Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Incentive Misalignments in Programmatic Advertising: Evidence from a Randomized Field Experiment

Thomas W. Frick, Rodrigo Belo, Rahul Telang

informs
https://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Vol. 69, No. 3, March 2023, pp. 1665–1686
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Incentive Misalignments in Programmatic Advertising: Evidence from a Randomized Field Experiment

Thomas W. Frick,[a] Rodrigo Belo,[b,c,*] Rahul Telang[d]

[a] Marketing Analyst, EPOS Group A/S, 2750 Ballerup, Denmark; [b] Nova School of Business and Economics, Universidade Nova de Lisboa, 2775-405 Carcavelos, Portugal; [c] Rotterdam School of Management, Erasmus University, 3062 PA Rotterdam, Netherlands; [d] The Heinz College, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213
*Corresponding author

Contact: thomas.frick.academia@gmail.com, https://orcid.org/0000-0002-4456-9093 (TWF); rbelo@rsm.nl, https://orcid.org/0000-0001-6839-5086 (RB); rtelang@andrew.cmu.edu, https://orcid.org/0000-0003-1069-1514 (RT)

Received: January 24, 2019
Revised: July 17, 2020; April 19, 2021; December 15, 2021
Accepted: February 1, 2022
Published Online in Articles in Advance: May 26, 2022

https://doi.org/10.1287/mnsc.2022.4438

Abstract. In programmatic advertising, firms outsource the bidding for ad impressions to ad platforms. Although firms are interested in targeting consumers that respond positively to advertising, ad platforms are usually rewarded for targeting consumers with high overall purchase probability. We develop a theoretical model that shows if consumers with high baseline purchase probability respond more positively to advertising, then firms and ad platforms agree on which consumers to target. If, conversely, consumers with low baseline purchase probability are the ones for which ads work best, then ad platforms target consumers that firms do not want to target—the incentives are misaligned. We conduct a large-scale randomized field experiment, targeting 208,538 individual consumers, in a display retargeting campaign. Our unique data set allows us to both causally identify advertising effectiveness and estimate the degree of incentive misalignments between the firm and ad platform. In accordance with the contracted incentives, the ad platform targets consumers that are more likely to purchase. Importantly, we find no evidence that ads are more effective for consumers with higher baseline purchase probability, rendering the ad platform's bidding suboptimal for the firm. A welfare analysis suggests that the ad platform's bidding optimization leads to a loss in profit for the firm and an overall decline in welfare. To remedy the incentive misalignment, we propose a solution in which the firm restricts the ad platform to target only consumers that are profitable based on individual consumer-level estimates for baseline purchase probability and ad effectiveness.

History: Accepted by Anandhi Bharadwaj, information systems.

Funding: This work was funded by Vereniging Trustfonds Erasmus Universiteit Rotterdam [Grant 97090/17.16/741/oe], Fundacão para a Ciência e a Tecnologia [UID/ECO/00124/2019, UIDB/00124/2020, UIDP/00124/2020, and Social Sciences DataLab - PINFRA/22209/2016], POR Lisboa, and POR Norte [Social Sciences DataLab, PINFRA/22209/2016].
Supplemental Material: The online appendices are available at https://doi.org/10.1287/mnsc.2022.4438.

Keywords: economics of advertising • field experiment • incentive misalignment • programmatic advertising

1. Introduction

Firms serve millions of digital ad impressions to consumers every month. Digital ad spaces are usually sold via auctions in ad exchanges, in which advertisers bid for the opportunity to show an ad to a consumer. This high-frequency process requires firms to place an individual bid for every available ad slot they are interested in. Most firms do not have the capacity to participate in these auctions directly and thus outsource the bidding process to an ad platform. Ad platforms provide ad allocation tools that firms can use to automate the purchasing of ad impressions. This service is called programmatic advertising. In programmatic advertising, most opportunities to display an ad are auctioned off, and the ad platform decides, on behalf of the firm, how much to bid for an ad impression. This decision is based on massive amounts of data on individual consumers' characteristics and online behavior.

Contracts in the digital ad industry have evolved so that ad platforms only receive payment when specific user behavior—usually a click or purchase—can be linked to an ad impression (Feng and Xie 2012). Advertisers have moved from paying for ad impressions (CPM) to paying only for when consumers click on ads to reach their websites (CPC) to eventually paying ad platforms only when consumers addressed with ads conduct a predefined target action (CPA), usually a purchase (Hu et al. 2016).[1] Firms have now commonly implemented CPA-based contracts, in which they set their willingness to pay for each purchase from a consumer addressed with advertising. However, not all purchases by consumers targeted with ads correspond to actual value created through advertising: some consumers would purchase even without being served an ad.

The value of advertising—that is, the increment in profits caused by advertising—although well defined from a conceptual point of view, is difficult to assess in practice, especially for firms. Causally identifying the value of advertising requires platforms to run sophisticated randomized experiments using data not easily available to firms (Johnson et al. 2017a). Besides, ad platforms typically report only aggregate success measures such as the absolute number of purchases or clicks associated with an ad, which do not capture the actual benefit of advertising to firms, as they confound the causal impact of advertising with consumers' baseline probability to respond positively to ads.

In this paper, we build a stylized model to highlight the potential incentive misalignments between firms and ad platforms in CPA contracts and empirically assess this misalignment with the help of a large-scale field experiment. Our model shows that, although firms are more interested in addressing consumers who are more receptive to digital advertising, ad platforms are rewarded for targeting consumers with high overall purchase probability. If consumers with high baseline purchase probability are those for which ads work best, then the firm and ad platform agree on which consumers should be targeted, and the incentives are aligned. If, on the other hand, consumers with low baseline purchase probability are the ones for which ads work best, then the ad platform will target consumers that the firm does not want to target—the incentives will be misaligned. Although some studies have found evidence consistent with a positive relationship between baseline purchase probability and ad effectiveness (Johnson et al. 2017a), others have found no such pattern (Blake et al. 2015). Thus, the magnitude of the incentive misalignment between firms and ad platforms depends on the relationship between consumers' baseline purchase probability and ad effectiveness and is ultimately an empirical question.

To assess this relationship empirically, we run a large-scale randomized field experiment in ad retargeting. We partnered with a European e-retailer and randomly allocated 208,538 individual consumers to being served either retargeting ads of the e-retailer or orthogonal charity ads, also called public service announcements (PSAs). This way, we can identify the causal impact of retargeting advertising on consumers' purchase probability. Using these data, we estimate the difference in ad effectiveness for consumers who are more or less aggressively retargeted by the ad platform. We find that the ad platform acts according to the contracted incentives and hence systematically targets consumers who are more likely to purchase. Although advertising does generally increase consumers' purchase probability, we do not find evidence that consumers that are more aggressively targeted by the ad platform are more receptive to ads: there is no significant relationship between a consumer's baseline purchase probability and the increase in purchase probability caused by ads. This renders the ad platform's targeting strategy suboptimal for the firm.

This work contributes to the literature on digital advertising: we follow the calls for more research from the areas of information systems (Choi et al. 2020) and marketing (Gordon et al. 2021) by expanding the understanding of incentives specified in CPA contract designs and assessing the economic implications of these incentives for firms and ad platforms. First, we extend the understanding of the implications of incentives specified in CPA contracts for firms and ad platforms. Although previous research indicated that CPA contracts are unfavorable for firms (Johnson and Lewis 2015, Xu et al. 2016), our work provides a more nuanced picture of contracted incentives by unraveling that the implications of CPA contracts are contingent on consumers' baseline conversion probability, the incremental effect of ads, and the relationship between these two concepts. Our model clarifies that the assessment of the incentive misalignment requires firms to understand the distribution of consumers in the two-dimensional space of baseline purchase probability and advertising effectiveness.

Second, our empirical case allows us to investigate the contingencies that determine the implications of the incentives specified in CPA contracts. Our unique data from a large-scale randomized field experiment allows us to observe variation in the bids placed by the ad platform's optimization algorithm under a CPA contract without introducing selection bias into treatment and control group, overcoming a major issue for the identification of ad effectiveness discussed in related work (Johnson et al. 2017a, Gordon et al. 2019). Our analysis provides an empirical example of how firms are affected by CPA contracts contingent on consumers' baseline purchase probability and advertising effectiveness. Therefore, this analysis represents a clear translation of our theoretical evaluation of the problem to an actual empirical case.

