For completeness, in Appendix B.3 we also show that for first-price auctions the publisher's revenue can also increase by disabling microtargeting (Proposition 12). For other selling mechanisms, similar intuition applies.

advertiser's and the direct advertiser's expected payoffs, respectively. Table 1 summarizes all the notation described above. Figure 13 summarizes the main results.

The rest of the paper is structured as follows. In Section 4, we present the main results and insights for the main model described above. Then, in Sections 5 and 6 we show the robustness of our results in more general settings that include a larger number of advertisers and arbitrary distributions G. In Section 5 we do it analytically for the results where a proof is feasible despite the lack of a closed-form bidding function.[8] In Section 6 we start by showing the existence of a pure symmetric equilibrium for the general model;[9] we then numerically approximate the bidding function for some general examples and show the robustness of the results for the general case. All proofs are relegated to Appendices A.1 and B.1 and a summary of all key formulas can be found in Appendix B.2. Lastly, in Appendix B.3 we analyze a variation of the model, where the impression is sold with a first-price auction instead of a second-price auction, and we show the robustness of the main results.

4 Analysis and Main Insights

In this section, we start with the main model with two advertisers. In subsection 4.1 we consider the common-value case (CV) and in subsection 4.2 we consider the independent-values case (IV). In subsection 4.3 we compare and discuss the differences between the common-value and independent-values cases in terms of publisher's revenue, conversion rates, and advertisers' payoffs.

4.1 Common-value case

Under the common-value case in the CT and FI settings, both advertisers will have the same information, therefore, in the second-price auction they will truthfully bid their valuation (in FI) or their expected valuation (in CT, where they do not know the actual valuation) (see e.g., Krishna, 2009). However, in the IA setting the exchange advertiser is more informed than the direct advertiser. As a consequence of this asymmetry, the exchange advertiser will still bid their true valuation, but the

[8] More specifically, the bidding function of the direct advertisers is the solution to a differential equation that does not always have a closed-form solution (see equation 4 and the proof of Lemma 3).

[9] More specifically, we show that there is a pure symmetric equilibrium bidding strategy for the direct advertisers under the common-value IA setting.

Information settings	
FI	Full Information. The publisher provides the behavioral data to all the advertisers.
IA	Microtargeting with Information Asymmetry. Only the exchange advertisers have access to behavioral data.
CT	Contextual Targeting. No advertiser has access to behavioral data.
Behavioral-value settings	
CV	Common Value. The behavioral value is the same for all the advertisers.
IV	Independent Values. The behavioral values are independent between advertisers.
Parameters	
κ	Behavioral Weight. The importance of behavioral data compared to contextual data on the advertisers' valuations.
p	Behavioral Probability. The probability that the behavioral value b of a random consumer is high for an advertiser.
Market Metrics	
V_{τ}^{σ}	Expected conversion rate under the behavioral-value setting $\sigma \in \{\mathrm{CV}, \mathrm{IV}\}$ and the information setting $\tau \in \{\mathrm{FI}, \mathrm{IA}, \mathrm{CT}\}$.
W_{τ}^{σ}	Publisher's expected revenue under the behavioral-value setting $\sigma \in \{\mathrm{CV}, \mathrm{IV}\}$ and the information setting $\tau \in \{\mathrm{FI}, \mathrm{IA}, \mathrm{CT}\}$.
E_{τ}^{σ}	Exchange advertiser's expected payoff under the behavioral-value setting $\sigma \in \{\mathrm{CV}, \mathrm{IV}\}$ and the information setting $\tau \in \{\mathrm{FI}, \mathrm{IA}, \mathrm{CT}\}$.
D_{τ}^{σ}	Direct advertiser's expected payoff under the behavioral-value setting $\sigma \in \{\mathrm{CV}, \mathrm{IV}\}$ and the information setting $\tau \in \{\mathrm{FI}, \mathrm{IA}, \mathrm{CT}\}$.
Others	
$v = \kappa b + (1-\kappa)c$	Advertiser's valuation for behavioral value b and contextual value c. It is also used as a proxy for conversion rate.
$\beta(c)$	Equilibrium bidding function of a direct advertiser under the common-value IA setting, where the direct advertiser does not know b but they know c.
Generalizations	
Number of advertisers	
n_1	Number of exchange advertisers. They have access to behavioral data, except in the contextual-targeting information setting (CT).
n_2	Number of direct advertisers. They only have access to contextual data, except in the full-information setting (FI).
$n = n_1 + n_2$	Total number of advertisers.
Distributions	
G	Contextual Distribution. The CDF of the distribution of the contextual value of a random consumer for an advertiser.

Table 1: Summary of Notation

direct advertiser might not always do that. We start off with Lemma 1 on the bidding function of the direct advertiser.

Lemma 1 (Advertisers' bidding behavior)**.** *Under the common-value* IA *setting, the exchange advertiser bids their true valuation, while given contextual value $c \in [0,1]$ the direct advertiser's bidding function is*

$$\beta(c) := \begin{cases} (1-\kappa)c, & \text{if } 0 \le c < \min\left\{\frac{\sqrt{1-p}}{\sqrt{1-p}+\sqrt{p}}, \frac{\sqrt{1-p}\kappa}{1-\kappa}\right\}, \\ \kappa p + (1-\kappa)c, & \text{if } \min\left\{\frac{\sqrt{1-p}}{\sqrt{1-p}+\sqrt{p}}, \frac{\sqrt{1-p}\kappa}{1-\kappa}\right\} \le c < \max\left\{\frac{\sqrt{1-p}}{\sqrt{1-p}+\sqrt{p}}, 1-\frac{\sqrt{p}\kappa}{1-\kappa}\right\}, \\ \kappa + (1-\kappa)c, & \text{if } \max\left\{\frac{\sqrt{1-p}}{\sqrt{1-p}+\sqrt{p}}, 1-\frac{\sqrt{p}\kappa}{1-\kappa}\right\} \le c \le 1. \end{cases} \tag{1}$$

The intuition behind Lemma 1 is the following. If the contextual value c of the direct advertiser is relatively low, they bid as if the common behavioral value b is 0. This is because if they assume some other value $b = x > 0$, they risk overpaying for the impression in the case where $b = 0$ and $(1-\kappa) \cdot c < (1-\kappa) \cdot c' < \kappa \cdot x + (1-\kappa) \cdot c$ (where c' is the contextual value of the exchange advertiser), where they end up with a negative payoff of $(1-\kappa) \cdot (c - c')$. When $c < \min\left\{\frac{\sqrt{1-p}}{\sqrt{1-p}+\sqrt{p}}, \frac{\sqrt{1-p}\kappa}{1-\kappa}\right\}$, this risk is too high to take. However, when the contextual value c is high $\left(c \ge \max\left\{\frac{\sqrt{1-p}}{\sqrt{1-p}+\sqrt{p}}, 1-\frac{\sqrt{p}\kappa}{1-\kappa}\right\}\right)$, it is very likely that $c > c'$, therefore they are not afraid to bid as if $b = 1$, because they have a higher incentive to win and avoid losing impressions with good behavioral values. For medium values of c, both the risks of overpaying for a bad impression and losing a good impression are too high to make any assumption about b, therefore the advertiser simply bids their expected valuation (note that the expected value of b is p).

Note that as κ increases, i.e. as the behavioral data becomes more important, the middle interval of c where the direct advertiser bids their expected valuation shrinks, and for $\kappa \ge 1/2$ it disappears, i.e. the advertiser either underbids or overbids depending on c (see also Figure 1 where the bidding function is shown for different values of κ). Since the role of information asymmetry is more important for larger values of κ and it is where the more interesting results occur, for simplicity for some of the analysis, we focus on the case where $\kappa \ge 1/2$, unless otherwise noted.

Note also that when the value of p is low, the region of c where the underbidding occurs is wider compared to the overbidding region, but the amount of underbidding (κp) is smaller compared to the amount of overbidding ($\kappa(1-p)$). On the other hand, when p is high, overbidding is more

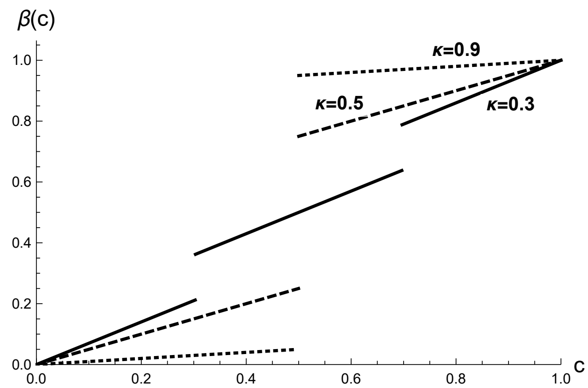
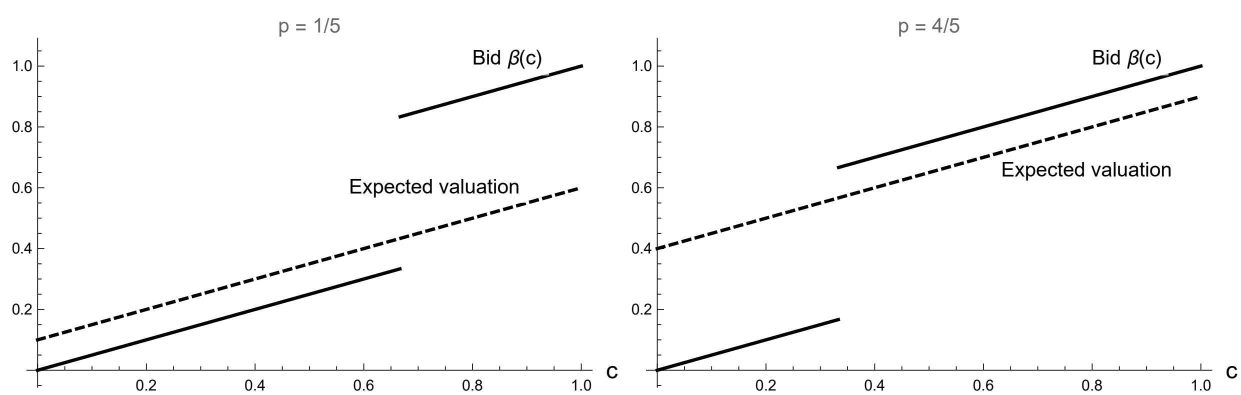


Figure 1: Bidding function of the direct advertiser for different values of κ (solid line for $\kappa = 0.3$, dashed line for $\kappa = 0.5$, and dotted line for $\kappa = 0.9$), $n_1 = n_2 = 1$, $p = 1/2$, and $G(x) = x$. Notice that for large contextual values c, as κ increases there is more overbidding, while for small contextual values c, as κ increases there is more underbidding.

common but the amount of overbidding is lower. This is illustrated in Figure 2.

In Lemma 2 of Section 5 we show a generalization of Lemma 1 for any $n_1 \geq 1$, any distribution G, $p \in [0, 1]$, and $\kappa \geq 1/2$. Lemma 3 in Section 6 is a further generalization for the more general case with $n_2 \geq 1$ (where the bidding function does not always have a closed-form expression). The same intuition as for Lemma 1 applies to Lemma 2 and Lemma 3 as well.



(a) $p = 1/5$.

(b) $p = 4/5$.

Figure 2: Bidding function of the direct advertiser (solid line) compared to their expected valuation (dashed line), for different values of p, $n_1 = n_2 = 1$, $\kappa = 1/2$, and $G(x) = x$. Notice that for small values of p (left) the region of overbidding is smaller than the region of underbidding, but the amount of overbidding ($\kappa(1 - p)$) is larger than the amount of underbidding (κp). For large values of p (right) the opposite happens.

The bidding function of Lemma 1 sometimes results in an inefficient market under the IA information setting. More specifically, both the underbidding and the overbidding can result in lower conversion rate compared to the CT setting (where every advertiser bids their expected valuation). This is illustrated in Example 1.

Example 1 (Inefficiency of non-truthful bidding)**.** Let $\kappa = p = 1/2$. Then the direct advertiser bids $c/2$ if $c < 1/2$ and $(1 + c)/2$ if $c \geq 1/2$, where c is their contextual value. The following two examples illustrate the inefficiency caused by the non-truthful bidding of the direct advertiser. They show that both underbidding and overbidding can result in lower conversion rates.

- **Inefficiency of underbidding**

  (IA setting) Suppose that the common behavioral value is high, i.e. $b = 1$, the exchange advertiser has contextual value $c_1 = 1/6$, and the direct advertiser has contextual value $c_2 = 1/3$. The actual valuations of the two advertisers are $v_1 = (1 + c_1)/2 = 7/12$ for the exchange advertiser and $v_2 = (1 + c_2)/2 = 8/12$ for the direct advertiser. The exchange advertiser bids their actual valuation $\beta_1 = 7/12$, but the direct advertiser underbids, i.e. $\beta_2 = c_2/2 = 2/12$. As a result, the direct advertiser loses the auction even though they have a higher valuation. Thus, the conversion rate ends up being lower than what it could be in a more efficient auction.

  (CT setting) If none of the advertisers knew the behavioral value b, then both advertisers would bid their expected valuations, i.e. $\beta_1 = 1/4 + c_1/2 = 4/12$ and $\beta_2 = 1/4 + c_2/2 = 5/12$. Thus, the advertiser with the highest valuation would win, leading to a higher conversion rate.

- **Inefficiency of overbidding**

  (IA setting) Suppose that the common behavioral value is low, i.e. $b = 0$, the exchange advertiser has contextual value $c_1 = 5/6$, and the direct advertiser has contextual value $c_2 = 2/3$. The actual valuations of the two advertisers are $v_1 = c_1/2 = 5/12$ for the exchange advertiser and $v_2 = c_2/2 = 4/12$ for the direct advertiser. The exchange advertiser bids their actual valuation $\beta_1 = 5/12$, but the direct advertiser overbids, i.e. $\beta_2 = (1+c_2)/2 = 10/12$. As a result, the direct advertiser wins the auction even though they have a lower valuation. Thus, the conversion rate ends up being lower than what it could be in a more efficient auction.

  (CT setting) If none of the advertisers knew the behavioral value b, then both advertisers would bid their expected valuations, i.e. $\beta_1 = 1/4 + c_1/2 = 8/12$ and $\beta_2 = 1/4 + c_2/2 = 7/12$.

Thus, the advertiser with the highest valuation would win, leading to a higher conversion rate.

As we can see in Example 1, there are cases where under the IA setting the advertiser with the highest valuation does not win, either due to the underbidding or due to the overbidding of the direct advertiser. In contrast, under the CT setting, the highest-valuation advertiser always wins, because every bidder bids their expected valuation, and the winner is determined based on the contextual values. This results in a higher conversion rate for the CT setting, as shown in Proposition 1.

For the publisher's revenue, things are less clear. On the one hand, the underbidding that occurs under IA can hurt the publisher, but on the other hand, the overbidding can benefit the publisher because it can increase the prices. Surprisingly, the opposite can happen too; underbidding can sometimes increase publisher's revenue, and overbidding can decrease it, as illustrated in Example 2.

Example 2 (The effects of non-truthful bidding on revenue)**.** Let $\kappa = 1/2$ and $p = 1/3$. Then the direct advertiser bids $c/2$ if $c < 2 - \sqrt{2}$ and $(1 + c)/2$ if $c \geq 2 - \sqrt{2}$, where c is their contextual value. The following two examples illustrate that, counter-intuitively, underbidding can sometimes increase publisher's revenue, and overbidding can sometimes decrease it.

- **Underbidding can increase publisher's revenue**

  (IA setting) Suppose that the common behavioral value is high, i.e. $b = 1$, the exchange advertiser has contextual value $c_1 = 1/12$, and the direct advertiser has contextual value $c_2 = 1/2$. The actual valuations of the two advertisers are $v_1 = (1 + c_1)/2 = 13/24$ for the exchange advertiser and $v_2 = (1 + c_2)/2 = 18/24$ for the direct advertiser. The exchange advertiser bids their actual valuation $\beta_1 = 13/24$, but the direct advertiser underbids, i.e. $\beta_2 = c_2/2 = 6/24$. The exchange advertiser wins and pays β_2, therefore, the publisher's revenue is $6/24$.

  (CT setting) If none of the advertisers knew the behavioral value b, then both advertisers would bid their expected valuations, i.e. $\beta_1 = 1/6 + c_1/2 = 5/24$ and $\beta_2 = 1/6 + c_2/2 = 10/24$. Then the direct advertiser would win and pay β_1. Therefore, publisher's revenue would be $5/24$, which is lower than the revenue under the IA setting.

- **Overbidding can decrease publisher's revenue**

  (IA setting) Suppose that the common behavioral value is low, i.e. $b = 0$, the exchange advertiser has contextual value $c_1 = 5/6$, and the direct advertiser has contextual value $c_2 = 3/4$. The actual valuations of the two advertisers are $v_1 = c_1/2 = 10/24$ for the exchange advertiser and $v_2 = c_2/2 = 9/24$ for the direct advertiser. The exchange advertiser bids their actual valuation $\beta_1 = 10/24$, but the direct advertiser overbids, i.e. $\beta_2 = (1 + c_2)/2 = 21/24$. The direct advertiser wins and pays β_1, therefore, publisher's revenue is $10/24$.

  (CT setting) If none of the advertisers knew the behavioral value b, then both advertisers would bid their expected valuations, i.e. $\beta_1 = 1/6 + c_1/2 = 14/24$ and $\beta_2 = 1/6 + c_2/2 = 13/24$. Then the exchange advertiser would win and pay β_2. Therefore, publisher's revenue would be $13/24$, which is higher than the revenue under the IA setting.

Despite valuation instances like those in Example 2, in Proposition 1 we show that the overall publisher's expected revenue is higher under the CT setting.

Proposition 1 (Common-value and information asymmetry). *Under the common-value behavioral setting, the publisher can improve both the conversion rate and the expected revenue if it hides the behavioral information from all the advertisers. In other words, we have $V^{\mathrm{CV}}_{\mathrm{IA}} \leq V^{\mathrm{CV}}_{\mathrm{CT}}$ and $W^{\mathrm{CV}}_{\mathrm{IA}} \leq W^{\mathrm{CV}}_{\mathrm{CT}}$.*

Proposition 1 shows that if the publisher has some useful information about a consumer but cannot provide this information to all advertisers, it can achieve a higher conversion rate by hiding the information from everyone rather than giving it only to some advertisers. As an added benefit, the publisher can also simultaneously increase its revenue by hiding this information for all advertisers. The main reason this happens is the inefficiency of the non-truthful bidding of the direct advertiser under the IA setting, as illustrated in Example 1.

Given the result of Proposition 1, one may wonder if the same can happen when there is no information asymmetry between the advertisers. In other words, if all advertisers have access to the same information, is it still possible that less information can simultaneously increase the conversion rate and the publisher's revenue? In Proposition 2 we show that this cannot happen under the common-value behavioral setting.

Proposition 2 (Common-value and full information)**.** *Under the common-value behavioral setting, both the conversion rate and the expected revenue remain unchanged when all advertisers have access to the same information (i.e., when all advertisers have access to behavioral data or none of the advertisers have access to behavioral data). In other words, it holds that* $V_{\text{FI}}^{\text{CV}} = V_{\text{CT}}^{\text{CV}}$ *and* $W_{\text{FI}}^{\text{CV}} = W_{\text{CT}}^{\text{CV}}$.

The equality $V_{\text{FI}}^{\text{CV}} = V_{\text{CT}}^{\text{CV}}$ is relatively easy to see, whereas the equality $W_{\text{FI}}^{\text{CV}} = W_{\text{CT}}^{\text{CV}}$ is less straightforward. Under the common-value setting, since all advertisers have the same behavioral value, when all have the same information, the behavioral part of their bids is the same for everyone; therefore, the winner of the auction is purely determined by their contextual values in both the FI and the CT settings. As a result, the conversion rate remains unchanged.