Third, we provide an economic assessment of the incentive misalignment by estimating consumers' location in the two-dimensional space of baseline purchase probability and advertising effectiveness. Our welfare analysis provides evidence that the currently contracted incentives lead to a decrease in overall welfare, compared with a contract design that incentivizes ad platforms to target consumers based on the effectiveness of advertising. While related research discusses a move toward cost per incremental action (CPIA) contracts as a solution to the presented problem (Johnson and Lewis 2015, Lewis and Wong 2018), we provide evidence that ad platforms are likely to have no interest in switching to such a contract design. Under the current incentive scheme—that is, CPA contracts—ad platforms may be appropriating revenue from firms that CPIA contracts cannot compensate.

Fourth, we develop and describe a novel solution to the incentive misalignment in CPA contracts, namely *restricted cost per action* targeting (RCPA). In contrast to previously suggested solutions (Johnson and Lewis 2015, Xu et al. 2016, Lewis and Wong 2018), our solution can be implemented on the firm side without intervention from the ad platform. Our solution can help firms in restricting ad platforms to solely targeting consumers that are profitable to both the ad platform and the firm. Our theoretical description and empirical evaluation of the solution shows that this solution can lead to a significant improvement in return on investment (ROI) on ad spend for firms.

In summary, our work is the first to simultaneously (1) provide a clear theoretical framework and explanation for the incentive misalignment between the ad platform and the firm; (2) show that the actual magnitude of the incentive misalignment is contingent on the relation between baseline purchase probability and ad effectiveness; (3) delineate and execute an empirical strategy to recover the structural parameters of the theoretical framework and assess the actual incentive misalignment in a real-world empirical setting; (4) provide an estimate of the welfare loss due to such misalignment; and (5) propose a solution—inspired by the theoretical framework—that is feasible to be implemented by the firm, that is, without requiring an intervention of the ad platform.

This work has significant practical implications for the digital advertising industry. Although programmatic advertising and CPA-based contracts with ad platforms have become very popular, we show that they are likely not beneficial for firms. We show empirically that ad platforms follow the incentives specified in CPA contracts and target high-purchase-probability consumers to increase their profits. This behavior is harmful when firms pay more for ad impressions that are not more effective. Therefore, we present evidence showing that the current contract structure does not adequately serve firms' interests.

2. Related Literature

2.1. Advertising Effectiveness and Retargeting

In retargeting, ad platforms use consumers' browsing behavior on firms' websites to readdress consumers with relevant products on external sites (Lambrecht and Tucker 2013, Bleier and Eisenbeiss 2015). Compared with nontargeted banner advertising, using the observed browsing behavior of consumers presents advertisers with two main advantages. First, consumer browsing behavior allows advertisers to generate insights into what consumers are interested in. Using these insights, advertisers can present more relevant, personalized, advertising to consumers (Arora et al. 2008). Second, by only targeting consumers that have previously visited the firm's website, advertisers become less likely to address uninterested consumers with ads, potentially leading to a more efficient allocation of advertising budget (Schumann 2014). Both personalization and targeting can lead to more positive responses to advertising from consumers (Arora et al. 2008).

Studies on advertising effectiveness aim to answer the following question: how much more profit does a firm make when advertising than when not advertising (Lewis et al. 2015). In practice, the task of causally assessing the effect of advertising on firm performance is challenging because of three main factors. First, applying a method to identify the effect of advertising correctly remains technically challenging (Johnson et al. 2017a). Second, the ad-serving process involves multiple parties, creating difficulties for firms to access the information required to estimate ad effectiveness (Johnson and Lewis 2015). Third, the effect of advertising is usually small and explains only a fraction of the observed variation in consumer behavior, making it hard to estimate ad effectiveness precisely (Lewis and Reiley 2014b, Lewis et al. 2015).

In the advertising industry, contracts in which firms reward ad platforms for observed consumer actions, such as clicks, are common practice. Nonetheless, such an approach to success attribution introduces several issues. Ads can also affect consumers that do not click them. Tying ad effectiveness to clicks can result in an underestimation of ad effectiveness (Lewis et al. 2015). Next to that, some consumers would have returned to the advertiser's website without seeing and clicking on an ad, leading to an upward bias in the advertiser's estimate of advertising's contribution to generate website traffic (Johnson et al. 2017a). Another common pitfall in the industry is to tie rewards to the last click before the purchase, assuming that all advertising effectiveness comes from the last clicked ad. This may lead to an upward bias in the estimate of the ad effectiveness of the focal ad while underestimating the contribution of other advertising (Lewis et al. 2015). Related research has pointed out the existence of externalities between ads when displayed simultaneously (Agarwal and Mukhopadhyay 2016). Some firms try to overcome the limitations of observational approaches to assess ad effectiveness by taking advantage of quasi-experiments. Here, some consumers are targeted with ads, whereas others are not. However, these quasi-experiments often cannot guarantee that consumers in the treated group are not systematically different from consumers in the untreated control group, introducing selection bias into estimates for ad effectiveness (Gordon et al. 2019).

To overcome the limitations of observational approaches in assessing advertising effectiveness, research has moved to apply randomized field experiments in which consumers are randomly allocated to being exposed to advertising or not (Blake et al. 2015). Although randomized trials offer a more accurate and unbiased way to estimate ad effectiveness, implementing such experiments is still challenging. Lewis and Reiley (2014b) show that achieving precision in estimating ad effectiveness requires a large sample since the economic effects of ads are small. Adding covariates to model specifications and reducing noise in the control group by excluding observations for consumers that would not have been treated, can improve the precision of estimates (Johnson et al. 2017b). From a technical perspective, consumers in the control group can commonly not be identified when not confronted with ads, leaving the researcher with no counterfactual (Sahni 2015). One way to overcome this technical challenge is to serve ads to members of the control group that are assumed to be orthogonal to the firm's ads (commonly charity ads), whereas the rest of the consumers are exposed to the firm's regular ads. The difference between consumers' purchase probabilities in the treatment and control groups can then be interpreted as the causal effect of advertisements on consumers' purchase probabilities (Johnson et al. 2017b).[2] Johnson et al. (2017a) propose a new methodology—called ghost ads—in which ad platforms run hidden auctions for consumers in the control group to determine which consumers would have seen an ad if treated and should, therefore, be included in the control group. Although this methodology allows for saving costs associated with displaying orthogonal ads for the control group, it needs to be implemented on the ad platform's side. Furthermore, although ad platforms start to provide firms with information on campaign-level estimates for advertising effectiveness, it remains difficult for firms to access individual-level consumer data from well-designed experiments that allow for the estimation of heterogeneous treatment effects.

2.2. Heterogeneous Advertising Effectiveness

Heterogeneous treatment effects describe a significant variation in the effect an experimental treatment transmits on an outcome variable contingent on other pretreatment variables (Athey and Imbens 2017). Translated into the advertising context, heterogeneous ad effects describe a significant difference in how consumers respond to ads based on variables that are not influenced by or collected before the confrontation with advertising. Relevant pretreatment variables could be consumers' gender or visit frequency to the advertiser's website before the experiment. Interestingly, these pretreatment variables are often also good predictors of consumers' baseline conversion probability. For example, consumers that visit an e-retailer's website more often are more likely to purchase from the e-retailer independent of advertising.

Blake et al. (2015) find—with the help of a large-scale field experiment—that search engine advertising is only effective for new and infrequent users compared with the effect on frequent purchasers. Sahni et al. (2019) find that the effectiveness of retargeting in making consumers return to the advertiser's website decreases with an increase in the time since a consumer has last visited the advertiser's website. Zantedeschi et al. (2017) show in the context of direct marketing that targeting the most responsive customers compared with targeting customers based on previous purchases, significantly increases returns to advertising. Johnson et al. (2016) find that the effectiveness of ads decreases with an increase in the distance between consumers' place of residence and an advertiser's physical stores. Lewis and Reiley (2014a) find that older consumers' in-store sales are more affected by display advertising. Other work investigates heterogeneity of ad effectiveness contingent on the ad creative (Schwartz et al. 2017; Simester et al. 2020a, b).

In our work, we investigate the presence of heterogeneous treatment effects of advertising contingent on a consumer's baseline purchase probability. Related research often investigates the presence of heterogeneous treatment effects conditional on pretreatment variables that are predictors for consumers' baseline purchase probability. Our research focuses on a more direct assessment, by recovering consumers' baseline purchase probability from the bids placed by an ad platform on behalf of the firm.

2.3. Auctions, Contract Design, and Incentives in Programmatic Advertising

Auctions are used in a wide variety of contexts, ranging from spectrum licenses, electricity markets, flowers, and other consumer goods (Adomavicius et al. 2012; Lu et al. 2016, 2019). Similarly, most digital advertising slots are nowadays auctioned off in so-called real-time bidding auctions, in which advertisers—or ad platforms on their behalf—bid for the right to display an ad on a specific slot. Real-time auctions allow advertisers to better adjust the allocation of their marketing budget according to current traffic patterns or market conditions compared with the upfront purchase of advertising space (Balseiro et al. 2014). Ad platforms usually use second-price auctions to sell off their ad space (Edelman et al. 2007). The optimal strategy in second-price auctions is to bid in accordance with your true valuation (Vickrey 1961, Clarke 1971, Groves 1973). Therefore, these auctions result in an efficient allocation of ad spaces as the advertiser that values the ad slot the most wins the right to display an ad[3] (Arnosti et al. 2016).

In programmatic advertising, firms outsource the ad allocation process to ad platforms. These outsourcing relationships are governed by contracts that commonly aim to incentivize ad platforms to maximize the outcome of marketing campaigns. Nevertheless, over the last years, a discussion has emerged around whether these incentives are specified in firms' actual interest. Both practitioners and researchers have started to question whether ad platforms "cherry-pick" consumers likely to convert independently of the effect of advertising (Johnson and Lewis 2015). In technical terms, "cherry-picking" implies that ad targeting has become focused on distinguishing between consumers who buy and consumers that do not buy instead of targeting consumers who are more affected by advertising.

Dalessandro et al. (2012) discuss the issue of both ad effectiveness measurement and attribution and suggest that a good attribution system needs to align the incentives of both firms and ad platforms that are contracted to serve ads. Last-click attribution, that is, assigning the credit for a conversion to the ad platform that served the ad the consumer clicked last in the purchase process, is an example of these misaligned incentives. Here, advertising outlets have an incentive to confront consumers with ads as late as possible in their purchase process without taking into account an increase in consumers' purchase probabilities. Depending on the contract, ad platforms can then charge a fee from the advertising firm for reported conversions.

Related work has suggested incentivizing ad platforms to optimize the bidding for ad impressions based on actual contributed uplift in purchase probabilities (Xu et al. 2016). Other work focuses on the pricing models currently implemented in contracts between firms and ad platforms in which ad platforms are rewarded based on the wrong absolute outcome, instead of the incremental increase in the outcome variable (Johnson and Lewis 2015). More recently, technical solutions to optimize the bidding for ad impressions based on the incremental effect of advertising have emerged (Lewis and Wong 2018).

Previous research provides theoretical insights on whether firms should choose CPM- or CPC-based contracts (Asdemir et al. 2012) or on the social welfare implications of CPC- and CPA-based contracts for firms and publishers (Hu et al. 2016). In contrast, we focus on the implications of CPA-based contracts for firms and ad platforms from both a theoretical and empirical perspective. Our work extends the literature on contract design in the context of digital advertising, by presenting a stylized model that reveals the potential for incentive misalignments contingent on which consumers a contracted ad platform targets on behalf of the firm.

Although previous studies have highlighted the existence of an incentive misalignment in CPA contracts between the ad platform and the firm (Johnson and Lewis 2015, Xu et al. 2016), the focus of our work lies on investigating whether the incentives specified in CPA-based contracts lead to a de facto misalignment, and if so, in quantifying it. Our work furthers the literature by unraveling that the implications of CPA contracts are contingent on consumers' baseline conversion probability and its relationship to the incremental effect of ads. We make clear that—to assess the economic implications of CPA contracts—firms need to understand consumers' distribution in the two-dimensional space of baseline purchase probability and advertising effectiveness. In contrast to previous research (Lewis and Wong 2018, Xu et al. 2016), we propose a new solution to remedy the incentive misalignment that can be implemented on the side of the firm without further intervention from the contracted ad platform and without the need to adjust contracted incentives (Johnson and Lewis 2015).

3. Model on Ad Targeting and Incentive Misalignments

In this section, we present a model of how advertising affects consumers' purchase probability and how this is related to both the firm's and ad platform's interest to target consumers, given the CPA contract design commonly implemented in programmatic advertising.

Consumer i's utility for a product, u_i, can be expressed by

$$u_i = v_i + ad_i \cdot \theta_i + \epsilon_i,$$

where v_i represents consumer i's valuation of the product independent of being served advertising for the product, and θ_i represents the increase in utility caused by advertising. This increase in utility can come from, for example, lowering a consumer's search costs or providing useful information for the advertised product. The parameter ad_i denotes whether a consumer is exposed to advertising or not. If the firm decides not to advertise to the consumer ($ad_i = 0$) the consumer's utility consists of only v_i and ϵ_i, which is the idiosyncratic error component of a consumer's utility. We assume that ϵ_i is independent and identically distributed (i.i.d.).

In the context of programmatic advertising, firms outsource the ad allocation—that is, the decision to which consumers to advertise—to ad platforms. Ad platforms act as intermediaries in the relationship between firm and consumer, and ad_i becomes the decision variable for the ad platform in programmatic advertising.

Firms and ad platforms usually write CPA-based contracts, where ad platforms are rewarded for purchases by consumers that have been addressed with advertising. In these contracts, firms agree to pay the fee f to ad platforms for every purchase by a consumer i that has been addressed with at least one ad. Commonly, f is chosen uniformly for all purchases and not differentiated between different purchase instances. This means that for every purchase reported, the ad platform is rewarded with f.[4]

3.1. Ad Platform

Given the incentive structure of the CPA contract, the expected profit realized by the ad platform, Π^A, when serving an ad to a consumer i is a function of the consumer's inherent valuation of the advertised product, v_i, and the effect of advertising, θ_i:

$$E[\Pi^A(ad_i, v_i, \theta_i) \mid ad_i = 1] = Pr(y = 1 \mid ad_i = 1, v_i, \theta_i) \cdot f,$$

where $Pr(y = 1 \mid ad_i, v_i, \theta_i)$ is consumer i's overall purchase probability, which consists of the consumer's inherent valuation of the advertised product, her baseline purchase probability v_i, and the effect of advertising, θ_i.

In case the ad platform does not serve advertising to consumer i, the ad platform's expected profit is q, which represents the expected profit of serving the next best ad to consumer i (i.e., the ad of another firm):

$$E[\Pi^A(ad_i, v_i, \theta_i) \mid ad_i = 0] = q.$$

Therefore, the ad platform wants to serve advertising to a consumer if its expected revenue from advertising for the focal firm is higher than the opportunity cost of serving other ads to the same consumer, q. The ad platform is willing to serve the ad to consumer i if and only if (iff)

$$E[\Pi^A(ad_i, v_i, \theta_i) \mid ad_i = 1] \geq E[\Pi^A(ad_i, v_i, \theta_i) \mid ad_i = 0].$$

For simplicity, we assume the consumer's utility is linear[5]:

$$Pr(y = 1 \mid ad_i, v_i, \theta_i) = v_i + ad_i\theta_i.$$

Then, the ad platform is willing to serve the ad to consumers for which the expected profit of serving the ad is greater than the expected profit of serving the next best ad:

$$(v_i + \theta_i) \cdot f \geq q. \tag{1}$$

Equation (1) makes the incentive misalignment in CPA-based contracts explicit. The ad platform has incentives to target consumers based on their overall purchase probability: consumers' baseline purchase probability together with the effect of advertising on their purchase probability $(v_i + \theta_i)$. In contrast, the firm derives benefits solely from the increase in purchase probability caused by advertising (θ_i). Equation (1) also makes clear that the CPA contract design allows the ad platform to extract rents from the firm it did not create. The fee paid by the firm to the ad platform has two components: one associated with the added value created by advertising, $\theta_i f$, and one associated with the consumer's intrinsic likelihood to purchase, which would have been realized independent of the ad, $v_i f$. This latter component is appropriated by the ad platform without having contributed to the creation of the respective value.