For the revenue, when the common behavioral value is high, publisher's revenue is higher under the CT setting because every advertiser bids above their actual valuation. In contrast, when the common behavioral value is low, the publisher's revenue is lower under the CT setting because every advertiser bids below their actual valuation. Due to the linearity of the expectation, the average revenue remains the same in the two information settings.

Note that as we move from the FI to the IA and then to the CT information setting, the overall information to the advertisers is reduced. As a result, the inequality $W_{\text{FI}}^{\text{CV}} \geq W_{\text{IA}}^{\text{CV}}$ agrees with the linkage principle (Milgrom and Weber, 1982) which would suggest that revealing information is better for the revenue, but the inequality $W_{\text{IA}}^{\text{CV}} \leq W_{\text{CT}}^{\text{CV}}$ violates the principle which happens due to the information asymmetry.[10]

In this section, we have seen that under the common-value setting there is a non-monotonic relationship between the amount of information available to the advertisers and the efficiency of the auction; as we reduce the information, efficiency (i.e. the conversion rate) first goes down and then goes up again. We have also seen that a similar effect occurs for the publisher's revenue. In Section 4.2 we show that this is no longer true when the behavioral values are independent.

[10] For some other cases where the principle is violated for different reasons, see e.g. Perry and Reny (1999); Fang and Parreiras (2003); Krishna (2009); Despotakis et al. (2017).

4.2 Independent-values case

In contrast to the common-value case, when the behavioral values of advertisers are independent, all advertisers will bid truthfully according to their (expected) valuation.

In the independent-values case, the intuitive result that less information to the advertisers decreases the conversion rate now holds. This is still not true for every valuation instance, as illustrated in Example 3, but it is true for the expected conversion rates, as shown in Proposition 3.

Example 3. Let $\kappa = p = 1/2$.

(IA setting) Suppose that the exchange advertiser has a behavioral value $b_1 = 1$ and a contextual value $c_1 = 3/8$, and the direct advertiser has a behavioral value $b_2 = 1$ and a contextual value $c_2 = 5/8$. The actual valuations of the two advertisers are $v_1 = (1 + c_1)/2 = 11/16$ for the exchange advertiser and $v_2 = (1 + c_2)/2 = 13/16$ for the direct advertiser. The exchange advertiser bids their actual valuation $\beta_1 = 11/16$, but the direct advertiser bids their expected valuation, i.e. $\beta_2 = 1/4 + c_2/2 = 9/16$. As a result, the direct advertiser loses the auction even though they have higher valuation.

(CT setting) If none of the advertisers knew their behavioral values b_i, then both advertisers would bid their expected valuations, i.e. $\beta_1 = 1/4 + c_1/2 = 7/16$ and $\beta_2 = 1/4 + c_2/2 = 9/16$. Thus, the advertiser with the highest valuation would win, leading to a higher conversion rate than the IA setting.

Despite valuation instances like those in Example 3, in Proposition 3 we show that the overall expected conversion rate increases with more information, under the independent-values setting.

Proposition 3 (Independent-values, conversion rates)**.** *Under the independent-values behavioral setting, the less information advertisers have overall, the lower the conversion rate is. More specifically,* $V^{\mathrm{IV}}_{\mathrm{FI}} \geq V^{\mathrm{IV}}_{\mathrm{IA}} \geq V^{\mathrm{IV}}_{\mathrm{CT}}$.

Proposition 3 shows that the dependence between the behavioral values of different advertisers is an essential element for the result of Proposition 1, since for independent values it no longer holds.

With respect to the publisher's revenue, the result is less intuitive. Proposition 4 shows that as we provide more information in general to advertisers, publisher revenue decreases.

Proposition 4 (Independent-values, publisher's revenues)**.** *Under the independent-values behavioral setting, the less information advertisers have overall, the higher publisher's revenue is. More specifically, we have* $W^{\mathrm{IV}}_{\mathrm{FI}} \leq W^{\mathrm{IV}}_{\mathrm{IA}} \leq W^{\mathrm{IV}}_{\mathrm{CT}}$.[11]

The result of Proposition 4 is sensitive to the number of advertisers (in contrast to the previous results that hold for arbitrary number of advertisers; see Section 5). What happens in general is that, for a small number of advertisers, less information is better, but for a large number of advertisers, more information is better. This is due to a version of the market-thinning effect (Levin and Milgrom, 2010). When there are few advertisers in the market, as they become more informed their values spread out, and there is less competition in the high valuations. But as the number of advertisers becomes larger and the competition increases, more information should improve publisher's revenue. More specifically, when $n_1 = n_2 = 1$ (i.e. there is one advertiser of each type) it holds that more information decreases revenue, but as n_1 and n_2 increase, at some point this stops being true. The exact threshold for the number of advertisers where monotonicity changes depends on the value of p, with a lower p increasing the threshold, the behavioral-value weight κ, with higher κ increasing the threshold, and the contextual distribution G (see Proposition 9 and Figure 7 for more details).[12]

4.3 Comparison of the Behavioral-Value Settings and Advertisers' Payoffs

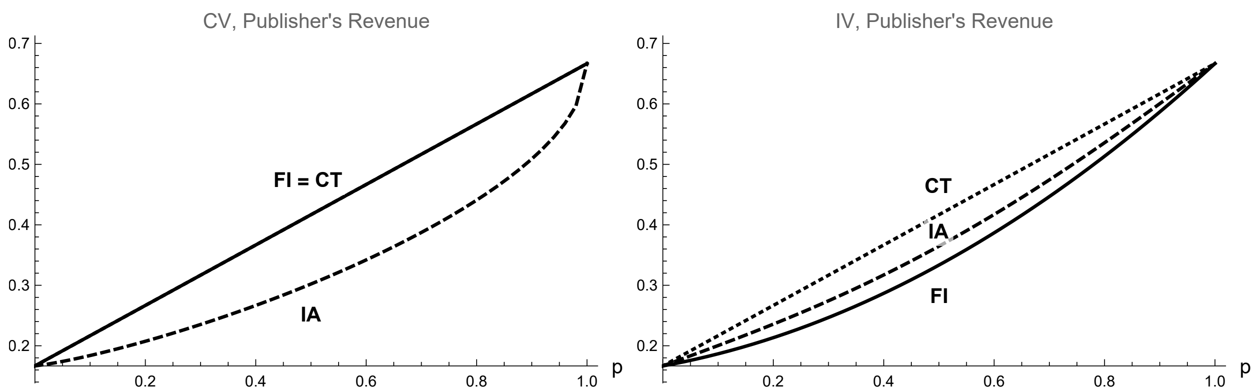
Now that we have the results for the simple model with two advertisers, we can compare the two behavioral-value settings (the common-value case and the independent-values cases) to each other in terms of their consequences for the publisher's revenue, conversion rates, and advertisers' payoffs.

We start with the publisher's revenue in Figure 3. In the two plots of Figure 3, we see the revenue under the three different information settings as the behavioral probability p changes in $[0, 1]$. In the common-value case in Figure 3(a), we can see that starting from the IA setting and eliminating the information asymmetry by going towards FI or CT, the publisher's revenue increases. This is due to the underbidding and overbidding behavior that occurs in IA, as discussed

[11] We want to highlight that this result holds for a low number of advertisers (e.g. two, like in the main model), but unlike the other results it does not always generalize for more advertisers. In Section 5.2 we consider the general case and discuss the details on this.

[12] It is interesting to note that there are also cases where the expected revenue is non-monotonic with respect to the total amount of information that is available to the advertisers. In other words, all six different orderings of $W^{\mathrm{IV}}_{\mathrm{FI}}$, $W^{\mathrm{IV}}_{\mathrm{IA}}$, and $W^{\mathrm{IV}}_{\mathrm{CT}}$ are possible under different conditions (see Figure 8).

in Section 4.1. In contrast to Figure 3(a), in the independent-values case in Figure 3(b) we observe a monotonic change in revenue. As we add information to the market (moving from CT to IA and then to FI), publisher's revenue goes down, as described in Proposition 4.



(a) Common-value case.

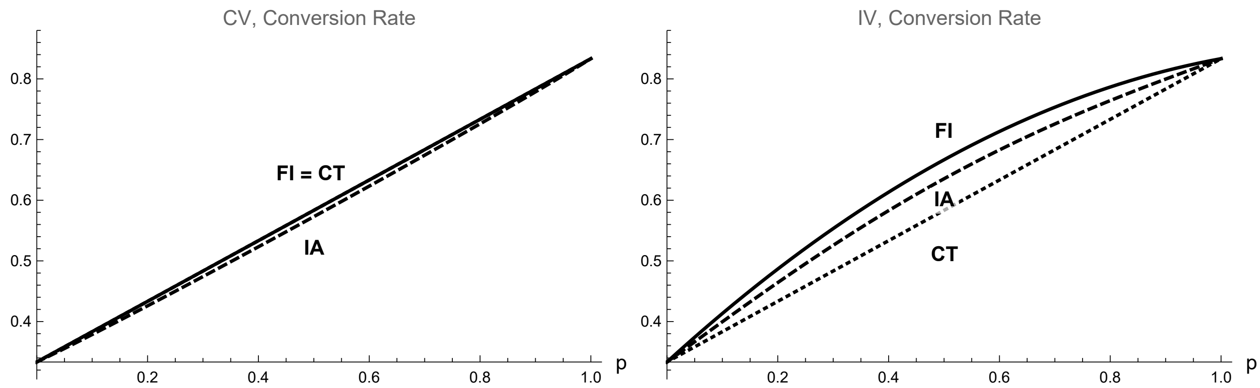
(b) Independent-values case.

Figure 3: Publisher's revenue under the different information settings for different values of $p \in [0, 1]$, $n_1 = n_2 = 1$, $\kappa = 1/2$, and $G(x) = x$.

Next, we move to the conversion rates in Figure 4. What we observe in Figure 4(a) is one of our main findings. What happens here is that in the IA setting, according to Lemma 1, a direct advertiser with high valuation often bids conservatively and loses to an exchange advertiser with a lower valuation. In addition, a direct advertiser with low valuation often bids aggressively and wins against an exchange advertiser with a higher valuation. Both of these bidding behaviors create an inefficient auction because an advertiser with lower valuation wins the consumer's impression, resulting in a lower conversion rate compared to the settings without information asymmetry.

Often in the literature, we see that in markets with thin competition when the publisher's revenue goes down (Figure 3(b)), conversion rate (Figure 4(b)) and the advertisers' payoffs (Figure 6) go up as we add information to the market (moving from CT to IA and then to FI). Here, we verify this for our model. However, Proposition 1 states that this is not the case when the behavioral values are correlated. In fact, both publisher revenue and conversion rate can move in the same direction, as we observe in Figures 3(a) and 4(a) in contrast to Figures 3(b) and 4(b).

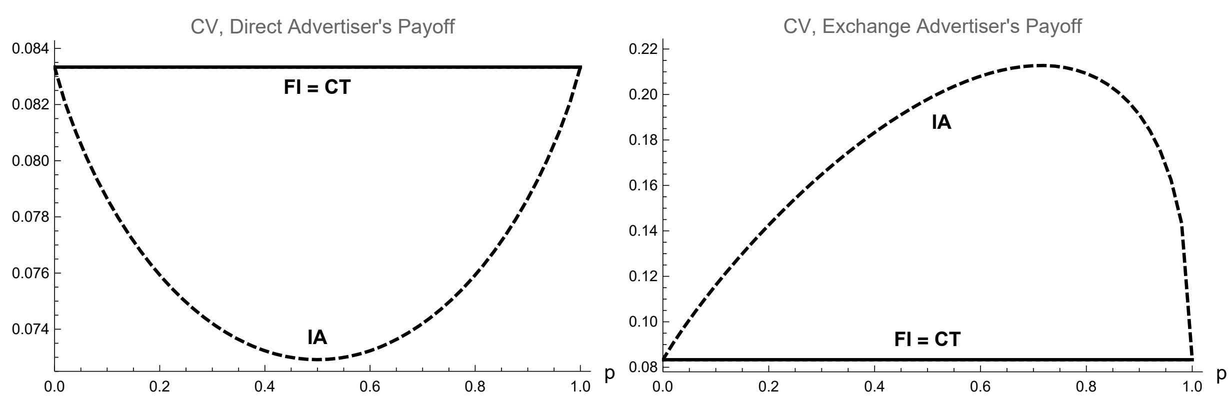
Regarding advertisers' payoffs, in Figure 5 we can see that in the common-value case they change non-monotonically both in terms of the information that is available to the advertisers and in terms of p. First, in Figure 5(a), we see that the direct advertiser's payoff decreases slightly in



(a) Common-value case. (b) Independent-values case.

Figure 4: Conversion rate under the different information settings as a function of $p \in [0, 1]$, for $n_1 = n_2 = 1$, $\kappa = 1/2$, and $G(x) = x$.

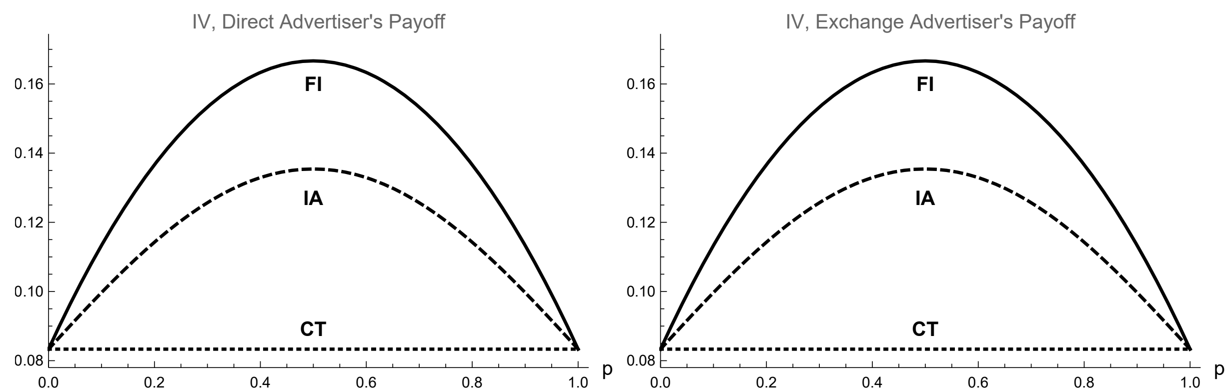
the IA setting compared to the FI and CT settings, while in Figure 5(b) we see that the exchange advertiser's payoff increases significantly. This is expected because the exchange advertiser has a strong competitive advantage in the IA setting, while in the FI and CT settings both advertisers are similar. Second, in terms of p, we see that in the IA setting, the direct advertiser's payoff is minimum for $p = 1/2$ where the uncertainty about the common behavioral value is maximized. However, the exchange advertiser's payoff is maximized for a value $p > 1/2$, which gives the exchange advertiser a higher probability of a high valuation in addition to the advantageous uncertainty of the direct advertiser.



(a) Direct advertiser. (b) Exchange advertiser.

Figure 5: Advertisers' payoffs in the common-value case under the different information settings for different values of $p \in [0, 1]$, $n_1 = n_2 = 1$, $\kappa = 1/2$, and $G(x) = x$.

In contrast to Figure 5, in the independent-values case in Figure 6 we see that both payoffs go down monotonically as we remove information from the market. Furthermore, we see that both types of advertisers have identical payoffs in all settings under IV, including the IA setting where the exchange advertiser would normally be expected to have an advantage. Although the direct advertiser has more fluctuations in their payoff in IA under different realizations of the valuations, their average payoff is the same as the exchange advertiser's one, because the advantage of extra information is not that big when the valuations are independent. This perhaps surprising result is independent of any distributional assumptions, but it is a consequence of the fact that there are only two advertisers in the simple version of the model. In Section 6 we discuss the more general case, which is more intuitive in the sense that the exchange advertiser has a higher payoff under the IA setting, but still interesting in terms of how the payoff changes as a function of p (see Figure 10).



(a) Direct advertiser.

(b) Exchange advertiser.

Figure 6: Advertisers' payoffs in the independent-values case under the different information settings for different values of $p \in [0, 1]$, $n_1 = n_2 = 1$, $\kappa = 1/2$, and $G(x) = x$.

5 Generalizations

In this section, we consider a more general version of the main model, for an arbitrary behavioral distribution G (instead of uniform) and more than two advertisers. More specifically, there are $n \geq 2$ advertisers competing for the impression, a subset of $n_1 \leq n$ of them are exchange advertisers, and the remaining $n_2 = n - n_1$ are direct advertisers. For the common-value case, the results shown in Section 4.1 extend to the more general setting; this is discussed in subsection 5.1. In the independent-values case there are some interesting differences when we increase the number of advertisers, which we discuss in subsection 5.2.

5.1 Common-value case

Lemma 2 is an analog of Lemma 1 for arbitrary distributions G and more than one exchange advertisers.

Lemma 2 (Advertisers' bidding behavior)**.** *For any distribution G, any $n_1 \geq 1$, $n_2 = 1$, and $\kappa \geq 1/2$, under the common-value* IA *setting, all the exchange advertisers bid their true valuations while there exists $\underline{c}(p) \in [0,1]$ such that the direct advertiser's bidding function is*

$$\beta(c) := \begin{cases} (1-\kappa)c, & \text{if } 0 \leq c < \underline{c}(p), \\ \kappa + (1-\kappa)c, & \text{if } \underline{c}(p) \leq c < 1. \end{cases}$$

Moreover, $\underline{c}$ is independent of κ, and it is a continuously differentiable decreasing function in p, with $\underline{c}(0) = 1$, $\underline{c}(1/2) = n_1 \mathbb{E}[c \cdot G(c)^{n_1-1}]$, and $\underline{c}(1) = 0$.

Like in Lemma 1, we see that the direct advertiser sometimes underbids, for low values of c, and sometimes overbids, for high values of c. Also, as p increases, they overbid more than they underbid. Due to this non-truthful bidding, similar results to those in Section 4.1 continue to hold for $n_1 > 1$.[13] Propositions 5 and 6 generalize the results of Propositions 1 and 2 for arbitrary distributions G. Proposition 5 is shown here for any number of advertisers $n_1, n_2 \geq 1$. We further check the robustness of Proposition 6 for $n_1, n_2 > 1$ in Section 6.

Proposition 5. *For any distribution G, any $n_1, n_2 \geq 1$, and $\kappa \in [0,1]$, under the common-value case, we have that $V_{\text{IA}}^{\text{CV}} \leq V_{\text{FI}}^{\text{CV}} = V_{\text{CT}}^{\text{CV}}$.*

Proposition 6. *For any distribution G, $n_1 = n_2 = 1$, and $\kappa \geq 1/2$, under the common-value case, we have that $W_{\text{IA}}^{\text{CV}} \leq W_{\text{FI}}^{\text{CV}} = W_{\text{CT}}^{\text{CV}}$.*

5.2 Independent-values case

Propositions 7 and 8 generalize the results of Propositions 3 and 4. The intuition for Proposition 7 is similar to that in the simple model version (as the amount of information available to advertisers decreases, the efficiency of the auction decreases).

[13] Lemma 3 in Section 6 generalizes this result for $n_2 > 1$ as well.

Proposition 7. *For any distribution G, any $n_1, n_2 \geq 1$, and $\kappa \in [0,1]$, under the independent-values case, we have that $V_{\mathrm{FI}}^{\mathrm{IV}} \geq V_{\mathrm{IA}}^{\mathrm{IV}} \geq V_{\mathrm{CT}}^{\mathrm{IV}}$.*

Proposition 8. *For any distribution G, $n_1 = n_2 = 1$, and $\kappa \in [0,1]$, under the independent-values case, we have that $W_{\mathrm{FI}}^{\mathrm{IV}} \leq W_{\mathrm{IA}}^{\mathrm{IV}} \leq W_{\mathrm{CT}}^{\mathrm{IV}}$.*

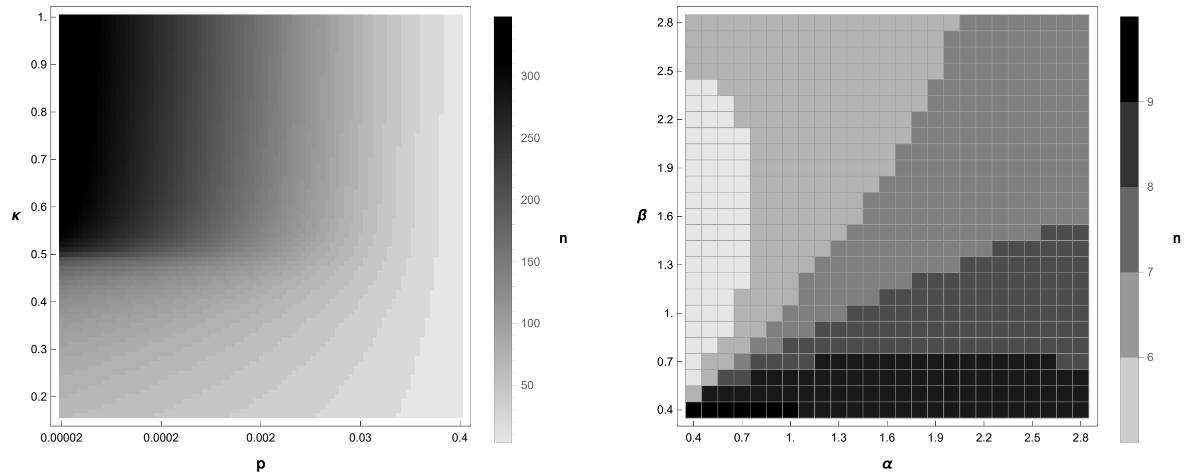
In contrast to the common-value setting, under independent behavioral values, publisher revenue behaves somewhat differently in general (for $n \geq 2$) than what we showed in Propositions 4 and 8. Proposition 9 describes the general phenomenon.

Proposition 9. *For any distribution G and $\kappa \geq 1/2$, under the independent-values case, we have $W_{\mathrm{FI}}^{\mathrm{IV}} \leq W_{\mathrm{CT}}^{\mathrm{IV}}$ for sufficiently small n, and $W_{\mathrm{FI}}^{\mathrm{IV}} > W_{\mathrm{CT}}^{\mathrm{IV}}$ for sufficiently large n. The threshold for n where the inequality is reversed depends on p, κ, and G.*

As described in Section 4.2, the market-thinning effect that occurs under the IV setting makes hiding information from advertisers beneficial for the publisher's revenue when the number of advertisers is low. However, when there is a sufficiently large number of advertisers, revealing more information increases revenue.

The threshold for the number of advertisers n where the inequality in Proposition 9 reverses depends on the parameters p and κ, and the distribution G. Figure 7 illustrates this. In Figure 7(a) we see that as p decreases and as κ increases, we need more and more advertisers to make the full-information setting give higher revenue than the contextual-targeting setting (i.e. $W_{\mathrm{FI}}^{\mathrm{IV}} \geq W_{\mathrm{CT}}^{\mathrm{IV}}$). In Figure 7(b) we see the thresholds for some examples of different Beta distributions for various parameters α and β. Generally, we observe that contextual distributions G with higher average have higher threshold. Also, if the average is low, a lower variance gives a higher threshold, but when the variance is high, a higher variance gives a higher threshold.

Given Proposition 9, a natural question to ask is how $W_{\mathrm{IA}}^{\mathrm{IV}}$ (the revenue under independent values with information asymmetry) compares to $W_{\mathrm{FI}}^{\mathrm{IV}}$ and $W_{\mathrm{CT}}^{\mathrm{IV}}$ under different conditions. The intuition behind Proposition 9 potentially suggests that $W_{\mathrm{FI}}^{\mathrm{IV}} \leq W_{\mathrm{IA}}^{\mathrm{IV}} \leq W_{\mathrm{CT}}^{\mathrm{IV}}$ for low n and $W_{\mathrm{FI}}^{\mathrm{IV}} \geq W_{\mathrm{IA}}^{\mathrm{IV}} \geq W_{\mathrm{CT}}^{\mathrm{IV}}$ for high n. Surprisingly, this is not always the case. In fact, as illustrated in Figure 8, all six different orderings between the revenues $W_{\mathrm{FI}}^{\mathrm{IV}}$, $W_{\mathrm{IA}}^{\mathrm{IV}}$, and $W_{\mathrm{CT}}^{\mathrm{IV}}$ are possible under different conditions. The information asymmetry between advertisers adds an additional element of complexity that the market-thinning effect alone is not sufficient to explain.



(a) $G(x) = x$, various p and κ.

(b) $G \sim \text{BETA}(\alpha, \beta)$, $p = 1/20$, $\kappa = 1/2$.

Figure 7: Minimum number of advertisers n such that $W_{\text{FI}}^{\text{IV}} \geq W_{\text{CT}}^{\text{IV}}$ for various values of p, κ, and different contextual distributions G.

The intuition behind Figure 8 is as follows. In Proposition 9 we saw that a low n makes hiding information from advertisers beneficial for the publisher, due to a thinner market. For a similar reason, a low p also makes hiding information beneficial. This is because when p is low, there is a low probability that the second highest bidder at the auction will have a behavioral value $b_i = 1$, which means that the clearing price will most likely be of the form $(1 - \kappa)c_i$ if all advertisers know their behavioral values. Thus, when p is low, the publisher prefers to hide information from as many advertisers as possible to make them bid their expected valuation $\kappa p + (1 - \kappa)c_i$ instead of their actual valuation. On the contrary, when p is high, the publisher prefers to reveal the behavioral information to as many advertisers as possible so that they can bid their (likely high) actual valuation. In other words, when p is low we have that $W_{\text{FI}}^{\text{IV}} < W_{\text{IA}}^{\text{IV}} < W_{\text{CT}}^{\text{IV}}$ (Region 1) and when p is high we have that $W_{\text{FI}}^{\text{IV}} > W_{\text{IA}}^{\text{IV}} > W_{\text{CT}}^{\text{IV}}$ (Region 2). This also explains why in Regions 1, 3, and 4 it is $W_{\text{FI}}^{\text{IV}} < W_{\text{CT}}^{\text{IV}}$, while in Regions 2, 5, and 6 it is $W_{\text{FI}}^{\text{IV}} > W_{\text{CT}}^{\text{IV}}$.

To understand why the IA setting generates higher revenue than the other two information settings in Regions 3 and 6 where p is medium and κ is high, let us consider the extreme case where $\kappa = 1$. In this extreme case, the valuations of the advertisers are just b_i, without a contextual element. Under the IA setting, the exchange advertisers will bid their actual values b_i, while the direct advertisers will bid their expected valuation which is just p. When p is high, there is a high chance that there will be at least two advertisers with high b_i's, so the publisher wants the

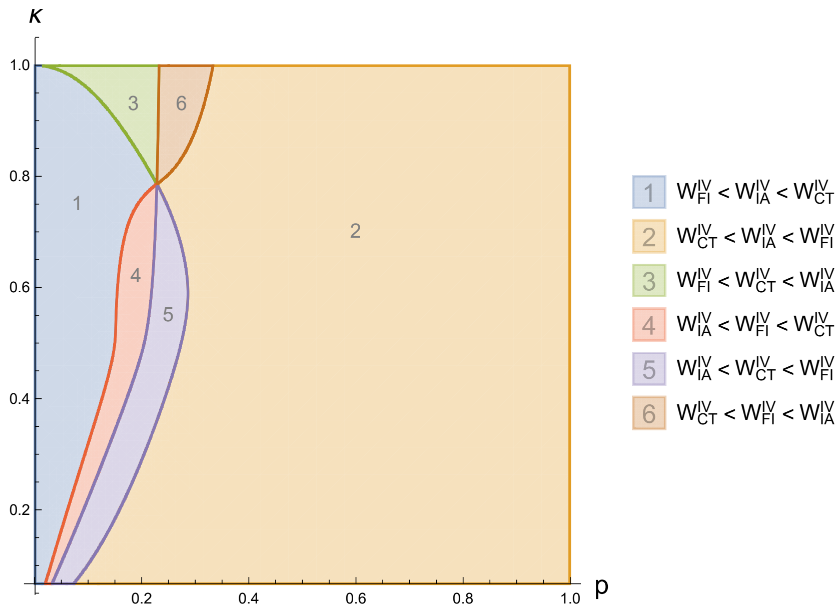


Figure 8: Publisher's revenues comparisons between different information settings under independent behavioral values, for various values of p and κ, $G(x) = x$, and $n_1 = n_2 = 2$.

advertisers to learn their b_i's to have a high chance of getting a clearing price of 1 (Region 2). When p is low, it is less likely that there will be at least two b_i's that are high, so the publisher prefers if the advertisers bid p instead of their b_i which is more likely 0. However, having many advertisers bidding p has no additional benefit compared to just two advertisers bidding p, since the clearing price will be p in both cases. Therefore, the optimal revenue for the publisher when p is low is achieved when there is information asymmetry, where the publisher guarantees a clearing price of at least p from the direct advertisers and there is also a (small) chance of something higher from the exchange advertisers (Regions 3 and 6).

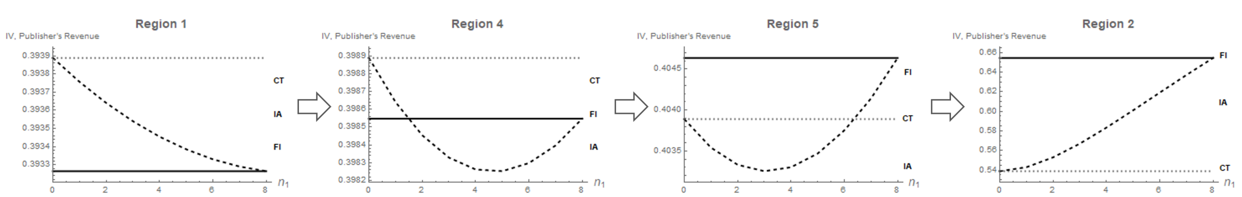


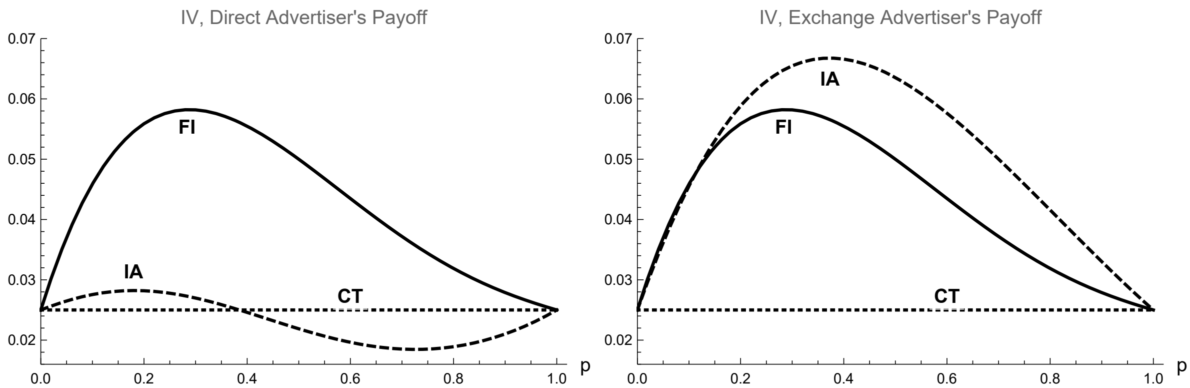
Figure 9: Publisher's revenues for the different information settings under independent behavioral values, for $n = 8$ advertisers, $n_1 \in [0, n]$, $n_2 = n - n_1$, $\kappa = 1/2$, $G(x) = x$, and $p \in \{0.01, 0.02, 0.03, 0.3\}$ (from left to right).

Finally, when κ is low, the importance of behavioral value on the advertisers' valuations is low. The contextual part of the valuations dominates in determining the winner. As a result, the benefit of information asymmetry described above, where it is good for the publisher to have both direct and exchange advertisers, is not essential anymore since the contextual values are known by both. In Regions 4 and 5, the IA setting has worse revenue than the other two settings because of a third effect.

Under the IA setting, there are two groups of advertisers, n_1 exchange advertisers and n_2 direct advertisers. If we fix the total number of advertisers $n = n_1 + n_2$, then we can think of the FI and the CT information settings as extreme versions of the IA setting. More specifically, FI is like IA with $(n_1, n_2) = (n, 0)$ and CT is like IA with $(n_1, n_2) = (0, n)$. With that view in mind, to understand how $W_{\text{FI}}^{\text{IV}}$, $W_{\text{IA}}^{\text{IV}}$, and $W_{\text{CT}}^{\text{IV}}$ compare to each other, it is useful to look at the function $W_{\text{IA}}^{\text{IV}}(n_1, n_2) = W_{\text{IA}}^{\text{IV}}(n_1, n - n_1)$ as n_1 goes from 0 to n, while everything else is fixed. In Figure 9 we can see some examples of this function (represented by the dashed line) for four different values of p, starting from a low p in the first plot on the left and increasing it towards the right (we also consider $n = 8$ advertisers to make the effect clearer). For $n_1 = 0$ the function gives the revenue under the CT setting (dotted line) and for $n_1 = n$ it gives the revenue under the FI setting (solid line). We see that for low p this function is decreasing and it gradually becomes increasing as p increases. While it transitions from decreasing to increasing, at some point, for medium values of p it becomes non-monotone (first decreasing and then increasing). This is the point where the IA setting can give lower revenue for the publisher than both the FI and the CT settings (Regions 4 and 5 in Figure 8).

The explanation for this is as follows. As n_1 increases from 0 to n, what we do is we move advertisers one by one from the group of direct advertisers to the group of exchange advertisers. The average of the bids in both groups is the same; therefore, the average bid is not affected as we move advertisers. However, what changes is the variance of the distribution of the bids. More specifically, the bids of the direct advertisers are more concentrated around the mean, while the bids of the exchange advertisers are more spread out. When we move the first few advertisers from the direct group to the exchange group, we make the bid distribution of the direct group slightly worse. However, the advertiser who determines the clearing price of the overall auction is still more likely in the direct group, as it has significantly more advertisers. Therefore, what happens is that as we start moving advertisers, we make the clearing price lower. However, after we reach a critical mass of advertisers in the exchange group, suddenly the clearing price will more likely be determined by the exchange group (i.e. there is a high chance that there will be at least two exchange advertisers with high b_i's). From that point onwards, as we make the exchange group larger, we increase the expected clearing price. This is the reason for the non-monotonicity of the function in the second and third plots of Figure 9. This transition phase is also what explains the existence of Regions 4 and 5 in Figure 8.

All three effects described above combined generate the six different regions we see in Figure 8.



(a) Direct advertiser. (b) Exchange advertiser.

Figure 10: Advertisers' payoffs in the independent-values case under the different information settings for different values of $p \in [0, 1]$, $n_1 = n_2 = 2$, $\kappa = 1/2$, and $G(x) = x$.

Advertisers' payoffs in the independent-values case. In Figure 10 we can see the payoffs of each type of advertiser for the three information settings and different values of p in $[0, 1]$ when the behavioral valuations of the advertisers are independent. The two plots of Figure 10 describe the more general behavior of the payoffs when there are more than two advertisers in total (in contrast to Figure 6 which was for one advertiser of each type). There are a few interesting things to note regarding the payoffs. First, in Figure 10(a), we see that for low values of p it is $D^{\text{IV}}_{\text{CT}} \leq D^{\text{IV}}_{\text{IA}} \leq D^{\text{IV}}_{\text{FI}}$, while for high values of p it is $D^{\text{IV}}_{\text{IA}} \leq D^{\text{IV}}_{\text{CT}} \leq D^{\text{IV}}_{\text{FI}}$. In other words, when p is low, a direct advertiser prefers the asymmetric setting where exchange advertisers have more information than them, over the contextual targeting setting where all advertisers have similar information. Second, in Figure 10(b), we see that for low values of p it is $E^{\text{IV}}_{\text{CT}} \leq E^{\text{IV}}_{\text{IA}} \leq E^{\text{IV}}_{\text{FI}}$, while for high values of p it is $E^{\text{IV}}_{\text{CT}} \leq E^{\text{IV}}_{\text{FI}} \leq E^{\text{IV}}_{\text{IA}}$. In other words, when p is low, an exchange advertiser prefers the full-information setting where direct advertisers have as much information as them, over the asymmetric setting where the exchange advertiser has more information than the direct advertisers.

The following result shows that these observations hold more generally.

Proposition 10. *For a uniform distribution G, any $n_1, n_2 \geq 1$, and $\kappa \geq 1/2$, when p is sufficiently low, it holds that $D^{\text{IV}}_{\text{CT}} \leq D^{\text{IV}}_{\text{IA}}$ and $E^{\text{IV}}_{\text{IA}} \leq E^{\text{IV}}_{\text{FI}}$.*

The intuition behind Proposition 10 is the following. As an advertiser, it is often advantageous for you if other advertisers gain more information than they currently have. This is because when an advertiser does not know their actual valuation they bid their expected valuation, but when p is low it is more likely than not that their actual valuation is lower than their expected valuation. In other words, when p is sufficiently low, you want the other advertisers to learn their actual valuations because then it is very likely that they will lower their bids.

6 More Robustness Checks

In this section, we check the robustness of Proposition 6 (which is the remaining result not proven analytically for the case where $n_1, n_2 > 1$, due to the lack of a closed-form general bidding function for the direct advertisers under the common-value IA setting). We first start by establishing the existence of a pure strategy symmetric equilibrium bidding function for the general case.

Lemma 3 (Advertisers' bidding behavior)**.** *For any strictly increasing and smooth distribution G, any $n_1, n_2 \geq 1$, and $\kappa \geq 1/2$, under the common-value* IA *setting, all exchange advertisers bid their true valuations and there exists a pure strategy symmetric equilibrium bidding function β for the direct advertisers satisfying $\beta(c) \in \{(1-\kappa)c\} \cup [\kappa, \kappa + (1-\kappa)c]$ for $c \in [0,1]$.*

Lemma 3 is a generalization of Lemmas 1 and 2. Based on Lemma 3, we can numerically approximate the function β for any $n_1, n_2 \geq 1$ by solving the differential equation $\frac{\partial u(\tilde{\beta};\beta,c)}{\partial \tilde{\beta}}\Big|_{\tilde{\beta}=\beta(c)} = 0$, where u is defined in equation (4). In Figure 11 we can see one example of the equilibrium bidding function when there are two exchange advertisers and two direct advertisers. Like in Lemma 1, for small contextual values c, direct advertisers underbid, while for large values of c they overbid.