3.2. Firm

The firm is interested in targeting consumers that are positively affected by the ad, that is, consumers with high θ_i. Under a CPA-based contract, the firm's expected profit for a consumer i, Π^F, targeted with advertising $(ad_i = 1)$ can expressed by

$$E[\Pi^F(ad_i, v_i, \theta_i) \mid ad_i = 1] = Pr(y = 1 \mid ad_i = 1, v_i, \theta_i) \cdot (r - f)$$
$$= (v_i + \theta_i) \cdot (r - f),$$

where r represents the profit that the firm can extract from a consumer that purchases.[6] Logically, the firm will always choose $f < r$. Otherwise advertising cannot have a positive return on investment for the firm. The expected profit from consumers that are not targeted can be expressed by

$$E[\Pi^F(ad_i, v_i, \theta_i) \mid ad_i = 0] = Pr(y = 1 \mid ad_i = 0, v_i, \theta_i) \cdot r = v_i \cdot r.$$

Therefore, from the firm's perspective, consumer i should be served advertising if the expected profit generated when advertising to this consumer is higher than when this consumer is not addressed with advertising, that is, iff

$$E[\Pi^F(ad_i, v_i, \theta_i) \mid ad_i = 1] \geq E[\Pi^F(ad_i, v_i, \theta_i) \mid ad_i = 0] \Longleftrightarrow$$
$$(v_i + \theta_i) \cdot (r - f) \geq v_i \cdot r. \tag{2}$$

When restructuring this equation, we see that the firm only wants to advertise to a consumer i when the cost of advertising for this consumer (c_i):

$$c_i = (v_i + \theta_i) \cdot f$$

is smaller than the return on advertising for this consumer, $\theta_i \cdot r$:

$$c_i \leq \theta_i \cdot r.$$

The firm wants to target only consumers for which the benefits of advertising overcome the cost imposed by the ad platform. These costs are based on the overall purchase probability $(v_i + \theta_i)$ and not on the increment in purchase probability caused by advertising (θ_i). Next to that, this equation shows that contracting with the ad platform based on θ_i, instead of $v_i + \theta_i$, would solve the problem and align the incentives. However, θ_i is usually assumed noncontractible, as the ad platform is frequently the only player potentially able to properly estimate its value.[7]

3.3. Incentive Misalignment

From Expressions (1) and (2), it follows that the ad platform and the firm are interested in targeting different sets of consumers. To visualize the incentive misalignment, we solve Equations (1) and (2) for θ_i.

This leaves us for the ad platform with the targeting boundary

$$\theta_i \geq \frac{q}{f} - v_i, \tag{3}$$

and the for the firm with the targeting boundary

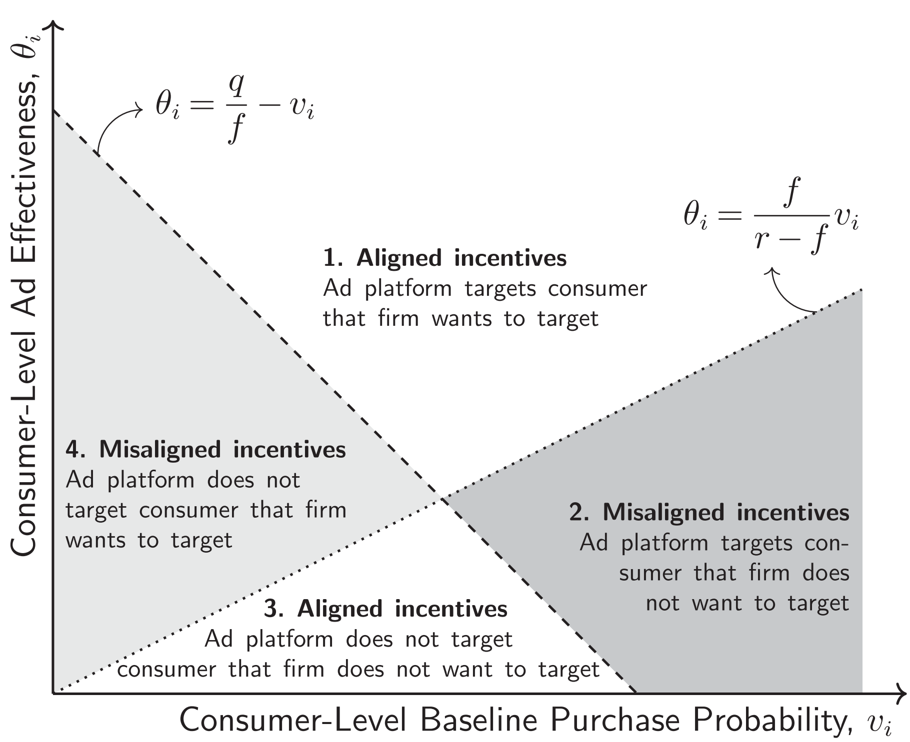
$$\theta_i \geq \frac{f}{r-f} v_i. \tag{4}$$

To visualize the conflicting interests in targeting between firm and ad platform, we plot the targeting boundaries (3) and (4) in the (v_i, θ_i) space. Figure 1 shows which consumers should be targeted in the interest of the firm and in accordance with the incentives for the ad platform.

Region 1 encompasses consumers that both the ad platform and the firm want to target. Region 3 encompasses consumers that neither the ad platform nor the firm want to target. Incentive misalignments arise when consumers are located in regions 2 and 4. Region 2 encompasses consumers that the ad platform targets but that the firm does not want to target. While these consumers are profitable to target for the ad platform, the added value from targeting them is lower than the cost of targeting them for the firm. Region 4 encompasses consumers that the ad platform does not target because of their low overall purchase probability but that the firm would like to target because of their relatively high sensitivity to ads.

The magnitude of the misalignment is a function of how consumers are distributed in the presented (v_i, θ_i) space. The monotonically increasing targeting boundary of the firm shows that if the impact of advertising on consumers' purchase probability (θ_i) is increasing with an increase in their baseline purchase probability (v_i), there might be no consumers that are interesting to target for the ad platform but not for the firm (region 2). If there are no consumers located in region 2, there is no negative effect of the incentives specified in the contract between firm and ad platform. We address this question empirically. With the help of a randomized field experiment, we can identify the effect of advertising on consumer's purchase probability. At the same time, we observe the ad platform's targeting behavior and analyze its implications on firms' benefits from advertising and overall welfare.

Figure 1. Regions with Targeting Conflicts



4. Empirical Strategy

Our empirical analysis has two goals. First, we are interested in estimating the effect of advertising on consumers' purchase probability. Second, and the focus of this paper, we want to measure the extent to which the potential incentive misalignment plays a role in our empirical setting, by assessing the relationship between θ and v.

We address our first goal by conducting a randomized field experiment. To assess the effect of advertising we need to induce random variation in whether consumers are addressed with advertising or not and observe their purchase behavior. As mentioned previously, the decision of whether to address a consumer with an ad (ad_i) is determined by the ad platform and thus endogenous. Moreover, because of the way programmatic advertising works, the ad platform only provides information to the firm for those consumers for which it has served the ad ($ad_i = 1$). Therefore, to identify the effect of advertising on consumers' purchase probability, we create exogenous variation in ad_i.

Practically, we induce this variation by randomly assigning consumers to seeing ads of the firm or ads for a charity organization, also called public service announcements (PSAs). Importantly, this variation is not the variation originating in the ad platform's decision to advertise to a consumer or not but rather variation within the group of consumers for which the ad platform has decided to advertise. Our identifying assumptions are as follows: (1) the charity ad is orthogonal to the firm's ad in its effect on consumers' purchase probability, meaning the charity ad does not affect consumers' propensity to purchase from the focal firm; (2) the ad platform assumes the same ad effectiveness for real ads and charity ads and does not select different types of consumers to see the firm's ads and charity ads; and (3) the ad platform does not adjust its estimate for ad effectiveness asymmetrically for treatment and control group; that is, it does not deviate in its targeting strategy between treatment and control group over time. Although we are unable to test assumption (1) directly, the approach of using charity ads and assuming they are orthogonal to the firm's ads has been widely used in the literature and is common practice in identifying ad effects in digital environments (Hoban and Bucklin 2015, Lewis and Rao 2015, Sahni 2015, Johnson et al. 2017b). We test assumption (2) by performing randomization checks on multiple covariates of interest and checking whether the allocation probability to the treatment remains constant across different heights of bids (see Online Appendix C). We test assumption (3) by comparing whether the ad platform adjusts its bidding for ad impressions differently between control and treatment group over time. Results in Online Appendix D support this assumption and point toward a symmetric optimization of the bidding for consumers in the treatment and control group.