Despite the lack of a closed-form bidding function, the intuition for the bidding behavior is the same as the one discussed in Section 4.1. As a result, Proposition 6 continues to hold for a large number of advertisers. In Figure 12 we can see a demonstration of this. In Figure 12(a) we consider different values of n, i.e. the total number of advertisers, and assuming that there is an

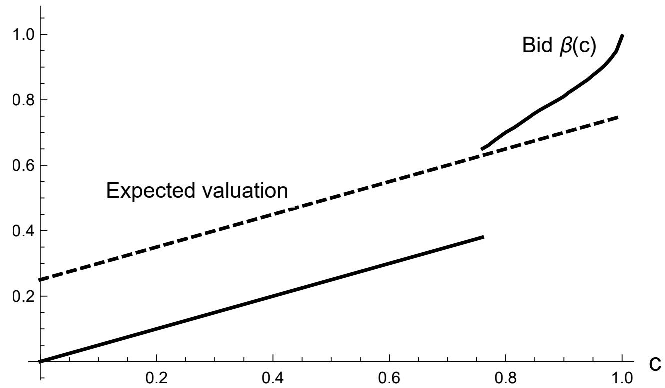
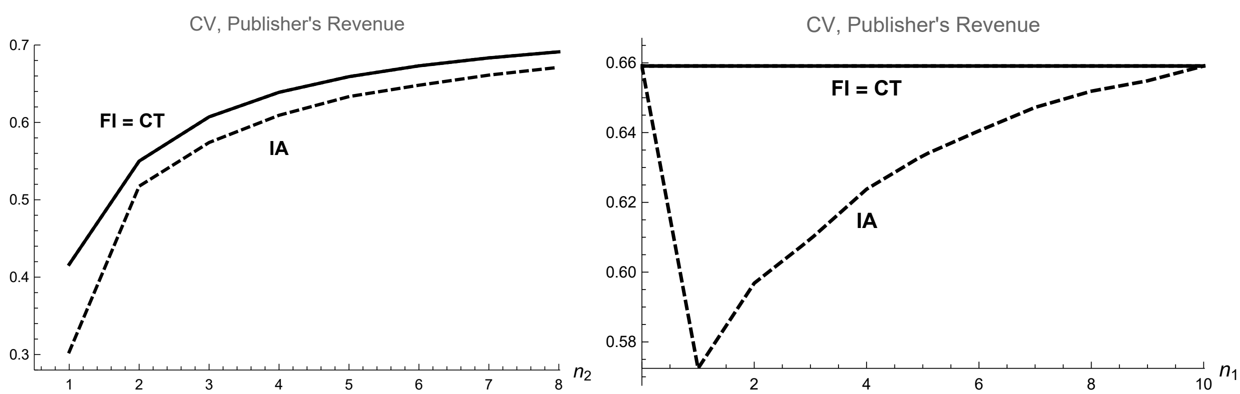


Figure 11: Bidding function of the direct advertisers (solid line) compared to their expected valuation (dashed line), for $n_1 = n_2 = 2$, $p = 1/2$, $\kappa = 1/2$, and $G(x) = x$.



(a) $n \in \{2, 4, 6, \ldots, 16\}$, $n_1 = n_2 = n/2$.

(b) $n = 10$, $n_1 \in \{0, 1, \ldots, n\}$, $n_2 = n - n_1$.

Figure 12: Publisher's revenue under the different information settings in the common-value case for different combinations of $n_1, n_2 \geq 1$, $p = 1/2$, $\kappa = 1/2$, and $G(x) = x$.

equal number of exchange and direct advertisers, we estimate the bidding function of the direct advertisers and calculate the publisher's revenue. We see that for all cases it is $W_{\text{IA}}^{\text{CV}} \leq W_{\text{FI}}^{\text{CV}} = W_{\text{CT}}^{\text{CV}}$. In Figure 12(b) we fix the total number of advertisers n and consider all different combinations of n_1 and n_2. As before, we establish that $W_{\text{IA}}^{\text{CV}} \leq W_{\text{FI}}^{\text{CV}} = W_{\text{CT}}^{\text{CV}}$ for all cases. Different choices for the number of advertisers and the other parameters generate similar plots (see also Appendix B.2 for all the key formulas used to generate the plots).

7 Conclusion

In this paper, we study the role of information asymmetry in microtargeted online advertising. Figure 13 provides a concise summary of the main results. We find that, under certain conditions, disallowing microtargeting can simultaneously increase ads' conversion rates and the publisher's revenue. This result aligns with recent anecdotal evidence and offers a potential explanation for what has been observed in practice.

Should the publisher **disable microtargeting** to maximize **conversion rate**?

	Asymmetric Information	Symmetric Information
Common Behavioral Value	**Yes** $V_{\mathrm{IA}}^{\mathrm{CV}} \leq V_{\mathrm{CT}}^{\mathrm{CV}}$	**No** $V_{\mathrm{FI}}^{\mathrm{CV}} = V_{\mathrm{CT}}^{\mathrm{CV}}$
Independent Behavioral Values	**No** $V_{\mathrm{IA}}^{\mathrm{IV}} \geq V_{\mathrm{CT}}^{\mathrm{IV}}$	**No** $V_{\mathrm{FI}}^{\mathrm{IV}} \geq V_{\mathrm{CT}}^{\mathrm{IV}}$

Should the publisher **disable microtargeting** to maximize **revenue**?

	Asymmetric Information	Symmetric Information
Common Behavioral Value	**Yes** $W_{\mathrm{IA}}^{\mathrm{CV}} \leq W_{\mathrm{CT}}^{\mathrm{CV}}$	**No** $W_{\mathrm{FI}}^{\mathrm{CV}} = W_{\mathrm{CT}}^{\mathrm{CV}}$
Independent Behavioral Values	**It depends.** (See Figure 8.)	

Figure 13: Summary of the main results.

More specifically, we demonstrate that when advertisers' behavioral valuations for a consumer are correlated, and some advertisers have access to the consumer's behavioral information while others do not, it is sometimes advantageous for the publisher, both in terms of revenue and conversion rate, to hide all behavioral data from all advertisers (i.e., to disable microtargeting and only allow targeting based on contextual information). This phenomenon does not occur if all advertisers share the same information or if the behavioral valuations of the advertisers are independent.

The rationale behind this result is that information asymmetry in a market where advertisers have correlated valuations can lead to inefficient bidding behavior. Some advertisers with high valuations might underbid due to concerns about overpaying, while others with low valuations may overbid for fear of losing valuable consumers. Both behaviors can result in a less efficient match between the consumer and the winning advertiser. Restricting access to behavioral information for all advertisers often leads to a less efficient market, because advertisers now have less information on which to base their bids. However, the level of inefficiency caused by the presence of asymmetric information is also significant; therefore, creating a level playing field among advertisers by eliminating microtargeting can lead to a more efficient market overall.

This paper has several important implications for key stakeholders within the online display advertising market. For publishers, our findings indicate that the information asymmetry among advertisers participating in advertising auctions may result in unexpected consequences on both revenue and ad conversion rates. Therefore, a careful approach towards the type and amount of information shared with advertisers is important, as more information is not always better, potentially implementing measures such as disabling third-party cookies and microtargeting for a more efficient ad market.

For advertisers, while more information about impressions they bid for is generally advantageous, less information for competitors may not always be beneficial. The bidding behavior of competitors could negatively impact them given the dynamics of the market, which underlines the need for a more fair market. Moreover, their bidding strategies may need to be flexible, potentially requiring underbidding or overbidding depending on the specific information available to them, for optimal results.

Regarding regulators, both the NPO (Edelman, 2020) and the New York Times (Davies, 2019) examples from Section 1 were responses from publishers to the General Data Protection Regulation (GDPR) in the European Union. It is interesting to note that such data privacy protection laws, initially conceived to protect consumers, can also inadvertently benefit publishers and advertisers. Therefore, careful design of these regulations can result in a win-win scenario for all parties involved.

Lastly, from a consumer standpoint, higher conversion rates usually imply more satisfactory ad content; therefore, disabling third-party tracking on a publisher's website can offer additional benefits beyond enhancing consumer privacy. It can also improve the relevance of the ads, supporting the growing trend of websites offering users the option to opt-out from tracking.

An interesting direction for future research is to examine the impact of microtargeting on consumer behavior. Consumers might prefer publishers who refrain from disclosing behavioral information to advertisers. This preference could influence publishers' decisions to enable or disable microtargeting, thereby affecting both publishers' revenue and conversion rates. If this is the case, it could serve as an additional mechanism that explains the increase in conversion rates when disallowing microtargeting.

References

Ada, S., Abou Nabout, N., and Feit, E. M. (2022). Context information can increase revenue in online display advertising auctions: Evidence from a policy change. *Journal of Marketing Research*, 59(5):1040–1058.

Athey, S. and Gans, J. S. (2010). The impact of targeting technology on advertising markets and media competition. *American Economic Review*, 100(2):608–13.

Bergemann, D. and Bonatti, A. (2011). Targeting in advertising markets: implications for offline versus online media. *The RAND Journal of Economics*, 42(3):417–443.

Brahim, N. B. E.-B., Lahmandi-Ayed, R., and Laussel, D. (2011). Is targeted advertising always beneficial? *International Journal of Industrial Organization*, 29(6):678–689.

Chandra, A. and Kaiser, U. (2014). Targeted advertising in magazine markets and the advent of the internet. *Management Science*, 60(7):1829–1843.

Chen, Y., Narasimhan, C., and Zhang, Z. J. (2001). Individual marketing with imperfect targetability. *Marketing Science*, 20(1):23–41.

Choi, H., Mela, C. F., Balseiro, S. R., and Leary, A. (2020). Online display advertising markets: A literature review and future directions. *Information Systems Research*, 31(2):556–575.

Choi, W. J., Jerath, K., and Sarvary, M. (2023). Consumer privacy choices and (un) targeted advertising along the purchase journey. *Journal of Marketing Research*, page 00222437221140052.

Choi, W. J. and Sayedi, A. (2019). Learning in online advertising. *Marketing Science*, 38(4):584–608.

Choi, W. J. and Sayedi, A. (2023a). Agency bidding in online advertising. *Available at SSRN 4324955*.

Choi, W. J. and Sayedi, A. (2023b). Open and private exchanges in display advertising. *Marketing Science*, 42(3):451–475.

Choi, W. J. and Sayedi, A. (2024). Agency market power and information disclosure in online advertising. *Marketing Science*.

Davies, J. (2019). After GDPR, The New York Times cut off ad exchanges in Europe and kept growing ad revenue. https://digiday.com/media/gumgumtest-new-york-times-gdpr-cut-off-ad-exchanges-europe-ad-revenue/.

De Corniere, A. and De Nijs, R. (2016). Online advertising and privacy. *The RAND Journal of Economics*, 47(1):48–72.

Deng, Y. and Mela, C. F. (2018). Tv viewing and advertising targeting. *Journal of Marketing Research*, 55(1):99–118.

Despotakis, S., Hafalir, I., Ravi, R., and Sayedi, A. (2017). Expertise in online markets. *Management Science*, 63(11):3895–3910.

Despotakis, S., Ravi, R., and Sayedi, A. (2021). First-price auctions in online display advertising. *Journal of Marketing Research*, 58(5):888–907.

Despotakis, S. and Yu, J. (2022). Multidimensional targeting and consumer response. *Management Science*, forthcoming.

Dong, X., Manchanda, P., and Chintagunta, P. K. (2009). Quantifying the benefits of individual-level targeting in the presence of firm strategic behavior. *Journal of Marketing Research*, 46(2):207–221.

D'Annunzio, A. and Russo, A. (2024). Intermediaries in the online advertising market. *Marketing Science*, 43(1):33–53.

Edelman, G. (2020). Can killing cookies save journalism? https://www.wired.com/story/can-killing-cookies-save-journalism/.

Esteban, L., Gil, A., and Hernandez, J. M. (2001). Informative advertising and optimal targeting in a monopoly. *The Journal of Industrial Economics*, 49(2):161–180.

Esteves, R.-B. and Resende, J. (2016). Competitive targeted advertising with price discrimination. *Marketing Science*, 35(4):576–587.

Fang, H. and Parreiras, S. O. (2003). On the failure of the linkage principle with financially constrained bidders. *Journal of Economic Theory*, 110(2):374–392.

Gal-Or, E., Gal-Or, M., May, J. H., and Spangler, W. E. (2006). Targeted advertising strategies on television. *Management Science*, 52(5):713–725.

Goldfarb, A. and Tucker, C. (2011). Online display advertising: Targeting and obtrusiveness. *Marketing Science*, 30(3):389–404.

Gordon, B. R., Jerath, K., Katona, Z., Narayanan, S., Shin, J., and Wilbur, K. C. (2021). Inefficiencies in digital advertising markets. *Journal of Marketing*, 85(1):7–25.

Hsiao, S. (2020). How our display buying platforms share revenue with publishers. https://blog.google/products/admanager/display-buying-share-revenue-publishers.

Hummel, P. and McAfee, R. P. (2016). When does improved targeting increase revenue? *ACM Transactions on Economics and Computation (TEAC)*, 5(1):1–29.

Iyer, G., Soberman, D., and Villas-Boas, J. M. (2005). The targeting of advertising. *Marketing Science*, 24(3):461–476.

Johnson, J. P. (2013). Targeted advertising and advertising avoidance. *The RAND Journal of Economics*, 44(1):128–144.

Krishna, V. (2009). *Auction theory*. Academic press.

Lee, D., Hosanagar, K., and Nair, H. S. (2018). Advertising content and consumer engagement on social media: Evidence from facebook. *Management Science*, 64(11):5105–5131.

Levin, J. and Milgrom, P. (2010). Online advertising: Heterogeneity and conflation in market design. *American Economic Review*, 100(2):603–07.

Long, F., Jerath, K., and Sarvary, M. (2022). Designing an online retail marketplace: Leveraging information from sponsored advertising. *Marketing Science*, 41(1):115–138.

McAfee, R. P. and McMillan, J. (1987). Auctions and bidding. *Journal of economic literature*, 25(2):699–738.

Milgrom, P. R. and Weber, R. J. (1982). A theory of auctions and competitive bidding. *Econometrica: Journal of the Econometric Society*, pages 1089–1122.

Perry, M. and Reny, P. J. (1999). On the failure of the linkage principle in multi-unit auctions. *Econometrica*, 67(4):895–900.

Rafieian, O. and Yoganarasimhan, H. (2021). Targeting and privacy in mobile advertising. *Marketing Science*, 40(2):193–218.

Ravichandran, D. and Korula, N. (2019). Effect of disabling third-party cookies on publisher revenue. https://services.google.com/fh/files/misc/disabling_third-party_cookies_publisher_revenue.pdf.

Roy, S. (2000). Strategic segmentation of a market. *International Journal of Industrial Organization*, 18(8):1279–1290.

Sayedi, A. (2018). Real-time bidding in online display advertising. *Marketing Science*, 37(4):553–568.

Shen, Q. and Miguel Villas-Boas, J. (2018). Behavior-based advertising. *Management Science*, 64(5):2047–2064.

Shin, J. and Shin, W. (2022). A theory of irrelevant advertising: An agency-induced targeting inefficiency. *Management Science*.

Shin, J. and Yu, J. (2021). Targeted advertising and consumer inference. *Marketing Science*, 40(5):900–922.

Snelders, E., Worp, L., and Sheng, S. (2020). A future without advertising cookies? https://www.ster.nl/media/h5ehvtx3/ster_a-future-without-advertising-cookies.pdf.

Summers, C. A., Smith, R. W., and Reczek, R. W. (2016). An audience of one: Behaviorally targeted ads as implied social labels. *Journal of Consumer Research*, 43(1):156–178.

Trusov, M., Ma, L., and Jamal, Z. (2016). Crumbs of the cookie: User profiling in customer-base analysis and behavioral targeting. *Marketing Science*, 35(3):405–426.

Tucker, C. E. (2014). Social networks, personalized advertising, and privacy controls. *Journal of marketing research*, 51(5):546–562.

Zhang, K. and Katona, Z. (2012). Contextual advertising. *Marketing Science*, 31(6):980–994.

Zia, M. and Rao, R. C. (2019). Search advertising: Budget allocation across search engines. *Marketing Science*, 38(6):1023–1037.

A Appendix

A.1 Proofs of Lemmas 1–2 and Propositions 1–5

Proof of Lemma 1

The direct advertiser's expected utility when their contextual value is c and they bid β is:

$$u(\beta, c) := p(1-\kappa)\int_0^{\max\left\{\frac{\beta-\kappa}{1-\kappa},0\right\}}(c-c')d(G(c')^{n_1}) + (1-p)(1-\kappa)\int_0^{\min\left\{\frac{\beta}{1-\kappa},1\right\}}(c-c')d(G(c')^{n_1}).$$

Suppose that $\kappa < 1-\kappa$. First, let us consider the case when $0 \leq \beta < \kappa$, we have $u(\beta, c) = (1-p)(1-\kappa)\int_0^{\frac{\beta}{1-\kappa}}(c-c')d(G(c')^{n_1})$, which means, $\frac{\partial u}{\partial \beta} = n_1(1-p)\left(c - \frac{\beta}{1-\kappa}\right)G\left(\frac{\beta}{1-\kappa}\right)^{n_1-1}G'\left(\frac{\beta}{1-\kappa}\right) = 0 \implies \beta = (1-\kappa)c$. For, the case when $\kappa < \beta \leq 1-\kappa$, we have $u(\beta, c) = p(1-\kappa)\int_0^{\frac{\beta-\kappa}{1-\kappa}}(c-c')d(G(c')^{n_1}) + (1-p)(1-\kappa)\int_0^{\frac{\beta}{1-\kappa}}(c-c')d(G(c')^{n_1})$, which means, $\frac{\partial u}{\partial \beta} = n_1 p\left(c - \frac{\beta-\kappa}{1-\kappa}\right)G\left(\frac{\beta-\kappa}{1-\kappa}\right)^{n_1-1}G'\left(\frac{\beta-\kappa}{1-\kappa}\right) + n_1(1-p)\left(c - \frac{\beta}{1-\kappa}\right)G\left(\frac{\beta}{1-\kappa}\right)^{n_1-1}G'\left(\frac{\beta}{1-\kappa}\right)$. When $n_1 = 1$ and $G(x) = x$, $\frac{\partial u}{\partial \beta} = 0$ implies that $\beta = \kappa p + (1-\kappa)c$. For the case when $1-\kappa < \beta \leq 1$, we have $u(\beta, c) = p(1-\kappa)\int_0^{\frac{\beta-\kappa}{1-\kappa}}(c-c')d(G(c')^{n_1}) + (1-p)(1-\kappa)\int_0^1(c-c')d(G(c')^{n_1})$, which means, $\frac{\partial u}{\partial \beta} = n_1 p\left(c - \frac{\beta-\kappa}{1-\kappa}\right)G\left(\frac{\beta-\kappa}{1-\kappa}\right)^{n_1-1}G'\left(\frac{\beta-\kappa}{1-\kappa}\right) = 0 \implies \beta = \kappa + (1-\kappa)c$. The global maximum of u occurs either at $\beta = (1-\kappa)c, \kappa p + (1-\kappa)c, \kappa + (1-\kappa)c$, or at one of the singular points $\beta = \kappa, 1-\kappa$. Let

$$u_1(c) := u(\beta = (1-\kappa)c, c) = \begin{cases} \frac{1}{2}(1-p)(1-\kappa)c^2, & c \leq \frac{\kappa}{1-\kappa}, \\ \frac{1}{2}(1-\kappa)c^2 - \frac{p}{2}\left(\frac{\kappa^2}{1-\kappa}\right), & \frac{\kappa}{1-\kappa} < c \leq 1, \end{cases}$$

$$u_2(c) := u(\beta = \kappa p + (1-\kappa)c, c) = \begin{cases} \frac{1}{2}(1-p)(1-\kappa)c^2 - \frac{1}{2}(1-p)p^2\left(\frac{\kappa^2}{1-\kappa}\right), & c \leq \frac{(1-p)\kappa}{1-\kappa}, \\ \frac{1}{2}(1-\kappa)c^2 - \frac{1}{2}(1-p)p\left(\frac{\kappa^2}{1-\kappa}\right), & \frac{(1-p)\kappa}{1-\kappa} < c \leq 1 - \frac{p\kappa}{1-\kappa}, \\ \frac{1}{2}p(1-\kappa)c^2 - \frac{1}{2}p(1-p)^2\left(\frac{\kappa^2}{1-\kappa}\right) & \\ \qquad +(1-p)(1-\kappa)\left(c - \frac{1}{2}\right), & 1 - \frac{p\kappa}{1-\kappa} < c \leq 1, \end{cases}$$

$$u_3(c) := u(\beta = \kappa + (1-\kappa)c, c) + \begin{cases} \frac{1}{2}(1-\kappa)c^2 - \frac{1-p}{2}\left(\frac{\kappa^2}{1-\kappa}\right), & c \leq 1 - \frac{\kappa}{1-\kappa}, \\ \frac{1}{2}p(1-\kappa)c^2 + (1-p)(1-\kappa)\left(c - \frac{1}{2}\right), & 1 - \frac{\kappa}{1-\kappa} < c \leq 1, \end{cases}$$

$$u_4(c) := u(\beta = \kappa, c) = (1-p)(1-\kappa)\int_0^{\frac{\kappa}{1-\kappa}}(c-c')d(G(c')^{n_1}) = (1-p)\kappa c - \frac{1}{2}(1-p)\left(\frac{\kappa^2}{1-\kappa}\right),$$

$$\begin{aligned} u_5(c) := u(\beta = 1-\kappa, c) &= p(1-\kappa)\int_0^{\frac{1-2\kappa}{1-\kappa}}(c-c')d(G(c')^{n_1}) + (1-p)(1-\kappa)\int_0^1(c-c')d(G(c')^{n_1}) \\ &= p(1-2\kappa)c - \frac{1}{2}p\left(\frac{(1-2\kappa)^2}{1-\kappa}\right) + (1-p)(1-\kappa)\left(c - \frac{1}{2}\right). \end{aligned}$$