The second goal of our empirical analysis relates to assessing the potential incentive misalignment for firms. This incentive misalignment is contingent on both the targeting behavior of the ad platform, as well as the relationship between v_i and θ_i. We assess the targeting behavior of the ad platform by observing how much the ad platform bids for ad impressions for different consumers. To assess the relationship of v_i and θ_i, we assume that θ_i is a linear function of v_i:

$$\theta_i \equiv \beta_2 + \beta_3 v_i. \tag{5}$$

Imposing this structure allows us to rewrite the equation for consumers' purchase probability and reduce the number of degrees of freedom: we now have only one unobserved variable, v_i. We estimate the probability that a consumer purchases ($y = 1$) as a function of how much s/he values the product, v_i, and how sensitive s/he is to the ad, θ_i:

$$Pr(y = 1 \mid ad_i, v_i, \theta_i) = v_i + ad_i\theta_i.$$

Given our assumption that θ_i is a linear function of v_i, we get

$$Pr(y = 1 \mid ad_i, v_i, \theta_i) = v_i + \beta_2 ad_i + \beta_3 ad_i v_i. \tag{6}$$

In this specification, β_2 represents the effect of serving an ad on the purchase probability, and β_3 represents the change in θ_i contingent on v_i. The β_3 coefficient also represents the extent to which there is an incentive misalignment between the ad platform and the firm. If β_3 is positive, consumers with a higher valuation for the product are also those that are more sensitive to advertising. Consequently, the consumers with higher θ_i that the firm wants to target are also those with higher $v_i + \theta_i$. This would result in a relatively low degree of incentive misalignment. If, on the contrary, β_3 is negative, the incentive misalignment is significant.

To estimate Equation (6), we would need to know v_i. As we do not directly observe this value, we take advantage of our setting and indirectly use the ad platform's own estimate, $\hat{v}_i$, to estimate our parameter of interest, β_3. We do so by leveraging the fact that the ad platform's bid is a function of the ad platform's own estimate of v_i and β_3.

Most online ad impressions are auctioned off in second-price sealed-bid auctions (Arnosti et al. 2016), resulting in an efficient outcome from the ad platform's perspective: the ad with the highest expected return is the one that is displayed to the consumer. For these auctions, the optimal bid equals the bidder's true valuation of the impression (given that one ad is auctioned off at a time; Edelman et al. 2007). In other words, the ad platform's optimal bid for an ad impression for a consumer can be expressed by

$$bid_i^A = \widehat{Pr}(y = 1 \mid ad_i^A = 1, v_i, \theta_i) \cdot f, \tag{7}$$

in which $\widehat{Pr}(y = 1 \mid ad_i^A = 1, v_i, \theta_i)$ is the purchase probability of consumer i estimated by the ad platform assuming the ad is served, and ad_i^A represents the decision by the ad platform of whether to serve the ad or not. We distinguish between ad_i^A, that is, whether the ad platform has decided to serve the firm's ad, and ad_i, that is, whether the firm's ad is actually served. Otherwise, the consumers see PSA ads. We simplify the notation and assume bid_i^A is normalized with $f = 1$.[8]

Our empirical context allows us to observe the bid placed by the ad platform on behalf of the firm. From this, we can recover the ad platform's estimate for how much the consumer values the product, $\hat{v}_i$, as a function of the bid, and the ad platform's estimates for the parameters β_2 and β_3, $\hat{\beta}_2$ and $\hat{\beta}_3$, respectively.[9] The ad platform bids under the assumption that it is serving a valid ad on behalf of the firm, that is, that $ad_i^A = ad_i = 1$ for all ads served. Therefore,

$$bid_i^A = \hat{\beta}_2 + (1 + \hat{\beta}_3)\hat{v}_i \Longleftrightarrow$$
$$\hat{v}_i = \frac{1}{1 + \hat{\beta}_3} bid_i^A - \frac{\hat{\beta}_2}{1 + \hat{\beta}_3}.$$

Replacing v_i with the expression for $\hat{v}_i$ in the equation for consumers' probability of Purchase (6), we get

$$Pr(y = 1 \mid ad_i) = -\frac{\hat{\beta}_2}{1 + \hat{\beta}_3} + \frac{1}{1 + \hat{\beta}_3} bid_i^A + \left(\beta_2 - \frac{\beta_3 \hat{\beta}_2}{1 + \hat{\beta}_3}\right) ad_i + \frac{\beta_3}{1 + \hat{\beta}_3} ad_i \cdot bid_i^A,$$

or

$$Pr(y = 1 \mid ad_i) = \varphi_0 + \varphi_1 bid_i^A + \varphi_2 ad_i + \varphi_3 ad_i \cdot bid_i^A, \tag{8}$$

where $\varphi_0 \equiv -\hat{\beta}_2/(1 + \hat{\beta}_3)$, $\varphi_1 \equiv 1/(1 + \hat{\beta}_3)$, $\varphi_2 \equiv \beta_2 - \beta_3\hat{\beta}_2/(1 + \hat{\beta}_3)$, and $\varphi_3 = \beta_3/(1 + \hat{\beta}_3)$.

Assuming that the ad platform's estimates for the parameters of interest, namely β_2 and β_3, are consistent, we get $\varphi_0 = -\beta_2/(1 + \beta_3)$, $\varphi_1 = 1/(1 + \beta_3)$, $\varphi_2 = \beta_2/(1 + \beta_3)$, and $\varphi_3 = \beta_3/(1 + \beta_3)$. With these, we can estimate Equation (8) and recover the structural parameters from the reduced-form parameters. In general, just estimating Equation (8) and naively interpreting the reduced-form parameters as the structural parameters lead to a bias in the estimates for β_2 and β_3 proportional to the magnitude of β_3. This bias can result in attenuation (when $\beta_3 > 0$) or inflation (when $\beta_3 < 0$). If β_3 is close to zero, this bias is negligible.

Although the ad platform has an incentive to target consumers when $(v_i + \theta_i) \cdot f \geq q$, it is likely that the ad platform considers only v_i when deciding whom to target. This assumption seems plausible as the magnitude of ad effectiveness has been found to be small in digital advertising (Lewis and Rao 2015) and is likely to be relatively small compared with v_i.[10]

If the ad platform only takes consumers' valuation, v_i, into account, or, in other words, if the ad platform assumes $\hat{\beta}_2 = \hat{\beta}_3 = 0$, then Equation (8) is reduced to

$$Pr(y = 1 \mid ad_i) = bid_i^A + \beta_2 ad_i + \beta_3 ad_i \cdot bid_i^A. \tag{9}$$

For the reasons mentioned previously, the reduced-form equation can provide a good approximation to the structural parameters. However, there's no structural interpretation for the parameters φ_0 and φ_1 in the reduced-form regression. In fact, only φ_2 and φ_3 have a structural meaning and can be identified.

5. Experiment and Analysis

We conducted a large-scale randomized field experiment in collaboration with a major European e-retailer focusing on selling consumer electronics. The firm sells a wide range of products including home appliances, laptops, smartphones, and cameras and follows an online-first strategy: the vast majority of their sales happen online, with their physical stores being used mostly for customer service. The firm is a major player in the regions it operates. To protect our partner firm's anonymity, we refer to our partner only as the "firm."

The firm's main goal for the experiment was to investigate the effectiveness of its display retargeting advertising campaigns to generate sales. Retargeting uses consumers' browsing behavior on the firm's website to show products that a consumer has browsed on external sites (Lambrecht and Tucker 2013, Bleier and Eisenbeiss 2015). This ad targeting strategy focuses on consumers that have already shown interest in the firm's products. This context is especially valuable to investigate a potential incentive misalignment between the firm and the ad platform. Ad platforms have rich data on these consumers, which can be used to inform the targeting of consumers and optimize the bidding for ad impressions.

The experiment ran in late spring in 2016 for 49 consecutive days. Consumers who had visited at least one product category page of our partner firm's website within the last 14 days were eligible to participate. This time window was updated on a rolling basis, meaning that consumers could leave or re-enter the experiment based on their activity. Participating consumers were randomly allocated to either being treated with retargeting ads (80% probability of assignment) or with PSAs (20% probability of assignment) that advertised the donation to a charity. Ads and PSAs were displayed on third-party websites. A total of 208,538 consumers, identified by cookies, participated in the experiment. Consumers remained in the same treatment group for the whole duration of the experiment.

Table 1. Descriptive Statistics

Variable	Observations	Mean	Standard deviation	Minimum	Maximum
Purchase	208,538	0.05547	0.22890	0	1
Visit	208,538	0.36744	0.48211	0	1
Ad	208,538	0.80031	0.39977	0	1
Bid	208,538	0.01344	0.01790	0.00016	0.09124
Cost	208,538	0.00419	0.00554	0.00001	0.09082
Impressions	208,538	20.10890	15.80474	1	221

Our partner firm decided on the 80%–20% split between treatment and control group to keep the experiment costs low, as PSAs cost the same as retargeting ads but generate zero return.[11] PSAs are assumed to be orthogonal to our retargeting ads in their effect on consumers, allowing us to measure the causal impact of retargeting ads (Johnson et al. 2017a).[12]

Our partner firm used a tool made available by the contracted ad platform that allowed for user segmentation based on cookies. The contracted ad platform is among the largest ad platforms in ad inventory and revenue and is representative of the market in terms of its practices. The tool allowed for a random allocation of consumers that visit our partner's website to either treatment or control group, by placing the respective cookies on consumers' devices. Upon a retargeting opportunity, the ad platform was placed a bid to serve an ad on behalf of the firm. Crucially, the tool does not use information about whether a consumer is in the treatment or control group when determining its bid for an ad impression. Although individuals were randomly assigned to either treatment or control groups, the optimization was done jointly, thus making these consumers comparable in all dimensions.

To assess the firm's advertising effectiveness, we obtained data on purchases by consumers and whether they returned to the firm's website after being treated. For our analysis, we aggregate two main data sources. An impression-level data set gives us information on all impressions displayed per individual consumer, how much the ad platform bid for an ad impression, and how much the impression cost, the second-highest bid. An activity-level data set provides information on the consumer's activity on the firm's page. Table 1 presents the descriptive statistics for our consumer-level data set.

The *purchase* variable is a binary variable indicating whether a consumer purchased after seeing an ad (or PSA for the control group). The *visit* variable is a binary variable indicating whether a consumer returned to the firm's website after seeing an ad. *ad* is the binary treatment indicator, with $ad_i = 1$ if a consumer is in the treatment group. *bid* represents the first successful bid placed for a consumer and *cost* the cost of the first ad impression.[13] *impressions* gives the number of impressions a consumer received throughout the experiment.

Table 2 shows the count statistics for our experiment split for treatment and control group. Model-free evidence shows that consumers that are treated with advertising are significantly more likely to purchase ($\Delta M^{purchase} = 0.0024$, $t = 1.9633$, $p = 0.0496$). Furthermore, we find that advertising has a significant and positive effect on consumers' probability to return to the firm's website ($\Delta M^{visit} = 0.0178$, $t = 6.7868$, $p < 0.0001$).

In Table 3, we compare treatment and control groups along a number of variables that can serve as randomization checks. We find no significant difference between treatment and control group for the bid, the cost for the first ad impression, or the number of impressions served to a consumer over the experiment's duration. We also compare consumers in the treatment and control group along a number of variables related to consumers' behavior before the first ad impression. We compare the number of tracked activities, visits, and activity duration before the first ad treatment and find no difference between control and treatment groups. We also investigate whether there is a difference between consumers regarding the types of pages visited before the start of the experiment. We find no difference in the number of visited product categories, visited product pages, and the number of shopping cart visits. This supports our assumption that consumers are randomly allocated to treatment and control groups.

The histogram in Figure 2 shows a significant variation of the bid placed by the ad platform, with a coefficient of variation (ratio of standard deviation to mean)

Table 2. Count Statistics for Control and Treatment Group

Group	Assignment probability	Consumers	Impressions	Purchases	Visits
Treatment	0.8	166,895	3,355,613	9,339	61,918
Control	0.2	41,643	837,856	2,229	14,708
Total	1.0	208,538	4,193,469	11,568	76,626

Table 3. Comparison Control and Treatment Group: Randomization Checks

	Mean			
Variable	Control	Treatment	t statistic	p value
Bid	0.0134	0.0134	0.2403	0.8101
Cost	0.0042	0.0042	0.6282	0.5298
Impressions	20.1200	20.1061	0.1610	0.8721
Activities	5.8074	5.8515	0.8757	0.3812
Visits	1.3580	1.3618	0.6063	0.5443
Activity duration	7.1067	7.0947	0.1400	0.8887
Number of visited product categories	1.7877	1.7994	0.6643	0.5065
Number of visited product pages	2.8329	2.8849	1.9236	0.0544
Number of shopping cart visits	0.1115	0.1098	0.3210	0.7482

larger than one ($CV = 1.332$). This variation is crucial for our identification strategy.[14]

We move forward to estimating our theoretical model. We assess whether the ad platform behaves according to the incentives laid out in the theory section and how consumers with different baseline purchase probabilities react to advertising.

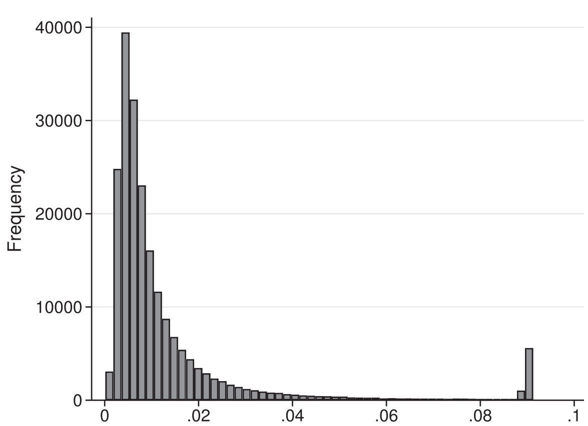
5.1. Main Results

We first estimate the reduced-form linear probability model (LPM) from Equation (9):

$$Pr(y = 1 \mid ad_i) = \varphi_0 + \varphi_1 bid_i^A + \varphi_2 ad_i + \varphi_3 ad_i \cdot bid_i^A + \epsilon_i,$$

where bid_i^A represents the first successful bid for the opportunity to display an ad to a consumer i placed by the ad platform in Euro, ad_i represents a binary variable indicating whether a consumer was addressed with retargeting ads ($ad_i = 1$) or PSA ads ($ad_i = 0$), and ϵ_i represents the idiosyncratic error term. The interaction between bid_i^A and ad_i represents the focal aspect of this analysis. Under the assumption that the ad platform optimizes its bidding based only on consumers' baseline purchase probability (and not on the ad effectiveness), the coefficient φ_3 has the same interpretation as in Equation (5): it represents the relationship between baseline purchase probability and ad effectiveness. If this coefficient is positive, then we can conclude that consumers with a higher likelihood of purchasing are also those that are most receptive to ads. This means that optimizing only based on purchase probability is not necessarily bad, as it goes hand in hand with ad effectiveness. If, on the other hand, this coefficient is zero or negative, it means that targeting based on purchase probability is not in line with the firm's interest.

Figure 2. Histogram Bid



Columns (1) and (2) in Table 4 show the results of the linear probability model estimates. We find that the ad platform bids higher for consumers that are more likely to purchase, as visible by the positive and significant coefficient for *bid*. We also find evidence for a positive effect of the ad treatment (ad_i). This effect is significant at the 10% level when we omit the interaction between *ad* and *bid* (column (1)) and significant at the 5% level when including it (column (2)). We find a negative but not statistically significant effect for the interaction between *ad* and *bid*.[15] This outcome stands against the notion that consumers with a higher baseline purchase probability (and higher bid) are more affected by advertising.[16] When estimating ad effectiveness operationalized as consumers' probability of returning to the firm's website after seeing an ad, an upper-funnel success measure, we find consistent results (see Online Appendix E).

Strictly speaking, and according to our theoretical model, the bid represents an endogenous decision made by the ad platform, which means its coefficient is not identified.[17] The coefficient should be considered only as an additional control variable. In line with the model derived in the theory section, we also run our estimation using the specification from Equation (9) by constraining the coefficient for *bid*. These estimates, presented in column (3) of Table 4, are very similar to those in column (2), with the main difference being that the coefficients for ad effectiveness and interaction with bid are now statistically significant at the 1% level. We find again that although advertising has a positive impact on consumers' purchase probabilities, consumers that receive higher bids seem to be less receptive to ads.