Clearly, $u_4(c) \leq u_1(c)$ (in fact u_4 is tangent to $\frac{1}{2}(1-p)(1-\kappa)c^2$ at $c = \frac{\kappa}{1-\kappa}$) and $u_5(c) \leq u_3(c)$ (in fact u_5 is tangent to $\frac{1}{2}p(1-\kappa)c^2 + (1-p)(1-\kappa)\left(c - \frac{1}{2}\right)$ at $c = 1 - \frac{\kappa}{1-\kappa}$), so we can ignore u_4 and u_5. Then $\beta = (1-\kappa)c$ when $u_1(c) > u_2(c), u_3(c)$, $\beta = \kappa p + (1-\kappa)c$ when $u_2(c) > u_1(c), u_3(c)$, and $\beta = \kappa + (1-\kappa)c$ when $u_3(c) > u_1(c), u_2(c)$, and we break ties arbitrarily. We note that u_1, u_2, u_3 are all continuous in c and that $\frac{du_1}{dc} \leq \frac{du_2}{dc} \leq \frac{du_3}{dc}$, therefore u_1 can only be overtaken by

u_2, u_3 and u_2 can only be overtaken by u_3, and u_3 cannot be overtaken. So, for $\kappa < 1-\kappa$, there must exist $\underline{c}$ and $\bar{c}$ such that $\beta(c) = (1-\kappa)c$ if $c < \underline{c}$, $\beta(c) = \kappa p + (1-\kappa)c$ if $\underline{c} < c < \bar{c}$, and $\beta(c) = \kappa + (1-\kappa)c$. Let us now find c_{12}, the point where u_2 overtakes u_1, suppose that $\frac{(1-p)\kappa}{1-\kappa} \leq c_{12} \leq \frac{\kappa}{1-\kappa}$: $\frac{1}{2}(1-p)(1-\kappa)c_{12}^2 = \frac{1}{2}(1-\kappa)c_{12}^2 - \frac{1}{2}(1-p)p\left(\frac{\kappa^2}{1-\kappa}\right) \implies c_{12} = \frac{\sqrt{1-p}\kappa}{1-\kappa} \in \left[\frac{(1-p)\kappa}{1-\kappa}, \frac{\kappa}{1-\kappa}\right]$. We do not need to further check other intervals due to the uniqueness of the intersection point. Similarly, we find the location of the point c_{23} where u_3 overtakes u_2: $1 - \frac{\kappa}{1-\kappa} \leq c_{23} \leq 1 - \frac{p\kappa}{1-\kappa}$: $\frac{1}{2}p(1-\kappa)c_{23}^2 + (1-p)(1-\kappa)\left(c_{23} - \frac{1}{2}\right) = \frac{1}{2}(1-\kappa)c_{23}^2 - \frac{1}{2}(1-p)p\left(\frac{\kappa^2}{1-\kappa}\right)$. This is a quadratic equation in c_{23} that has roots: $1 \pm \frac{\sqrt{p}\kappa}{1-\kappa}$. We take the negative root $c_{23} = 1 - \frac{\sqrt{p}\kappa}{1-\kappa} \in \left[1 - \frac{\kappa}{1-\kappa}, 1 - \frac{p\kappa}{1-\kappa}\right]$. Finally, we consider the point c_{13} where u_3 overtakes u_1, suppose that $1 - \frac{\kappa}{1-\kappa} < c_{13} < \frac{\kappa}{1-\kappa}$: $\frac{1}{2}(1-p)(1-\kappa)c_{13}^2 = \frac{1}{2}p(1-\kappa)c_{13}^2 + (1-p)(1-\kappa)\left(c_{13} - \frac{1}{2}\right)$. This is a quadratic equation in c_{13} with two roots: $\frac{\sqrt{1-p}}{\sqrt{1-p} \pm \sqrt{p}}$. Since $c_{13} \in [0,1]$, we take the positive root: $c_{13} = \frac{\sqrt{1-p}}{\sqrt{1-p}+\sqrt{p}}$. Finally, we take $\underline{c} := \min\{c_{12}, c_{13}\}$, and $\bar{c} := \max\{c_{13}, c_{23}\}$, this also ensures that c_{13} is relevant only if $1 - \frac{\kappa}{1-\kappa} < c_{23} < c_{13} < c_{12} < \frac{\kappa}{1-\kappa}$.

Now, suppose that $\kappa \geq 1-\kappa$. Let us consider the case when $0 \leq \beta \leq 1-\kappa$, we have $u(\beta, c) = (1-p)(1-\kappa)\int_0^{\frac{\beta}{1-\kappa}}(c-c')d(G(c')^{n_1})$, as before, $\frac{\partial u}{\partial c} = 0$ implies $\beta = (1-\kappa)c$. For $1-\kappa \leq \beta \leq \kappa$, we find that $u(\beta, c) = (1-p)(1-\kappa)\int_0^1(c-c')d(G(c')^{n_1})$, which is a constant in β. For $\kappa < \beta \leq 1$, we have $u(\beta, c) = p(1-\kappa)\int_0^{\frac{\beta-\kappa}{1-\kappa}}(c-c')d(G(c')^{n_1}) + (1-p)(1-\kappa)\int_0^1(c-c')d(G(c')^{n_1})$, which means $\frac{\partial u}{\partial \beta} = 0$ implies $\beta = \kappa + (1-\kappa)c$. This time we let

$$u_1(c) := u(\beta = (1-\kappa)c, c) = \frac{1}{2}(1-p)(1-\kappa)c^2,$$

$$u_2(c) := u(\beta = \kappa + (1-\kappa)c, c) = \frac{1}{2}p(1-\kappa)c^2 + (1-p)(1-\kappa)\left(c - \frac{1}{2}\right).$$

And, as before, $u_3(c) := u(\beta = \kappa, c)$, $u_4(c) := u(\beta = 1-\kappa, c)$, which we can check that they satisfy $u_3(c) \leq u_1(c)$ and $u_4(c) \leq u_2(c)$, so we can ignore them. Since $\frac{du_1}{dc} = (1-p)(1-\kappa) \leq \frac{du_2}{dc} = (1-\kappa)c$, we conclude that u_2 can only overtake u_1 and cannot be overtaken. Hence, there exists $\underline{c} = \bar{c}$ such that $\beta(c) = (1-\kappa)c$ if $c < \underline{c}$ and $\beta(c) = \kappa + (1-\kappa)c$ if $c > \bar{c}$. Further inspection reveals that $\underline{c} = \bar{c} = c_{13} = \frac{\sqrt{1-p}}{\sqrt{1-p}+\sqrt{p}}$ as previously found. This completes the proof. ■

Proofs of Propositions 1, 2, 3, and 4

Propositions 1, 2, 3, and 4 are special cases of Propositions 5, 6, 7, and 8. We present the proofs of the more general statements next. ■

Proof of Lemma 2

When $\kappa \geq 1/2$, we have $\kappa \geq 1-\kappa$, and we only need to consider two cases: $0 \leq \beta \leq 1-\kappa$ where the direct advertiser expected utility is $u(\beta, c) = (1-p)(1-\kappa)\int_0^{\frac{\beta}{1-\kappa}}(c-c')d(G(c')^{n_1})$ and $\kappa \leq \beta \leq 1$ where the direct advertiser expected utility is $u(\beta, c) = p(1-\kappa)\int_0^{\frac{\beta-\kappa}{1-\kappa}}(c-c')d(G(c')^{n_1}) + (1-p)(1-\kappa)\int_0^1(c-c')d(G(c')^{n_1})$. We do not need to consider the $1-\kappa < \beta < \kappa$ case since $u(\beta, c) = (1-p)(1-\kappa)\int_0^1(c-c')d(G(c')^{n_1})$ is constant in β over that domain. It follows that if $0 \leq \beta \leq 1-\kappa$, then $\frac{\partial u}{\partial \beta} = n_1(1-p)\left(c - \frac{\beta}{1-\kappa}\right)G\left(\frac{\beta}{1-\kappa}\right)^{n_1-1}G'\left(\frac{\beta}{1-\kappa}\right) = 0 \implies \beta = (1-\kappa)c$. Similarly, if $\kappa \leq \beta \leq 1$, then $\frac{\partial u}{\partial \beta} = n_1 p\left(c - \frac{\beta-\kappa}{1-\kappa}\right)G\left(\frac{\beta-\kappa}{1-\kappa}\right)^{n_1-1}G'\left(\frac{\beta-\kappa}{1-\kappa}\right) = 0 \implies \beta = \kappa + (1-\kappa)c$. For any fixed c, the global maximum of u occurs either at $\beta = (1-\kappa)c, \kappa + (1-\kappa)c$ or at one of the singular points $\beta = \kappa, 1-\kappa$. Let

$$u_1(c) := u(\beta = (1-\kappa)c, c) = (1-p)(1-\kappa)\int_0^c (c-c')d(G(c')^{n_1}),$$

$$u_2(c) := u(\beta = \kappa + (1-\kappa)c, c) = p(1-\kappa)\int_0^c (c-c')d(G(c')^{n_1}) + (1-p)(1-\kappa)\int_0^1 (c-c')d(G(c')^{n_1}),$$

$$u_3(c) := u(\beta = \kappa, c) = (1-p)(1-\kappa)\int_0^{\frac{\kappa}{1-\kappa}} (c-c')d(G(c')^{n_1}),$$

$$u_4(c) := u(\beta = 1-\kappa, c) = p(1-\kappa)\int_0^{1-\frac{\kappa}{1-\kappa}} (c-c')d(G(c')^{n_1}) + (1-p)(1-\kappa)\int_0^1 (c-c')d(G(c')^{n_1}).$$

Since $u_3(c)$ is the tangent line to $u_1(c)$ at $c = \frac{\kappa}{1-\kappa}$ and $u_4(c)$ is the tangent line to $u_2(c)$ at $c = 1 - \frac{\kappa}{1-\kappa}$, and both u_1, u_2 are convex, we have $u_3 \leq u_1$ and $u_4 \leq u_2$, so we can ignore u_3, u_4. Next, we note that $\frac{du_1}{dc} = (1-p)(1-\kappa)G(c)^{n_1} < p(1-\kappa)G(c)^{n_1} + (1-p)(1-\kappa) = \frac{du_2}{dc}$ for all $c \in [0,1]$. We conclude that u_1 can only be overtaken by u_2. Note also that $u_1(0) = 0 > u_2(0) = -(1-p)(1-\kappa)\int_0^1 c' d(G(c')^{n_1})$ and $u_2(1) = p(1-\kappa)\int_0^c(1-c')d(G(c')^{n_1}) + u_1(1) > u_1(1)$, so the intersection point $\underline{c}(p) \in [0,1]$ exists and is unique. For a given distribution G, we can find $\underline{c}$ from the relation $u_1(\underline{c}) = u_2(\underline{c})$. Equivalently, the relation for $\underline{c}$ may be written as

$$\int_0^1 (\underline{c} - c')d(G(c')^{n_1}) = \frac{1-2p}{1-p}\int_0^{\underline{c}} (\underline{c} - c')d(G(c')^{n_1}). \tag{2}$$

Clearly, $1-\kappa$ cancels out and $\underline{c}$ is independent of κ. Furthermore, $u_1 - u_2$ is continuously differentiable in p and in $\underline{c}$ with nonvanishing derivative, and hence $\underline{c}$ is continuously differentiable in p by the Implicit Function Theorem. Differentiating $u_1 - u_2 = 0$ with respect to p, we get $-\frac{1}{1-p}\int_0^{\underline{c}}(\underline{c}-c')d(G(c')^{n_1}) = \left[(1-p)(1-G(\underline{c})^{n_1}) + pG(\underline{c})^{n_1}\right]\frac{d\underline{c}}{dp}$. The factor in the square bracket is positive and also $\int_0^{\underline{c}}(\underline{c}-c')d(G(c')^{n_1}) > 0$, hence $\frac{d\underline{c}}{dp} < 0$.

From (2) we can see that $p = 0$ implies $\int_{\underline{c}}^1(\underline{c}-c')d(G(c')^{n_1}) = 0$, which holds exactly if $\underline{c}(0) = 1$ as the integral is < 0 for all $\underline{c} < 1$. Similarly, we can see that the LHS of (2) is bounded in $[-1,1]$, while the RHS approaches $-\infty$ as $p \to 1^-$, unless $\underline{c} \to 0$, which must be the case. Hence $\underline{c}(1) = 0$. Lastly, the RHS of (2) vanishes when $p = 1/2$, therefore, we are left with $\int_0^1(\underline{c}-c')d(G(c')^{n_1}) = 0$ or $\underline{c} = \int_0^1 c'd(G(c')^{n_1}) = \mathbb{E}[n_1 cG(c)^{n_1-1}]$, as claimed. ■

Proof of Proposition 5

The statement of the proposition holds in a more general setting, which we prove in Lemma 4.

Lemma 4. *Consider any auction mechanism M such that, whenever bidders are symmetric and independent, in equilibrium M allocates the impression to the highest-valuation bidder. Suppose bidder i's valuation is given by some function of random variables corresponding to behavioral and contextual data: $v_i = v(b_i, c_i)$, where we assume v is increasing in c_i. Then under the common-value case $b_1 = b_2 = \cdots =: b$, for any distribution G, any $n_1, n_2 \geq 1$, and with selling mechanism M, we have $V_{\mathrm{IA}}^{\mathrm{CV}} \leq V_{\mathrm{FI}}^{\mathrm{CV}} = V_{\mathrm{CT}}^{\mathrm{CV}}$.*

Proof. Under both the full-information and the contextual-targeting settings, when the behavioral value b is common among all the bidders, we have that the bidders are symmetric with their valuations determined by the independently drawn contextual values c_i. Therefore, the impression is allocated to the bidder with the highest c_i under M. Under full information, the valuation of any bidder i is $v_i = v(b_i, c_i)$. Under the contextual-targeting setting, the expected valuation of any bidder i is $\mathbb{E}[v_i] = \mathbb{E}[v(b_i, c_i)|c_i]$. It follows that the expected conversion rate is

$$V_{\mathrm{FI}}^{\mathrm{CV}} = \mathbb{E}_b\left[(n_1+n_2)\int_0^1 v(b,c)G(c)^{n_1+n_2-1}G'(c)dc\right] = (n_1+n_2)\int_0^1 \mathbb{E}\left[v(b,c)|c\right]G(c)^{n_1+n_2-1}G'(c)dc = V_{\mathrm{CT}}^{\mathrm{CV}},$$

where in the second equality we applied Fubini's Theorem. Since the mechanism M under full information ensures that the bidder with the highest valuation will win, it must be the case that

$V^{\mathrm{CV}}_{\mathrm{FI}}$ is the highest possible conversion rate under any information setting. In particular, $V^{\mathrm{CV}}_{\mathrm{IA}} \leq V^{\mathrm{CV}}_{\mathrm{FI}} = V^{\mathrm{CV}}_{\mathrm{CT}}$. $\square$

Proposition 5 follows by applying Lemma 4 to the case where M is a second-price auction and $v(b_i, c_i) := \kappa b_i + (1-\kappa)c_i$. Since $b_i = 1$ with probability p and $b_i = 0$ with probability $1-p$, it also follows that $\mathbb{E}[v(b_i, c_i)|c_i] = \kappa p + (1-\kappa)c_i$. $\blacksquare$

B Online Appendix

B.1 Proofs of Propositions 6–9 and Lemma 3

Proof of Proposition 6

The fact that $W_{\mathrm{CT}}^{\mathrm{CV}} = W_{\mathrm{FI}}^{\mathrm{CV}}$ is general and can be seen by directly comparing their expressions. For $n_1 = n_2 = 1$ and $\kappa \geq 1/2$ we may simplify (5) to:

$$
\begin{aligned}
W_{\mathrm{IA}}^{\mathrm{CV}} =& 2p\int_{\underline{c}}^{1}(\kappa+(1-\kappa)c)(1-G(c))G'(c)dc + 2(1-p)\int_{0}^{\underline{c}}(1-\kappa)c(1-G(c))G'(c)dc \\
&+ p\int_{0}^{\underline{c}}(1-\kappa)cG'(c)dc + p\int_{0}^{\underline{c}}(\kappa+(1-\kappa)c)(1-G(\underline{c}))G'(c)dc + (1-p)\int_{\underline{c}}^{1}(1-\kappa)c(1-G(\underline{c}))G'(c)dc \\
=& W_{\mathrm{FI}}^{\mathrm{CV}} + p\Big(\int_{0}^{\underline{c}}(\kappa+(1-\kappa)c)G(c)G'(c)dc - \int_{0}^{\underline{c}}(\kappa+(1-\kappa)c)(G(\underline{c})-G(c))G'(c)dc - \kappa G(\underline{c})\Big) \\
&+ (1-p)\Big(\int_{\underline{c}}^{1}(1-\kappa)c(G(c)-G(\underline{c}))G'(c)dc - \int_{\underline{c}}^{1}(1-\kappa)c(1-G(c))G'(c)dc\Big) \\
=:& W_{\mathrm{FI}}^{\mathrm{CV}} + W_{\Delta},
\end{aligned}
$$

where W_{Δ} is defined to be the sum of the first and the second bracket. Let us show that $W_{\Delta} \leq 0$ for all $p \in [0,1]$. First, we note that (2) can be written equivalently as

$$
pn_1\int_{0}^{\underline{c}} cG(c)^{n_1-1}G'(c)dc + (1-p)n_1\int_{\underline{c}}^{1} cG(c)^{n_1-1}G'(c)dc = \underline{c}pG(\underline{c})^{n_1} + \underline{c}(1-p)(1-G(\underline{c})^{n_1}). \quad (3)
$$

Then, we have

$$
\begin{aligned}
W_{\Delta} =& \frac{1}{2}\kappa pG(\underline{c})^2 + (1-\kappa)p\int_{0}^{\underline{c}} cG(c)G'(c)dc \\
&- \kappa pG(\underline{c})^2 - (1-\kappa)pG(\underline{c})\int_{0}^{\underline{c}} cG'(c)dc + \frac{1}{2}\kappa pG(\underline{c})^2 + (1-\kappa)p\int_{0}^{\underline{c}} cG(c)G'(c)dc - \kappa pG(\underline{c}) \\
&+ (1-\kappa)(1-p)\int_{\underline{c}}^{1} cG(c)G'(c)dc - (1-\kappa)(1-p)G(\underline{c})\int_{\underline{c}}^{1} cG'(c)dc - (1-\kappa)(1-p)\int_{\underline{c}}^{1} c(1-G(c))G'(c)dc \\
=& 2(1-\kappa)p\int_{0}^{\underline{c}} cG(c)G'(c)dc + (1-\kappa)(1-p)\int_{\underline{c}}^{1} cG(c)G'(c)dc \\
&- (1-\kappa)G(\underline{c})p\int_{0}^{\underline{c}} cG'(c)dc - (1-\kappa)G(\underline{c})(1-p)\int_{\underline{c}}^{1} cG'(c)dc \\
&- (1-\kappa)(1-p)\int_{\underline{c}}^{1} c(1-G(c))G'(c)dc - \kappa pG(\underline{c}) \\
\leq& (1-\kappa)p\underline{c}G(\underline{c})^2 + (1-\kappa)(1-p)\int_{\underline{c}}^{1} cG(c)G'(c)dc \\
&- (1-\kappa)G(\underline{c})\left(\underline{c}pG(\underline{c}) + \underline{c}(1-p)(1-G(\underline{c}))\right) - (1-\kappa)(1-p)\int_{\underline{c}}^{1} c(1-G(c))G'(c)dc \\
&+ (1-\kappa)(1-p)\underline{c}(1-G(\underline{c})) - (1-\kappa)p\int_{0}^{\underline{c}} cG'(c)dc - (1-\kappa)(1-p)\int_{\underline{c}}^{1} cG'(c)dc.
\end{aligned}
$$

The last inequality can be explained as follows. We rewrite the first line using the inequalities: $2(1-\kappa)p\int_0^{\underline{c}} cG(c)G'(c)dc \le (1-\kappa)p\underline{c}\int_0^{\underline{c}} d(G(c)^2) = (1-\kappa)p\underline{c}G(\underline{c})^2$. The second line follows from (3) with $n_1 = 1$. Lastly, we rewrite $\kappa pG(\underline{c})$ using (3) and the fact that $1-\kappa \le \kappa$, $\underline{c} \le 1$: $\kappa pG(\underline{c}) \ge (1-\kappa)p\underline{c}G(\underline{c}) = (1-\kappa)p\int_0^{\underline{c}} cG'(c)dc + (1-\kappa)(1-p)\int_{\underline{c}}^1 cG'(c)d - (1-\kappa)\underline{c}(1-p)(1-G(\underline{c}))$.