Table 4. Reduced-Form and Structural Regressions LPM

	Dependent variable: *Purchase*				
	Reduced form			Structural form	
	(1)	(2)	(3)	(4)	(5)
Bid (φ_1)	0.8328***	0.9310***	1	0.9310***	1
	(0.0371)	(0.0839)	(*constrained*)	(0.0839)	(*constrained*)
Ad (β_2)	0.0024*	0.0041**	0.0050***	0.0036***	0.0016
	(0.0012)	(0.0016)	(0.0013)	(0.0012)	(0.0015)
Ad × *bid* (β_3)		−0.1227	−0.1917***	−0.1093	0.0741
		(0.0935)	(0.0414)	(0.0742)	(0.0968)
Constant	0.0423***	0.0410***	0.0401***	0.0410***	0.0410***
	(0.0012)	(0.0014)	(0.0011)	(0.0014)	(0.0014)
Observations	208,538	208,538	208,538	208,538	208,538
R^2	0.004	0.004	0.004	0.004	0.004
Adjusted R^2	0.004	0.004	0.004	0.004	0.004

Note. Robust standard errors in parentheses.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

As mentioned, these findings can only be interpreted under the assumption that the ad platform does not take the ad effectiveness (θ_i) into account when deciding on its bids. Although this assumption seems plausible, we move forward by considering the case where the ad platform targets consumers based on both consumers' valuation of the product (v_i) and the ad effectiveness (θ_i).

Considering that the ad platform takes θ_i into account when deciding whom to target leads us to Equation (8). We estimate this equation with the help of a nonlinear least-squares estimator. We start by allowing parameter φ_1 to be free and then repeat the estimation with a constrained φ_1 (columns (4) and (5) of Table 4, respectively).

Again, we find no evidence for a positive β_3. Although structural estimates are imprecise for the constrained structural specification (column (5)), the unconstrained structural specification shows a statistically significant coefficient for ad effectiveness at the 1% level. The interaction coefficient remains negative and not significant. Moreover, the coefficients in the unconstrained structural estimates (column (4)) are very close to the coefficients of the respective reduced-form estimates (column (2)). This similarity is an indication that β_3 is not causing a large bias in the results of the reduced-form estimates. This means that even when giving the ad platform "the benefit of the doubt," that is, that it takes the ad effectiveness (θ_i) into account when deciding how much to bid for consumers' ad impressions, we find no evidence that the targeting is beneficial for the firm.

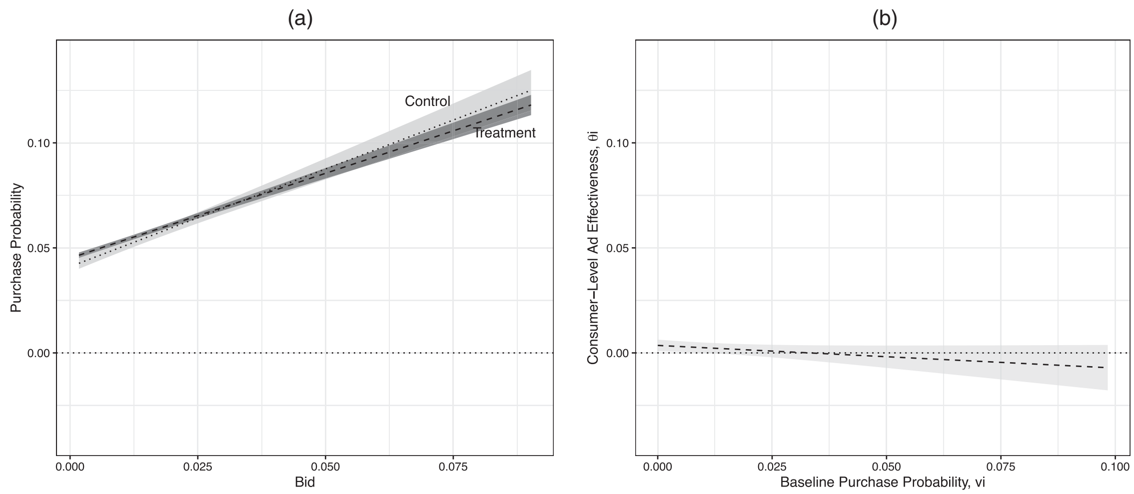
Although our specification in column (3) yields a negative and statistically significant interaction effect between *ad* and *bid*, both the reduced-form and unrestricted structural estimates are imprecise. A power analysis for our experiment indicates that while we are sufficiently powered to detect the effect of advertising, we are underpowered to detect the respective effect for the interaction of *ad* and *bid* (see Online Appendix G). The lack of precision of these estimates is in line with studies that have shown the difficulty of estimating the effects of digital advertising (Lewis and Rao 2015). This lack of precision makes it difficult to interpret our results, especially our coefficient of interest, β_3.

Therefore, we move toward the interpretation of the confidence interval of the effect of advertising conditional on the bid placed by the ad platform and the economic interpretation of this confidence interval. We create a visual representation of our reduced-form estimates along with 95% confidence intervals that allow for a more intuitive economic interpretation of our findings.

Figure 3(a) depicts the estimated purchase probability for treated and control users as a function of the ad platform's bid. The positive slopes of the estimated purchase probabilities for both treatment and control group corroborate the idea that the ad platform targets consumers with higher absolute purchase probability—the ad platform can predict and bid higher for consumers with a higher purchase probability. According to our estimates, the ad effectiveness is positive and statistically significant for consumers that receive low bids. As visible in the crossing of the lines for the estimated purchase probabilities and the overlap in confidence intervals for higher bids, the ad effect vanishes for consumers that received higher bids. Although the ad platform targets consumers that are likely to purchase independently of the firm's advertising by bidding higher for their impressions, the targeted consumers seem to be unaffected by the firm's advertising.

Figure 3(b) shows the estimates for the consumer-level ad effectiveness (θ_i) as a function of the baseline purchase probability (v_i). This line corresponds to the difference between treatment and control estimates

Figure 3. Visualization of Ad Effectiveness



Notes. (a) Estimated purchase probability for treatment and control as a function of bid. (b) Estimated ad effectiveness as a function of estimated baseline purchase probability

presented in Figure 3(a) but displayed in the (v_i, θ_i) space.[18] We notice that the magnitude of the ad effectiveness is small compared with the baseline purchase probability (Figure 3(a)), on average a 4.5% increase in consumers' purchase probability. This figure shows that ad effectiveness is positive and statistically significant for consumers with very low baseline purchase probability (low v_i). The point estimate for ad effectiveness decreases and loses significance with an increase in consumers' baseline purchase probability. Moreover, the estimated upper bound of the 95% confidence interval for the ad effectiveness remains virtually unchanged. Moving beyond the interpretation of the nonsignificant estimate of the coefficient for β_3 by considering the visualization of the confidence interval, it is highly unlikely that ad effectiveness increases with baseline purchase probability. Thus, although our estimates for ad effectiveness are relatively imprecise, by integrating baseline purchase probability with ad effectiveness, we see that these two quantities are unlikely to have a positive relationship.

5.2. Robustness Checks and Additional Analyses

We conduct several robustness checks and additional analyses to ensure the robustness of our main results presented in Table 4. First, we repeat the estimation process using a logit link function. In Online Appendix A, we present the respective equations and the estimations. The results are qualitatively similar to those obtained with the LPM estimates.

To further assess the robustness of our findings, we explore the possibility that we do not find a significant interaction effect in our analysis because our experiment has low power. We assess the functional form of the relationship between bids placed by the ad platform and the increase in consumers' purchase probabilities (see Online Appendix H). We find no evidence for an increasing trend in the relationship between bids placed by the ad platform and ad effectiveness for both purchases and visits. This finding gives us confidence that the reason for not finding a significant coefficient for the interaction between *bid* and *ad* is not an under-powered experiment.

To strengthen the argument that there is no significant correlation between consumers' baseline purchase probability and the increase in purchase probabilities caused by ads, we build a predictive model estimating consumers' purchase probabilities prior to the ad treatment (see Online Appendix I). The purpose of this analysis is to investigate the relationship between consumers' baseline purchase probabilities, instead of the bid placed by ad platforms, and the increase in purchase probabilities caused by ads more directly. We use these predicted purchase probabilities to test whether consumers with a higher predicted purchase probability react more positively to our ad treatment. Consistent with our claim, we find no evidence for consumers with higher predicted purchase probabilities being influenced more positively by the ad treatment.

We run additional robustness checks in which we analyze the relationship between the mean bid, the maximum bid, the median bid, and the cumulative bid placed for an ad impression for consumer *i* and their impact on consumers' purchase probability (see Online Appendix J). When using a different operationalizing of the ad platform's targeting behavior, that is, the ad platform's bidding for ad impressions on behalf of the firm, our results remain robust.

Notably, auction participants do not pay their actual winning bid to serve an ad impression but are charged the second highest bid in the ad auctions. Our results remain robust when analyzing the impact of the cost of the first impression, the average cost of an impression per consumer, and the overall cost for impressions per consumer (see Online Appendix K).