Back to the main calculation, after some cancellations, the last inequality becomes:

$$\begin{aligned}
W_\Delta \le& (1-\kappa)(1-p)\underline{c}(1-G(\underline{c}))^2 - 2(1-\kappa)(1-p)\int_{\underline{c}}^1 c(1-G(c))G'(c)dc - (1-\kappa)p\int_0^{\underline{c}} cG'(c)dc \\
\le& (1-\kappa)(1-p)\underline{c}(1-G(\underline{c}))^2 - 2(1-\kappa)(1-p)\underline{c}\int_{\underline{c}}^1 (1-G(c))G'(c)dc - (1-\kappa)p\int_0^{\underline{c}} cG'(c)dc \\
=& -(1-\kappa)p\int_0^{\underline{c}} cG'(c)dc \le 0.
\end{aligned}$$

Therefore, we have that $W_\Delta \le 0$ as needed. In fact, we can see from $W_\Delta \le -(1-\kappa)p\int_0^{\underline{c}} cG'(c)dc$, that the equality holds exactly when $p = 0, 1$, i.e. $W_{\mathrm{IA}}^{\mathrm{CV}}|_{p=0,1} = W_{\mathrm{FI}}^{\mathrm{CV}}|_{p=0,1} = W_{\mathrm{CT}}^{\mathrm{CV}}|_{p=0,1}$. ■

Proof of Proposition 7

Let's consider $N := n_1 + n_2$ advertisers which are divided into two disjoint subsets $A = \{a_1^A, \cdots, a_{n_1}^A\}$ and $B = \{a_1^B, \cdots, a_{n_2}^B\}$, $A \coprod B = \{1, 2, \cdots, n_1 + n_2\}$. Set A contains exchange advertisers with full information and hence bid their true valuation. Set B contains direct advertisers with only the contextual value, and hence bid the expected value $\kappa p + (1-\kappa)c$. Let's consider an instance of an auction where the contextual values in set B are given by $c_1 > \cdots > c_{n_2}$, whereas the highest bid in A is given by the bidder a^* with valuation $\kappa b^* + (1-\kappa)c^*$. Independently, we also draw $b_1, \cdots, b_{n_2}$ behavioral values for the direct advertisers in B.

First, we consider the case where a_1^B from B is the winner: $\kappa p + (1-\kappa)c_1 > \kappa b^* + (1-\kappa)c^*$. Suppose that we moved an advertiser $a_i^B \ne a_1^B$ from set B to set A, keeping all the contextual values fixed. After the move, either a_i^B becomes the winner or nothing changes. Suppose a_i^B becomes the winner, this means $b_i = 1$, $\kappa + (1-\kappa)c_i > \kappa b^* + (1-\kappa)c^*$ and $\kappa + (1-\kappa)c_i > \kappa p + (1-\kappa)c_1 \implies c_1 - c_i < \frac{\kappa(1-p)}{1-\kappa}$.

With probability p we have $b_1 = 1$, and in this case, the change in the winner's valuation Δv_w is given by $\mathbb{E}[\Delta v_w | b_1 = 1] = (\kappa + (1-\kappa)c_i) - (\kappa + (1-\kappa)c_1) = -(1-\kappa)(c_1 - c_i) > -\kappa(1-p)$.

With probability $1-p$ we have $b_1 = 0$, and in this case, the change in the winner's valuation is given by $\mathbb{E}[\Delta v_w | b_1 = 0] = (\kappa + (1-\kappa)c_i) - (1-\kappa)c_1 = \kappa - (1-\kappa)(c_1 - c_i) > \kappa - \kappa(1-p) = \kappa p$.

Therefore, the expected change of the winner's valuation is $\mathbb{E}[\Delta v_w] = \mathbb{E}[\Delta v_w | b_1 = 1]\mathbb{P}[b_1 = 1] + \mathbb{E}[\Delta v_w | b_1 = 0]\mathbb{P}[b_1 = 0] > -p \cdot \kappa(1-p) + (1-p) \cdot \kappa p = 0$.

Suppose that we moved an advertiser a_1^B from set B to A, keeping all the drawn contextual values fixed. After the move, either a_1^B remains the winner, hence nothing changes, or it is not. If a_1^B is no longer a winner, then either a^* is the winner, in that case, we have an increase in the winner's valuation since $\kappa b^* + (1-\kappa)c^* > \kappa b_1 + (1-\kappa)c_1$. Otherwise, a_2^B is now the winner, so we must have $b_1 = 0$ and $\kappa p + (1-\kappa)c_2 > (1-\kappa)c_1 \implies c_1 - c_2 < \frac{\kappa p}{1-\kappa}$.

With probability p we have $b_2 = 1$, and in this case, the change in the winner's valuation is given by $\mathbb{E}[\Delta v_w | b_2 = 1] = (\kappa + (1-\kappa)c_2) - (1-\kappa)c_1 = \kappa - (1-\kappa)(c_1 - c_2) > \kappa - \kappa p = \kappa(1-p)$.

With probability $1-p$ we have $b_2 = 0$, and in this case, the change in the winner's valuation is given by $\mathbb{E}[\Delta v_w | b_2 = 0] = (1-\kappa)c_2 - (1-\kappa)c_1 = -(1-\kappa)(c_1 - c_2) > -\kappa p$.

Therefore, the expected change of the winner's valuation is $\mathbb{E}[\Delta v_w] = \mathbb{E}[\Delta v_w | b_1 = 1]\mathbb{P}[b_1 = 1] + \mathbb{E}[\Delta v_w | b_1 = 0]\mathbb{P}[b_1 = 0] > p \cdot \kappa(1-p) - (1-p) \cdot \kappa p = 0$.

Now, we consider the case where a^* is the winner: $\kappa b^* + (1-\kappa)c^* > \kappa p + (1-\kappa)c_1$. If a winner changed by moving an a_i^B advertiser from the set B to the set A, keeping all the drawn contextual and behavioral values fixed, then the moved advertiser must have $\kappa b_i + (1-\kappa)c_i > \kappa b^* + (1-\kappa)c^*$. Therefore, the winner's valuation can only increase in this case.

It follows that the conversion rate increases or remains the same for every advertiser we move from set B to set A. We conclude that $V_{\text{FI}}^{\text{CV}} \geq V_{\text{IA}}^{\text{CV}} \geq V_{\text{CT}}^{\text{CV}}$. ∎

Proof of Proposition 8

Let's denote by $w_{\text{FI}}^{\text{IV}}, w_{\text{IA}}^{\text{IV}}, w_{\text{CT}}^{\text{IV}}$ the revenue under each information setting for an instant of auction, so that we have $W_{\text{FI}}^{\text{IV}} := \mathbb{E}[w_{\text{FI}}^{\text{IV}}], W_{\text{IA}}^{\text{IV}} := \mathbb{E}[w_{\text{IA}}^{\text{IV}}], W_{\text{CT}}^{\text{IV}} := \mathbb{E}[w_{\text{CT}}^{\text{IV}}]$. Consider an instance of auction where the contextual value of the exchange advertiser is c_1 and the contextual value of the direct advertiser is c_2. Both c_1, c_2 are drawn independently from the distribution G. First, we consider $\mathbb{E}\left[w_{\text{FI}}^{\text{IV}} | c_1, c_2\right]$, there are two cases: the case $\max\{c_1, c_2\} > \min\{c_1, c_2\} + \frac{\kappa}{1-\kappa}$, for which we find: $\mathbb{E}\left[w_{\text{FI}}^{\text{IV}} | c_1, c_2\right] = (\kappa + (1-\kappa)\min\{c_1, c_2\}) \cdot p + (1-\kappa)\min\{c_1, c_2\} \cdot (1-p) = \kappa p + (1-\kappa)\min\{c_1, c_2\}$ and the case: $\min\{c_1, c_2\} + \frac{\kappa}{1-\kappa} > \max\{c_1, c_2\}$, where we have

$$\mathbb{E}\left[w_{\text{FI}}^{\text{IV}} | c_1, c_2\right] = (\kappa + (1-\kappa)\min\{c_1, c_2\}) \cdot p^2 + (1-\kappa)\min\{c_1, c_2\} \cdot (1-p) + (1-\kappa)\max\{c_1, c_2\} \cdot (1-p)p$$

$$=\kappa p^2 + (1-\kappa)\min\{c_1,c_2\}\cdot(1-p+p^2) + (1-\kappa)\max\{c_1,c_2\}\cdot(1-p)p.$$

Next, we consider $\mathbb{E}\left[w_{\mathrm{IA}}^{\mathrm{IV}}|c_1,c_2\right]$, there are three cases: the case $c_1 > c_2 + \frac{p\kappa}{1-\kappa}$, for which we find $\mathbb{E}\left[w_{\mathrm{IA}}^{\mathrm{IV}}|c_1,c_2\right] = \kappa p + (1-\kappa)c_2$, the case: $c_2 + \frac{p\kappa}{1-\kappa} > c_1 > c_2 - \frac{(1-p)\kappa}{1-\kappa}$, where we have $\mathbb{E}\left[w_{\mathrm{IA}}^{\mathrm{IV}}|c_1,c_2\right] = (\kappa p + (1-\kappa)c_2)\cdot p + (1-\kappa)c_1\cdot(1-p)$, and the case: $c_2 - \frac{(1-p)\kappa}{1-\kappa} > c_1$, where we have $\mathbb{E}\left[w_{\mathrm{IA}}^{\mathrm{IV}}|c_1,c_2\right] = (\kappa + (1-\kappa)c_1)\cdot p + (1-\kappa)c_1\cdot(1-p) = \kappa p + (1-\kappa)c_1$.

Lastly, in all cases we have that $\mathbb{E}\left[w_{\mathrm{CT}}^{\mathrm{IV}}|c_1,c_2\right] = \kappa p + (1-\kappa)\min\{c_1,c_2\}$.

Now we can check that for all possible pairs of c_1, c_2 we have $\mathbb{E}\left[w_{\mathrm{FI}}^{\mathrm{IV}}|c_1,c_2\right] \leq \mathbb{E}\left[w_{\mathrm{IA}}^{\mathrm{IV}}|c_1,c_2\right] \leq \mathbb{E}\left[w_{\mathrm{IA}}^{\mathrm{CT}}|c_1,c_2\right]$. Taking an expectation over all possible c_1, c_2 we have that $W_{\mathrm{FI}}^{\mathrm{IV}} \leq W_{\mathrm{IA}}^{\mathrm{IV}} \leq W_{\mathrm{CT}}^{\mathrm{IV}}$ as claimed. ■

Proof of Proposition 9

From Proposition 8 we already know that for $n_1 = n_2 = 1$ we have $W_{\mathrm{FI}}^{\mathrm{IV}} \leq W_{\mathrm{CT}}^{\mathrm{IV}}$ for all $p \in [0,1]$. Now, fix $p \in [0,1]$ and consider the $n_1, n_2 > 0$ case. Note that using integration by-parts we can rewrite $W_{\mathrm{FI}}^{\mathrm{IV}}$ as:

$$\begin{aligned}
W_{\mathrm{FI}}^{\mathrm{IV}} =&(n_1+n_2)(1-\kappa)(1-p)^{n_1+n_2-1}p - (n_1+n_2)\kappa(1-p)^{n_1+n_2-1}p \\
&+ (1-\kappa)\int_0^1 \left(c - \frac{1-(1-p)G(c)}{(1-p)G'(c)}\right) d\left((1-p)G(c)\right)^{n_1+n_2} \\
&+ \int_0^1 \left(\kappa + (1-\kappa)c - \frac{(1-\kappa)(1-G(c))}{G'(c)}\right) d\left(pG(c) + (1-p)\right)^{n_1+n_2} \\
=& -(n_1+n_2)(2\kappa-1)(1-p)^{n_1+n_2-1}p + (1-\kappa)(1-p)^{n_1+n_2}\mathbb{E}_{c\sim G(c)^{n_1+n_2}}\left[c - \frac{1-(1-p)G(c)}{(1-p)G'(c)}\right] \\
&+ \mathbb{E}_{c\sim(pG(c)+(1-p))^{n_1+n_2}}\left[\kappa + (1-\kappa)c - \frac{(1-\kappa)(1-G(c))}{G'(c)}\right].
\end{aligned}$$

Where $\mathbb{E}_{c\sim F(c)}[.]$ denotes the expected value with c distributed by $F(c)$. Similarly, we can rewrite $W_{\mathrm{CT}}^{\mathrm{IV}}$ as

$$\begin{aligned}
W_{\mathrm{CT}}^{\mathrm{IV}} &= \int_0^1 \left(\kappa p + (1-\kappa)c - \frac{(1-\kappa)(1-G(c))}{G'(c)}\right) dG(c)^{n_1+n_2} \\
&= \mathbb{E}_{c\sim G(c)^{n_1+n_2}}\left[\kappa p + (1-\kappa)c - \frac{(1-\kappa)(1-G(c))}{G'(c)}\right].
\end{aligned}$$

When n_1 and n_2 are large, the densities of distributions $G(c)^{n_1+n_2}$ and $(pG(c)+(1-p))^{n_1+n_2}$ become concentrated around $c = 1$. Therefore, $W_{\mathrm{CT}}^{\mathrm{IV}}$ tends towards $\kappa p + (1-\kappa)$. On the other hand, the first and second terms in $W_{\mathrm{FI}}^{\mathrm{IV}}$ tend to zero due to $(1-p)^{n_1+n_2}$ but the last term tends to $\kappa + (1-\kappa) = 1$. Hence, we have $W_{\mathrm{FI}}^{\mathrm{IV}} > W_{\mathrm{CT}}^{\mathrm{IV}}$ for all sufficiently large n_1 and n_2. ■

Proof of Proposition 10

Using the formulas in the Online Appendix B.2 for $G(x) = x$ and $\kappa \geq 1/2$, we derive that

$$\frac{\partial}{\partial p}\left(D_{\mathrm{IA}}^{\mathrm{IV}} - D_{\mathrm{CT}}^{\mathrm{IV}}\right)\Big|_{p=0} = \frac{n_1}{n_1+n_2+1}\left(\frac{2\kappa}{n_1+n_2-1} - \frac{1}{n_1+n_2}\right)$$

and

$$\frac{\partial}{\partial p}\left(E_{\mathrm{FI}}^{\mathrm{IV}} - E_{\mathrm{IA}}^{\mathrm{IV}}\right)\Big|_{p=0} = \frac{n_2}{n_1+n_2+1}\left(\frac{2\kappa}{n_1+n_2-1} - \frac{1}{n_1+n_2}\right).$$

Since $\frac{2\kappa}{n_1+n_2-1} - \frac{1}{n_1+n_2} \geq \frac{1}{n_1+n_2-1} - \frac{1}{n_1+n_2} > 0$ when $\kappa \geq 1/2$, both derivatives are strictly positive at $p = 0$. Therefore, there is a neighborhood of $p = 0$, where the result holds. ■

Proof of Lemma 3

The exchange advertisers will always bid their true valuation as it is a weakly dominant strategy to do so. Therefore, for the remainder, we will focus on the nontrivial part, which is the direct advertisers' bidding strategy.

The expected utility for a direct advertiser with contextual value c from bidding $\tilde{\beta}$ when all other $n_2 - 1$ direct advertisers follow the strategy β is given by:

$$\begin{aligned}
u(\tilde{\beta};\beta,c) :=\, & p\Bigg[(1-\kappa)\int_0^{\max\left\{\frac{\tilde{\beta}-\kappa}{1-\kappa},0\right\}} (c-c')G(\sup\beta^{-1}[0,\kappa+(1-\kappa)c'])^{n_2-1}\left(n_1 G(c')^{n_1-1}G'(c')\right)dc' \\
&+ \int_0^{\sup\beta^{-1}[0,\tilde{\beta})} (\kappa+(1-\kappa)c-\beta(c'))G\left(\max\left\{\frac{\beta(c')-\kappa}{1-\kappa},0\right\}\right)^{n_1}\left((n_2-1)G(c')^{n_2-2}G'(c')\right)dc'\Bigg] \\
&+ (1-p)\Bigg[(1-\kappa)\int_0^{\min\left\{\frac{\tilde{\beta}}{1-\kappa},1\right\}} (c-c')G(\sup\beta^{-1}[0,(1-\kappa)c'])^{n_2-1}\left(n_1 G(c')^{n_1-1}G'(c')\right)dc' \\
&+ \int_0^{\sup\beta^{-1}[0,\tilde{\beta})} ((1-\kappa)c-\beta(c'))G\left(\min\left\{\frac{\beta(c')}{1-\kappa},1\right\}\right)^{n_1}\left((n_2-1)G(c')^{n_2-2}G'(c')\right)dc'\Bigg]. \qquad (4)
\end{aligned}$$

Let us restrict our attention to the bidding functions β that belong to the following class of functions: $\mathcal{F} := \left\{\beta \in L^1[0,1] \mid \beta \text{ is represented by a non-decreasing function } [0,1] \to [0,1]\right\}$.

Here, $L^1[0,1]$ denotes the usual Banach space of the equivalence classes of Lebesgue-integrable functions on $[0,1]$ equipped with the usual norm $\|f\|_{L^1} := \int_0^1 |f(x)|dx$. It is not hard to verify that $\mathcal{F}$ is a convex and compact subset of $L^1[0,1]$.

We note that the sign of each of the integrals in (4) is determined by the sign of $(c - c')$, $\kappa + (1-\kappa)c - \beta(c')$, and $(1-\kappa)c - \beta(c')$, respectively, all of which are increasing functions in c and decreasing in c'. Essentially, given β and c, finding the maximum $\tilde{\beta} = \tilde{\beta}_0$ of $u(\tilde{\beta};\beta,c)$ is to 'integrate until the integrands are negative'. The reality is slightly more subtle, as the upper limit of each integral are different non-linear functions of $\tilde{\beta}$.

Lemma 5. *Given $\beta \in \mathcal{F}$ and $c \in [0,1]$ then $u(\tilde{\beta};\beta,c)$ as a function of $\tilde{\beta} \in [0,1]$ achieves its global maximum inside $\{(1-\kappa)c\} \cup [\kappa, \kappa + (1-\kappa)c]$.*

Proof. First, let us observe where a maximum of $u(\tilde{\beta};\beta,c)$ cannot be located. If $\tilde{\beta}(c) \in (\kappa + (1-\kappa)c, 1]$ then the third term of (4) is constant in $\tilde{\beta}$. The first term is strictly decreasing for $\tilde{\beta} > \kappa + (1-\kappa)c$. The second and fourth terms are non-constant if $\beta(c) > \kappa + (1-\kappa)c$ for some c, but then these two terms decrease with $\tilde{\beta}$ because $\beta(c') > \kappa + (1-\kappa)c > (1-\kappa)c$ for $c' = \max \beta^{-1}[0,\tilde{\beta}) \geq \max \beta^{-1}[0, \kappa + (1-\kappa)c)$. Similarly, if $\tilde{\beta} \in [0, (1-\kappa)c)$ then only the third and fourth terms of (4) are non-constant in $\tilde{\beta}$. The third term is strictly increasing for $\tilde{\beta} < (1-\kappa)c$ and for any $c' = \max \beta^{-1}[0,\tilde{\beta}) \leq \max \tilde{\beta}^{-1}[0,(1-\kappa)c)$, which means $\beta(c') < (1-\kappa)c$, hence the fourth term is increasing.

If $\tilde{\beta} \in [(1-\kappa)c, \kappa]$, then every term of (4) is constant except for the fourth term which could be non-constant if $\beta(c) > (1-\kappa)c$ for some c, and in that case, the fourth term is decreasing. In other words, the maximum value of $u(\tilde{\beta};\beta,c)$ over $[(1-\kappa)c, \kappa]$ is reached at $\tilde{\beta} = (1-\kappa)c$. Since the fourth term of (4) is necessarily strictly decreasing, it is possible that $u(\tilde{\beta};\beta,c)$ also attains its maximum value at other points in $((1-\kappa)c, \kappa]$, this fact will serve no practical implication for us.

Next, we focus on the case where $\tilde{\beta} \in [\kappa, \kappa + (1-\kappa)c]$, and we shall show that $u(\tilde{\beta};\beta,c)$ also reaches its maximum over this interval. We note that $u(\tilde{\beta};\beta,c)$ is left-continuous because $\sup \beta^{-1}[0,\tilde{\beta})$ is left-continuous, and the point where it is not continuous is exactly where $\{c \mid \beta(c) = \tilde{\beta}_0\}$ has non-empty interior. In particular, let $\underline{b} := \inf\{c \mid \beta(c) = \tilde{\beta}_0\}$ and $\bar{b} := \sup\{c \mid \beta(c) = \tilde{\beta}_0\}$ then it follows that $(\underline{b}, \bar{b}) \subset S(\tilde{\beta}_0)$. In that case, we have $\sup \beta^{-1}[0,\tilde{\beta}) \leq \underline{b}$ for $\tilde{\beta} \leq \tilde{\beta}_0$ and $\sup \beta^{-1}[0,\tilde{\beta}) \geq \bar{b}$ for $\tilde{\beta} > \tilde{\beta}_0$. Given that $\tilde{\beta} \in [\kappa, \kappa + (1-\kappa)c]$, the third term of (4) is constant in a neighborhood of $\tilde{\beta}_0$. Let $\delta > 0$ be arbitrarily small, then the first term of (4) will take approximately the same value at $\tilde{\beta}_0$ and at $\tilde{\beta}_0 + \delta$. If $u(\tilde{\beta}_0;\beta,c) < u(\tilde{\beta}_0 + \delta;\beta,c)$ it must be the case that the sum of the second and fourth integrals is positive over $(\underline{b}, \bar{b})$. In particular, $\left(\kappa + (1-\kappa)c - \tilde{\beta}_0\right) G\left(\max\left\{\frac{\tilde{\beta}_0-\kappa}{1-\kappa}, 0\right\}\right)^{n_1} + \left((1-\kappa)c - \tilde{\beta}_0\right) G\left(\min\left\{\frac{\tilde{\beta}_0}{1-\kappa}, 1\right\}\right)^{n_1} > 0$.

But we also know that for all $c' \in \sup \beta^{-1}[0, \tilde{\beta}_0 + \delta)$ we have $\beta(c') < \tilde{\beta}_0 + \delta$, then from the inequality above we have that the sum of the integrands of the second and fourth integrals in (4) is positive immediately to the right of $\tilde{\beta}_0$ as $\delta > 0$ is arbitrary small. Since $\tilde{\beta}_0 \leq \kappa + (1-\kappa)c$, the integrand of the first integral in (4) is also positive. It follows that $u(\tilde{\beta}; \beta, c)$ continue to increase over some right neighbourhood of $\tilde{\beta}_0 + \delta$, hence $\sup_{\tilde{\beta}} u(\tilde{\beta}; \beta, c) > \lim_{\tilde{\beta} \to \tilde{\beta}_0^+} u(\tilde{\beta}; \beta, c)$. □

Lemma 5 allows us to define the best-response set-valued function as follows: $BR(\beta, c) := \arg\max_{\tilde{\beta} \in [0,1]} u(\tilde{\beta}; \beta, c)$. Let us also restrict our attention to β such that $\beta(c) \in \{(1-\kappa)c\} \cup [\kappa, \kappa + (1-\kappa)c]$.

Lemma 6. *The best-response function is closed-valued and* non-decreasing *in the sense that if* $c_1 < c_2$*, then* $\max BR(\beta, c_1) \leq \min BR(\beta, c_2)$.

Proof. The fact that $BR(\beta, c)$ is closed follows since according to Lemma 5, $u(\tilde{\beta}; \beta, c)$ is left-continuous and if $u(\tilde{\beta}; \beta, c)$ is discontinuous at $\tilde{\beta}_0$ then $\limsup_{\tilde{\beta} \to \tilde{\beta}_0^+} u(\tilde{\beta}; \beta, c)$ is always less than the global maximum value of u. In other words, if $\tilde{\beta}_i \in BR(\beta, c), i = 1, 2, \cdots$ and $\tilde{\beta}_i \to \tilde{\beta}_0 \in [0, 1]$ then $u(\tilde{\beta}_0; \beta, c) = u(\tilde{\beta}_i; \beta, c)$ for all i, which means $\tilde{\beta}_0 \in BR(\beta, c)$. Therefore, it makes sense to talk about the maximum and minimum of $BR(\beta, c)$.

Given any $\delta > 0$, we note that it is possible to write $u(\tilde{\beta}; \beta, c+\delta) = u(\tilde{\beta}; \beta, c) + \Delta(\tilde{\beta}; \beta, \delta)$, where $\Delta(\tilde{\beta}; \beta, c)$ is exactly given by (4) but with $(c-c')$, $(\kappa + (1-\kappa)c - \beta(c'))$, and $((1-\kappa)c - \beta(c'))$ factors replaced by δ, $(1-\kappa)\delta$, and $(1-\kappa)\delta$, respectively. Thus, $\Delta(\tilde{\beta}; \beta, \delta)$ is a non-decreasing function in $\tilde{\beta}$ and strictly increases over $[0, 1-\kappa] \cup [\kappa, 1]$. Then the fact that $BR(\beta, c)$ is non-decreasing follows from the following elementary argument. Let $\tilde{\beta}_0 = \max BR(\beta, c)$ then $u(\tilde{\beta}_0; \beta, c) \geq u(\tilde{\beta}; \beta, c)$ for all $\tilde{\beta} \in [0, \tilde{\beta}_0)$. Therefore, $u(\tilde{\beta}_0; \beta, c+\delta) > u(\tilde{\beta}; \beta, c+\delta)$ for all $\tilde{\beta} \in [0, \tilde{\beta}_0)$ by the strict monotonicity of $\Delta(\tilde{\beta}; \beta, \delta)$, which means any other global maxima of $u(\tilde{\beta}; \beta, c+\delta)$ must be in $[\tilde{\beta}_0, \kappa + (1-\kappa)c]$, proving the lemma. □

Using Lemma 6 it is now possible to define the best-response bidding function to the bidding β of all other $n_2 - 1$ direct advertisers: $BR : \mathcal{F} \to \mathcal{F}$, $\tilde{\beta} := BR(\beta) : c \mapsto \min BR(\beta, c)$, where we have slightly abused the notation, using both $\tilde{\beta}$ as a particular bidding value and the bidding function, and BR as both the best response bidding set-valued function and the best response bidding function-valued map. However, we hope that any ambiguity can be resolved by context.

Lemma 7. *The best-response function BR is continuous with respect to the L^1 norm.*

We will omit the technical proof, but the intuition is clear. Any two $\beta_1, \beta_2 \in \mathcal{F}$ non-decreasing functions which are 'close' together under L^1 norm must take similar values $\beta_1(c) \approx \beta_2(c)$ at any c they are both continuous. Moreover, the location of any discontinuous points of β_1 and β_2 must be similar. The same is true for their inverses $\sup \beta_1^{-1}[0, \tilde{\beta}) \approx \beta_2^{-1}[0, \tilde{\beta})$. Hence we can expect $u(\tilde{\beta}; \beta_1, c) \approx u(\tilde{\beta}; \beta_2, c)$ for all $\tilde{\beta}$ and c and therefore the maximum point of $u(.; \beta_1, c)$ should be close to the maximum point of $u(.; \beta_2, c)$.

From Lemma 7, the response function BR is continuous with respect to L^1 norm and maps a convex compact subset $\mathcal{F} \subset L^1[0,1]$ into itself. $L^1[0,1]$ is a normed-vector space, hence it is automatically a Hausdorff locally convex topological vector space. From the Kakutani-Fan-Glicksberg Theorem, we know that BR has a fixed point. ■

B.2 Key Formulas

Unless stated otherwise, all formulas in this section are valid for any given $\kappa \in [0,1]$, $p \in [0,1]$, $n_1, n_2 \geq 0$ and an arbitrary contextual-value distribution G on $[0,1]$.

Common-value case

To deal with any discontinuities of the bidding function β we let $\beta^{-1}[a,b)$ denote the inverse image of β i.e. a set I such that $x \in I \implies \beta(x) \in [a,b)$, and $\sup \beta^{-1}[a,b)$ denotes the supremum of this set.

Advertisers' conversion rate:

The advertisers' conversion rates under each information setting are given by:

$$\begin{aligned} V_{\text{FI}}^{\text{CV}} =& p(n_1+n_2) \int_0^1 (\kappa + (1-\kappa)c) G(c)^{n_1+n_2-1} G'(c) dc \\ &+ (1-p)(n_1+n_2) \int_0^1 (1-\kappa) c G(c)^{n_1+n_2-1} G'(c) dc, \end{aligned}$$

$$
\begin{aligned}
V_{\text{IA}}^{\text{CV}} =& pn_2 \int_0^1 (\kappa + (1-\kappa)c) G\left(\max\left\{\frac{\beta(c)-\kappa}{1-\kappa}, 0\right\}\right)^{n_1} G(c)^{n_2-1} G'(c) dc \\
&+ pn_1 \int_0^1 (\kappa + (1-\kappa)c) G(c)^{n_1-1} G\left(\sup \beta^{-1}[0, \kappa + (1-\kappa)c)\right)^{n_2} G'(c) dc \\
&+ (1-p) n_2 \int_0^1 (1-\kappa) c G\left(\min\left\{\frac{\beta(c)}{1-\kappa}, 1\right\}\right)^{n_1} G(c)^{n_2-1} G'(c) dc \\
&+ (1-p) n_1 \int_0^1 (1-\kappa) c G(c)^{n_1-1} G\left(\sup \beta^{-1}[0, (1-\kappa)c)\right)^{n_2} G'(c) dc,
\end{aligned}
$$

$$
V_{\text{CT}}^{\text{CV}} = (n_1 + n_2) \int_0^1 (\kappa p + (1-\kappa)c) G(c)^{n_1+n_2-1} G'(c) dc.
$$

We note that $V_{\text{FI}}^{\text{CV}} = V_{\text{CT}}^{\text{CV}}$.

Publisher's expected revenue:

The publisher's expected revenues for each information setting are given by:

$$
\begin{aligned}
W_{\text{FI}}^{\text{CV}} =& p(n_1+n_2)(n_1+n_2-1) \int_0^1 (\kappa + (1-\kappa)c)(1-G(c)) G(c)^{n_1+n_2-2} G'(c) dc \\
&+ (1-p)(n_1+n_2)(n_1+n_2-1) \int_0^1 (1-\kappa) c (1-G(c)) G(c)^{n_1+n_2-2} G'(c) dc,
\end{aligned}
$$

$$
\begin{aligned}
W_{\text{IA}}^{\text{CV}} =& pn_1 n_2 \int_0^1 (\kappa + (1-\kappa)c) \left(1 - G\left(\sup \beta^{-1}[0, \kappa + (1-\kappa)c)\right)\right) G(c)^{n_1-1} G\left(\sup \beta^{-1}[0, \kappa + (1-\kappa)c)\right)^{n_2-1} G'(c) dc \\
&+ pn_1 n_2 \int_0^1 \beta(c) \left(1 - G\left(\max\left\{\frac{\beta(c)-\kappa}{1-\kappa}, 0\right\}\right)\right) G\left(\max\left\{\frac{\beta(c)-\kappa}{1-\kappa}, 0\right\}\right)^{n_1-1} G(c)^{n_2-1} G'(c) dc \\
&+ pn_1(n_1-1) \int_0^1 (\kappa + (1-\kappa)c) (1-G(c)) G(c)^{n_1-2} G\left(\sup \beta^{-1}[0, \kappa + (1-\kappa)c)\right)^{n_2} G'(c) dc \\
&+ pn_2(n_2-1) \int_0^1 \beta(c) (1-G(c)) G\left(\max\left\{\frac{\beta(c)-\kappa}{1-\kappa}, 0\right\}\right)^{n_1} G(c)^{n_2-2} G'(c) dc \\
&+ (1-p) n_1 n_2 \int_0^1 (1-\kappa) c \left(1 - G\left(\sup \beta^{-1}[0, (1-\kappa)c)\right)\right) G(c)^{n_1-1} G\left(\sup \beta^{-1}[0, (1-\kappa)c)\right)^{n_2-1} G'(c) dc \\
&+ (1-p) n_1 n_2 \int_0^1 \beta(c) \left(1 - G\left(\min\left\{\frac{\beta(c)}{1-\kappa}, 1\right\}\right)\right) G\left(\min\left\{\frac{\beta(c)}{1-\kappa}, 1\right\}\right)^{n_1-1} G(c)^{n_2-1} G'(c) dc \\
&+ (1-p) n_1(n_1-1) \int_0^1 (1-\kappa) c (1-G(c)) G(c)^{n_1-2} G\left(\sup \beta^{-1}[0, (1-\kappa)c)\right)^{n_2} G'(c) dc \\
&+ (1-p) n_2(n_2-1) \int_0^1 \beta(c) (1-G(c)) G\left(\min\left\{\frac{\beta(c)}{1-\kappa}, 1\right\}\right)^{n_1} G(c)^{n_2-2} G'(c) dc,
\end{aligned} \tag{5}
$$

$$
W_{\text{CT}}^{\text{CV}} = (n_1+n_2)(n_1+n_2-1) \int_0^1 (\kappa p + (1-\kappa)c)(1-G(c)) G(c)^{n_1+n_2-2} G'(c) dc.
$$

We note that $W_{\text{FI}}^{\text{CV}} = W_{\text{CT}}^{\text{CV}}$.

Direct advertisers' payoff:

$$D_{\text{FI}}^{\text{CV}} = \frac{V_{\text{FI}}^{\text{CV}} - W_{\text{FI}}^{\text{CV}}}{n_1 + n_2} = \int_0^1 (p\kappa + (1-\kappa)c) G(c)^{n_1+n_2-1} G'(c) dc$$
$$- (n_1 + n_2 - 1) \int_0^1 (p\kappa + (1-\kappa)c)(1 - G(c)) G(c)^{n_1+n_2-2} G'(c) dc, \quad (6)$$

$$\begin{aligned}
D_{\text{IA}}^{\text{CV}} =& p \int_0^1 (\kappa + (1-\kappa)c) G\left(\max\left\{\frac{\beta(c)-\kappa}{1-\kappa}, 0\right\}\right)^{n_1} G(c)^{n_2-1} G'(c) dc \\
&+ (1-p) \int_0^1 (1-\kappa) c G\left(\min\left\{\frac{\beta(c)}{1-\kappa}, 1\right\}\right)^{n_1} G(c)^{n_2-1} G'(c) dc \\
&- p n_1 \int_0^1 (\kappa + (1-\kappa)c) \left(1 - G\left(\sup \beta^{-1}[0, \kappa + (1-\kappa)c)\right)\right) G(c)^{n_1-1} G\left(\sup \beta^{-1}[0, \kappa + (1-\kappa)c)\right)^{n_2-1} G'(c) dc \\
&- p(n_2 - 1) \int_0^1 \beta(c) (1 - G(c)) G\left(\max\left\{\frac{\beta(c)-\kappa}{1-\kappa}, 0\right\}\right)^{n_1} G(c)^{n_2-2} G'(c) dc \\
&- (1-p) n_1 \int_0^1 (1-\kappa) c \left(1 - G\left(\sup \beta^{-1}[0, (1-\kappa)c)\right)\right) G(c)^{n_1-1} G\left(\sup \beta^{-1}[0, (1-\kappa)c)\right)^{n_2-1} G'(c) dc \\
&- (1-p)(n_2 - 1) \int_0^1 \beta(c) (1 - G(c)) G\left(\min\left\{\frac{\beta(c)}{1-\kappa}, 1\right\}\right)^{n_1} G(c)^{n_2-2} G'(c) dc, \quad (7)
\end{aligned}$$

$$D_{\text{CT}}^{\text{CV}} = \frac{V_{\text{CT}}^{\text{CV}} - W_{\text{CT}}^{\text{CV}}}{n_1 + n_2} = D_{\text{FI}}^{\text{CV}}. \quad (8)$$

Exchange advertisers' payoff:

$$E_{\text{FI}}^{\text{CV}} = \frac{V_{\text{FI}}^{\text{CV}} - W_{\text{FI}}^{\text{CV}}}{n_1 + n_2} = D_{\text{FI}}^{\text{CV}} = \frac{V_{\text{CT}}^{\text{CV}} - W_{\text{CT}}^{\text{CV}}}{n_1 + n_2} = E_{\text{CT}}^{\text{CV}}, \quad (9)$$

$$\begin{aligned}
E_{\text{IA}}^{\text{CV}} =& p \int_0^1 (\kappa + (1-\kappa)c) G(c)^{n_1-1} G\left(\sup \beta^{-1}[0, \kappa + (1-\kappa)c)\right)^{n_2} G'(c) dc \\
&+ (1-p) \int_0^1 (1-\kappa) c G(c)^{n_1-1} G\left(\sup \beta^{-1}[0, (1-\kappa)c)\right)^{n_2} G'(c) dc \\
&- p n_2 \int_0^1 \beta(c) \left(1 - G\left(\max\left\{\frac{\beta(c)-\kappa}{1-\kappa}, 0\right\}\right)\right) G\left(\max\left\{\frac{\beta(c)-\kappa}{1-\kappa}, 0\right\}\right)^{n_1-1} G(c)^{n_2-1} G'(c) dc \\
&- pn(n_1 - 1) \int_0^1 (\kappa + (1-\kappa)c) (1 - G(c)) G(c)^{n_1-2} G\left(\sup \beta^{-1}[0, \kappa + (1-\kappa)c)\right)^{n_2} G'(c) dc \\
&- (1-p) n_2 \int_0^1 \beta(c) \left(1 - G\left(\min\left\{\frac{\beta(c)}{1-\kappa}, 1\right\}\right)\right) G\left(\min\left\{\frac{\beta(c)}{1-\kappa}, 1\right\}\right)^{n_1-1} G(c)^{n_2-1} G'(c) dc \\
&- (1-p)(n_1 - 1) \int_0^1 (1-\kappa) c (1 - G(c)) G(c)^{n_1-2} G\left(\sup \beta^{-1}[0, (1-\kappa)c)\right)^{n_2} G'(c) dc. \quad (10)
\end{aligned}$$

Independent-values case

In the independent-values case, all advertisers bid truthfully. An advertiser with a contextual value c but without behavioral information would bid the expected value $\kappa p + (1-\kappa)c$. An advertiser with full information would bid $v = \kappa b + (1-\kappa)c$ where v is distributed by $\tilde{G}(v) := \mathbb{P}[v' \leq v] = pG\left(\max\left\{\frac{v-\kappa}{1-\kappa}, 0\right\}\right) + (1-p)G\left(\min\left\{\frac{v}{1-\kappa}, 1\right\}\right)$.

Advertisers' conversion rate:

The advertisers' conversion rates under each information setting are given by:

$$V_{\mathrm{FI}}^{\mathrm{IV}} = (n_1+n_2)\int_0^1 v\tilde{G}(v)^{n_1+n_2-1}\tilde{G}'(v)dv,$$

$$\begin{aligned}V_{\mathrm{IA}}^{\mathrm{IV}} = n_1 &\int_0^1 v\tilde{G}(v)^{n_1-1}G\left(\min\left\{\max\left\{\frac{v-\kappa p}{1-\kappa},0\right\},1\right\}\right)^{n_2}\tilde{G}'(v)dv \\ &+ (1-p)n_2\int_0^1 (1-\kappa)c\tilde{G}(\kappa p+(1-\kappa)c)^{n_1}G(c)^{n_2-1}G'(c)dc \\ &+ pn_2\int_0^1 (\kappa+(1-\kappa)c)\tilde{G}(\kappa p+(1-\kappa)c)^{n_1}G(c)^{n_2-1}G'(c)dc,\end{aligned}$$

$$V_{\mathrm{CT}}^{\mathrm{IV}} = (n_1+n_2)\int_0^1 (\kappa p+(1-\kappa)c)G(c)^{n_1+n_2-1}G'(c)dc.$$

Publisher's expected revenue:

The publisher's expected revenues for each information setting are given by:

$$W_{\mathrm{FI}}^{\mathrm{IV}} = (n_1+n_2)(n_1+n_2-1)\int_0^1 v\left(1-\tilde{G}(v)\right)\tilde{G}(v)^{n_1+n_2-2}\tilde{G}'(v)dv,$$

$$\begin{aligned}W_{\mathrm{IA}}^{\mathrm{IV}} =& n_1(n_1-1)\int_0^1 v\left(1-\tilde{G}(v)\right)\tilde{G}(v)^{n_1-2}G\left(\min\left\{\max\left\{\frac{v-\kappa p}{1-\kappa},0\right\},1\right\}\right)^{n_2}\tilde{G}'(v)dv \\ &+ n_1n_2\int_0^1 v\left(1-G\left(\min\left\{\max\left\{\frac{v-\kappa p}{1-\kappa},0\right\},1\right\}\right)\right)\tilde{G}(v)^{n_1-1}G\left(\min\left\{\max\left\{\frac{v-\kappa p}{1-\kappa},0\right\},1\right\}\right)^{n_2-1}\tilde{G}'(v)dv \\ &+ n_1n_2\int_0^1 (\kappa p+(1-\kappa)c)\left(1-\tilde{G}(\kappa p+(1-\kappa)c)\right)\tilde{G}(\kappa p+(1-\kappa)c)^{n_1-1}G(c)^{n_2-1}G'(c)dc \\ &+ n_2(n_2-1)\int_0^1 (\kappa p+(1-\kappa)c)\left(1-G(c)\right)\tilde{G}\left(\kappa p+(1-\kappa)c\right)^{n_1}G(c)^{n_2-2}G'(c)dc,\end{aligned}$$

$$W_{\mathrm{CT}}^{\mathrm{IV}} = (n_1+n_2)(n_1+n_2-1)\int_0^1 (\kappa p+(1-\kappa)c)(1-G(c))G(c)^{n_1+n_2-2}G'(c)dc.$$

Direct advertisers' payoff:

$$D_{\mathrm{FI}}^{\mathrm{IV}} = \frac{V_{\mathrm{FI}}^{\mathrm{IV}} - W_{\mathrm{FI}}^{\mathrm{IV}}}{n_1 + n_2} = \int_0^1 v\tilde{G}(v)^{n_1+n_2-1}\tilde{G}'(v)dv$$
$$- (n_1 + n_2 - 1)\int_0^1 v\left(1 - \tilde{G}(v)\right)\tilde{G}(v)^{n_1+n_2-2}\tilde{G}'(v)dv, \tag{11}$$

$$\begin{aligned} D_{\mathrm{IA}}^{\mathrm{IV}} =& (1-p)\int_0^1 (1-\kappa)c\tilde{G}(\kappa p + (1-\kappa)c)^{n_1} G(c)^{n_2-1}G'(c)dc \\ &+ p\int_0^1 (\kappa + (1-\kappa)c)\tilde{G}(\kappa p + (1-\kappa)c)^{n_1} G(c)^{n_2-1}G'(c)dc \\ &- n_1\int_0^1 v\left(1 - G\left(\min\left\{\max\left\{\frac{v-\kappa p}{1-\kappa}, 0\right\}, 1\right\}\right)\right)\tilde{G}(v)^{n_1-1}G\left(\min\left\{\max\left\{\frac{v-\kappa p}{1-\kappa}, 0\right\}, 1\right\}\right)^{n_2-1}\tilde{G}'(v)dv \\ &- (n_2 - 1)\int_0^1 (\kappa p + (1-\kappa)c)\left(1 - G(c)\right)\tilde{G}\left(\kappa p + (1-\kappa)c\right)^{n_1} G(c)^{n_2-2}G'(c)dc, \end{aligned} \tag{12}$$

$$D_{\mathrm{CT}}^{\mathrm{IV}} = \frac{V_{\mathrm{CT}}^{\mathrm{IV}} - W_{\mathrm{CT}}^{\mathrm{IV}}}{n_1 + n_2} = \int_0^1 (\kappa p + (1-\kappa)c)G(c)^{n_1+n_2-1}G'(c)dc$$
$$- (n_1 + n_2 - 1)\int_0^1 (\kappa p + (1-\kappa)c)(1 - G(c))G(c)^{n_1+n_2-2}G'(c)dc.$$

Exchange advertisers' payoff:

$$E_{\mathrm{FI}}^{\mathrm{IV}} = \frac{V_{\mathrm{FI}}^{\mathrm{IV}} - W_{\mathrm{FI}}^{\mathrm{IV}}}{n_1 + n_2} = D_{\mathrm{FI}}^{\mathrm{IV}}, \qquad E_{\mathrm{CT}}^{\mathrm{IV}} = \frac{V_{\mathrm{CT}}^{\mathrm{IV}} - W_{\mathrm{CT}}^{\mathrm{IV}}}{n_1 + n_2} = D_{\mathrm{CT}}^{\mathrm{IV}}, \tag{13}$$

$$\begin{aligned} E_{\mathrm{IA}}^{\mathrm{IV}} =& \int_0^1 v\tilde{G}(v)^{n_1-1}G\left(\min\left\{\max\left\{\frac{v-\kappa p}{1-\kappa}, 0\right\}, 1\right\}\right)^{n_2}\tilde{G}'(v)dv \\ &- (n_1 - 1)\int_0^1 v\left(1 - \tilde{G}(v)\right)\tilde{G}(v)^{n_1-2}G\left(\min\left\{\max\left\{\frac{v-\kappa p}{1-\kappa}, 0\right\}, 1\right\}\right)^{n_2}\tilde{G}'(v)dv \\ &- n_2\int_0^1 (\kappa p + (1-\kappa)c)\left(1 - \tilde{G}(\kappa p + (1-\kappa)c)\right)\tilde{G}(\kappa p + (1-\kappa)c)^{n_1-1}G(c)^{n_2-1}G'(c)dc. \end{aligned} \tag{14}$$

B.3 First-Price Auction

A commonly used auction format in advertising auctions nowadays is the first-price auction format (see e.g., Despotakis et al., 2021). In this section, we modify our model to a first-price auction instead of a second-price auction for selling the impression, to test the robustness of our main findings.

In a first-price auction, since advertisers pay their own bid if they win, direct advertisers who do not have complete information about their valuations, are even more likely to bid conservatively (underbid) compared to a second-price auction. As a result, information asymmetry can lead to both a lower conversion rate and reduced publisher revenue, compared to the symmetric information settings, for similar reasons this happens in second-price auctions. In other words, our result that disabling microtargeting can simultaneously increase both the conversion rate and publisher revenue remains valid in common-value first-price auctions. Propositions 11 and 12 below confirm this.

Proposition 11. *For any distribution G, any $n_1, n_2 \geq 1$, $\kappa \in [0,1]$, under the common-value case, and when the auction format is first-price, we have that $V_{\mathrm{IA}}^{\mathrm{CV}} \leq V_{\mathrm{FI}}^{\mathrm{CV}} = V_{\mathrm{CT}}^{\mathrm{CV}}$.*

Proof. This result follows from Lemma 4, where the mechanism M is a first-price auction. □

Proposition 12. *For any distribution G, $n_1 > 1$, $n_2 \geq 1$, $\kappa \geq 1/2$, and sufficiently low p, under the common-value case, and when the auction format is first-price, we have that $W_{\mathrm{IA}}^{\mathrm{CV}} \leq W_{\mathrm{FI}}^{\mathrm{CV}} = W_{\mathrm{CT}}^{\mathrm{CV}}$.*

Proof. Suppose that, when there is information asymmetry, at equilibrium the direct and exchange advertisers use the bidding functions $\beta_D : [0,1] \to [0,1]$ and $\beta_E : [0,1]^2 \to [0,1]$, respectively. The expected utility of a direct advertiser with contextual value c who bids $\tilde{\beta}$ is

$$\begin{aligned} u_D(\tilde{\beta}; \beta_D, \beta_E, c) =& p\left(\kappa + (1-\kappa)c - \tilde{\beta}\right) G\left(\sup \beta_E^{-1}([0,\tilde{\beta}), b=1)\right)^{n_1} G(\sup \beta_D^{-1}([0,\tilde{\beta})))^{n_2-1} \\ &+ (1-p)\left((1-\kappa)c - \tilde{\beta}\right) G\left(\sup \beta_E^{-1}([0,\tilde{\beta}), b=0)\right)^{n_1} G(\sup \beta_D^{-1}([0,\tilde{\beta})))^{n_2-1}. \end{aligned}$$

The expected utility of an exchange advertiser with contextual value c and behavioral value b who bids $\tilde{\beta}$ is $u_E(\tilde{\beta}; \beta_D, \beta_E, c, b) = \left(\kappa b + (1-\kappa)c - \tilde{\beta}\right) G(\sup \beta_E^{-1}([0,\tilde{\beta}), b))^{n_1-1} G(\sup \beta_D^{-1}([0,\tilde{\beta})))^{n_2}$.

The expected publisher revenue is

$$\begin{aligned} W_{\mathrm{IA}}^{\mathrm{CV}} =& pn_2 \int_0^1 \beta_D(c) G(\sup \beta_E^{-1}([0,\beta_D(c)),1))^{n_1} G(c)^{n_2-1} G'(c)dc \\ &+ pn_1 \int_0^1 \beta_E(c,1) G(c)^{n_1-1} G(\sup \beta_D^{-1}([0,\beta_E(c,1))))^{n_2} G'(c)dc \\ &+ (1-p)n_2 \int_0^1 \beta_D(c) G(\sup \beta_E^{-1}([0,\beta_D(c)),0))^{n_1} G(c)^{n_2-1} G'(c)dc \\ &+ (1-p)n_1 \int_0^1 \beta_E(c,0) G(c)^{n_1-1} G(\sup \beta_D^{-1}([0,\beta_E(c,0))))^{n_2} G'(c)dc. \end{aligned}$$

For the other two information settings, $W_{\text{FI}}^{\text{CV}}$ and $W_{\text{CT}}^{\text{CV}}$, the revenue is identical to the second-price auction case by the revenue equivalence principle.

The conversion rate is

$$
\begin{aligned}
V_{\text{IA}}^{\text{CV}} =& pn_2 \int_0^1 (\kappa + (1-\kappa)c) G\left(\sup \beta_E^{-1}([0, \beta_D(c)), 1)\right)^{n_1} G(c)^{n_2-1} G'(c) dc \\
&+ pn_1 \int_0^1 (\kappa + (1-\kappa)c) G(c)^{n_1-1} G\left(\sup \beta_D^{-1}[0, \beta_E(c,1))\right)^{n_2} G'(c) dc \\
&+ (1-p)n_2 \int_0^1 (1-\kappa)c G\left(\sup \beta_E^{-1}([0, \beta_D(c)), 0)\right)^{n_1} G(c)^{n_2-1} G'(c) dc \\
&+ (1-p)n_1 \int_0^1 (1-\kappa)c G(c)^{n_1-1} G\left(\sup \beta_D^{-1}[0, \beta_E(c,0))\right)^{n_2} G'(c) dc.
\end{aligned}
$$

If p is sufficiently low,[14] the direct advertisers will choose to not compete with the exchange advertisers if $b = 1$ and will always bid as if $b = 0$ regardless of the actual value of b (which they do not know anyway). In this case, when $b = 0$ the game reduces to a symmetric first-price auction among all $n_1 + n_2$ advertisers, and when $b = 1$ the game reduces to a symmetric first-price auction among n_1 exchange advertisers. Consequently, we have a symmetric equilibrium where the bidding functions have simple analytical closed forms as follows:

$$
\begin{aligned}
\beta_D(c) &= (1-\kappa)\left(c - \int_0^c \left(\frac{G(t)}{G(c)}\right)^{n_1+n_2-1} dt\right), \\
\beta_E(c,b) &= \kappa b + (1-\kappa)\left(c - \int_0^c \left(\frac{G(t)}{G(c)}\right)^{n_1+(1-b)n_2-1} dt\right).
\end{aligned}
\tag{15}
$$

Under the perfect-information and the contextual-targeting settings, the bidders are symmetric and independent, thus it follows from the revenue equivalence principle that $W_{\text{IA}}^{\text{CV}}$ and $W_{\text{FI}}^{\text{CV}}$ are both equal to their second-price auction counterparts, hence they are equal to each other. We also note from (15) that $\beta_D(c) \leq \beta_E(c,b) \leq \kappa b + (1-\kappa)\left(c - \int_0^c \left(\frac{G(t)}{G(c)}\right)^{n_1+n_2-1} dt\right)$, for all c, b, where notice that the RHS is the bidding function under the perfect-information setting. The first inequality holds since $\beta_E(c,1) - \beta_D(c) \geq \kappa - (1-\kappa) \geq 0$ because $\kappa \geq 1/2$, and the second inequality holds since $\left(\frac{G(t)}{G(c)}\right)^{n_1+(1-b)n_2-1} \geq \left(\frac{G(t)}{G(c)}\right)^{n_1+n_2-1}$ for all $t, c \in [0,1]$ such that $t \leq c$. In other words, the bids of all advertisers are at least as high in the perfect-information setting as under information asymmetry, hence the revenue $W_{\text{FI}}^{\text{CV}}$ is at least as high as $W_{\text{IA}}^{\text{CV}}$. □

[14] E.g. if $\kappa p + (1-\kappa) < \kappa \implies p < 2 - 1/\kappa$, which is a sufficient but not necessary bound.

C Common Contextual Values

In this section, we consider the case where the contextual random variables c_i are not independent, but they are the same for all advertisers, i.e. $c_1 = c_2 = \ldots = c_n =: c$, where c is drawn from distribution G. As in the main model, we consider the following sub-cases. Let $\mathbb{E}[c] := \int_0^1 c dG(c)$ in the following.

- Common behavioral values (CV): All the advertisers have the same valuation $\kappa b + (1-\kappa)c$.
  - Full Information (FI): All advertisers bid $\kappa b + (1-\kappa)c$. For $n_1, n_2 \geq 1$: $V_{\text{FI}}^{\text{CV}} = (\kappa + (1-\kappa)\mathbb{E}[c]) \cdot p + (1-\kappa)\mathbb{E}[c] \cdot (1-p) = \kappa p + (1-\kappa)\mathbb{E}[c]$.
  - Information Asymmetry (IA): Exchange advertisers bid $\kappa b + (1-\kappa)c$. If there is at least one exchange advertiser, direct advertisers bid $(1-\kappa)c$. If there is no exchange advertiser, direct advertisers bid $\kappa p + (1-\kappa)c$. For $n_1, n_2 \geq 1$: $V_{\text{IA}}^{\text{CV}} = (\kappa + (1-\kappa)\mathbb{E}[c]) \cdot p + (1-\kappa)\mathbb{E}[c] \cdot (1-p) = \kappa p + (1-\kappa)\mathbb{E}[c]$.
  - Contextual Targeting (CT): All advertisers bid $\kappa p + (1-\kappa)c$. For $n_1, n_2 \geq 1$: $V_{\text{CT}}^{\text{CV}} = \kappa p + (1-\kappa)\mathbb{E}[c]$.

  We can see that $V_{\text{FI}}^{\text{CV}} = V_{\text{IA}}^{\text{CV}} = V_{\text{CT}}^{\text{CV}}$.

- Independent behavioral values (IV): Advertisers have valuations $\kappa b_i + (1-\kappa)c$.
  - Full Information (FI): All advertisers bid their valuation $\kappa b_i + (1-\kappa)c$. For $n_1, n_2 \geq 1$:

$$\begin{aligned} V_{\text{FI}}^{\text{IV}} &= (\kappa + (1-\kappa)\mathbb{E}[c]) \cdot \left(1 - (1-p)^{n_1+n_2}\right) + (1-\kappa)\mathbb{E}[c] \cdot (1-p)^{n_1+n_2} \\ &= \kappa\left(1 - (1-p)^{n_1+n_2}\right) + (1-\kappa)\mathbb{E}[c]. \end{aligned}$$

  - Information Asymmetry (IA): Exchange advertisers bid their valuation $\kappa b_i + (1-\kappa)c$. Direct advertisers bid $\kappa p + (1-\kappa)c$. For $n_1, n_2 \geq 1$:

$$\begin{aligned} V_{\text{IA}}^{\text{IV}} &= (\kappa + (1-\kappa)\mathbb{E}[c]) \cdot \left(1 - (1-p)^{n_1}\right) + (\kappa p + (1-\kappa)\mathbb{E}[c]) \cdot (1-p)^{n_1} \\ &= \kappa\left(1 - (1-p)^{n_1+1}\right) + (1-\kappa)\mathbb{E}[c]. \end{aligned}$$

  - Contextual Targeting (CT): All advertisers bid $\kappa p + (1-\kappa)c$. For $n_1, n_2 \geq 1$: $V_{\text{CT}}^{\text{IV}} = \kappa p + (1-\kappa)\mathbb{E}[c]$.

  Since $1 - (1-p)^n$ is an increasing function in $n \geq 0$, it follows that $V_{\text{FI}}^{\text{CV}} \geq V_{\text{IA}}^{\text{CV}} \geq V_{\text{CT}}^{\text{CV}}$.