One explanation that would justify the ad platform's bidding behavior is that the ad platform's optimization algorithm incorporates consumers' profit contributions (r) into its optimization. In case the profit contribution is negatively correlated with the increase in consumers' purchase probabilities, $corr(r, \theta_i) < 0$, there might be a valid reason for the algorithm to not bid higher for more receptive consumers but instead target consumers with higher profit potential. In Online Appendix L, we provide evidence indicating that the ad platform does not refrain from targeting high θ_i consumers for the sake of targeting consumers with high profit potential.

To further assess the robustness of our findings, we re-estimate our main model and include dummy variables to control for the ad format and ad placement of the first ad impression to account for the fact that the ad platform might bid differently contingent on the ad quality. Additionally, we re-estimate the model including time controls for hour of the day, day of the month, and day of the week of the first impression as dummy controls to control for heterogeneity in ad attractiveness based on timing. Online Appendix F presents these results. The focal coefficient estimates remain practically unchanged when controlling for heterogeneity in formats and timing of ads.

6. Welfare Analysis and Economic Implications

In this section, we perform a welfare analysis to provide an economic interpretation of the structural parameter β_3, which is the parameter for the relationship between v_i and θ_i. We estimate the differences in profits and total welfare between a regime in which consumers are targeted according to the firm's interest—hereafter the *firm regime*—and the status quo in CPA contracts, in which the ad platform decides which consumers to target—the *ad platform regime*.

To perform the welfare analysis, we need to understand which consumers would be targeted under the firm regime and which consumers would be targeted under the ad platform regime. To do so, we need to estimate the position of each consumer in the (v_i, θ_i) space. Furthermore, we need the values of the CPA fee f that the firm pays to the ad platform per reported purchase, the average profit r the firm attains per sale, and the opportunity cost q of targeting a consumer (see targeting boundaries in Figure 1). We populate these values with the information from our empirical case.

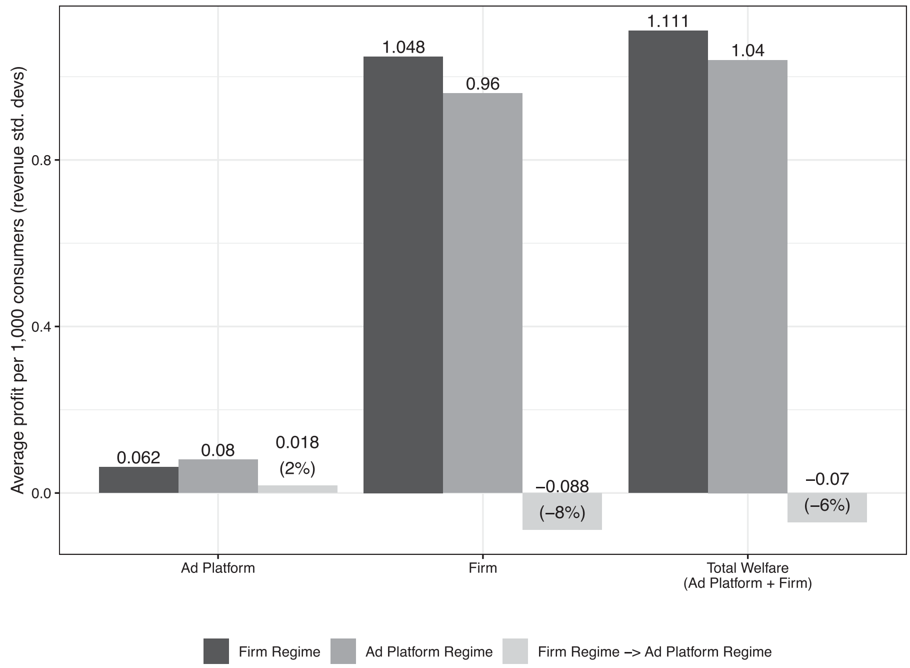
To protect the financial information of our partner firm, we standardize financial variables by dividing them by the standard deviation of the firm's revenue generated from purchases by consumers participating in the experiment. This linear transformation allows us to interpret percentage changes in the monetary outcome variables without revealing the firm's absolute revenue and profit figures. The firm pays about 2% of a standard deviation of the firm's revenue per reported purchase ($f = 0.02$). The average profit per purchase is about 27% of a standard deviation of the firm's revenue ($r = 0.27$). We use the median ad cost in our data as value for the ad platform's opportunity cost ($q = 0.003$).[19]

We recover v_i from the ad platform's first bid in line with Equation (8). Without further assumptions, we would be unable to recover θ_i. Therefore, we use our estimate for β_3, the correlation between v_i and θ_i and assume θ_i is normally distributed around this estimate with a standard deviation of 0.02,[20] approximately two times the standard deviation of the baseline purchase probability, v_i.[21]

Figure 4 shows how the average profit and total welfare per 1,000 consumers are affected by moving from targeting only the consumers the firm wants to target (firm regime) to the consumers the ad platform targets (ad platform regime). Importantly, the welfare analysis takes into account profit from purchases conducted by consumers that would have happened independent of being confronted with advertising, related to consumers' baseline purchase probability v_i. The horizontal axis of the figure distinguishes between the two stakeholders—the ad platform and the firm—and the total welfare. The differently colored bars represent results from the different regimes and their difference in percent of the total welfare. Moving from the firm regime to the ad platform regime leads to an increase in the ad platform's profit, at the expense of the firm's profit. Interestingly, the move leads to a decrease in total welfare. More specifically, the move to the ad platform regime results in a welfare loss of about 6% when compared with the firm regime. The currently common ad platform regime, that is, CPA contract design, leads to an inefficient outcome. The welfare loss originates from the fact that the ad platform regime—that is, a typical CPA contract—leads the ad platform to target consumers that are more likely to purchase, not those for which the ads are more effective. This results in an inefficient allocation of the ads—as they are not delivered to the consumers that would benefit the most from them—leading to an overall lower effect of advertising and consequently to an overall loss in welfare. The firm is the sole bearer of this loss, as the ad platform is actually better off in this regime.

Finally, we assess the profitability of our specific campaign from the perspective of the firm. Given the difference of 0.24 percentage points in conversion rates

Figure 4. Welfare Analysis—Comparison of Average Profit and Total Welfare per Consumer Under Firm and Ad Platform Regime



between treated and nontreated consumers, and an average profit per sale of 0.27 of a standard deviation of the firm's revenue, we find that the extra sales correspond to an increase in the firm's profits by about 0.66 of a standard deviation per 1,000 treated consumers or a total of 110.88 standard deviations of the firm's revenue.

However, the ad platform charged a fee of 0.02 of a standard deviation of the firm's revenue per reported purchase, which represents an average fee of 1.24 standard deviations per 1,000 treated consumers, or a total of 206.96 standard deviations of the firm's revenue for the treated consumers in the experiment. This means that the firm lost a total of 96.09 standard deviations of its revenue. Thus, from the total of 206.96 standard deviations appropriated by the ad platform in this campaign, only 53.6% correspond to value created by the ad platform itself. The remaining 46.4% correspond to fees paid to the ad platform originating in purchases that would have happened anyway.[22]

7. Restricted CPA Targeting

In this section, we propose a strategy to mitigate one of the sources of the incentive misalignment presented in our theoretical framework. Although the firm cannot influence the misalignment that originates from the ad platform not targeting consumers the firm wants to target (region 4 in Figure 1)—that is, the firm is not directly able to force the ad platform to target consumers the ad platform does not want to target—we propose a strategy that helps the firm to mitigate the misalignment caused by the ad platform targeting consumers the firm does not want to target (region 2 in Figure 1). We call this new strategy *restricted CPA* targeting (RCPA).

RCPA requires having access to two main sources of data: (1) user-level pretreatment characteristics (e.g., CRM data, demographics, email campaigns, user clickstream data on the website, etc.) and their purchase behavior; and (2) an experimental context in which consumers are randomly allocated to either the treatment or control groups with respect to whether they are exposed to the real ad or the PSA. The first source of data—user-level characteristics and purchase behavior—is readily available to many firms, especially in the context of retargeting, as consumers have usually already revealed some of their information and behavior to the focal firm (advertiser). One could also imagine a setting in which firms target already existing customers making use of information from a customer relationship management (CRM) system. The second source of data—the experimental context—can be the same as the one used by our focal firm in our study, which is readily available from the most popular ad platforms. Thus, firms do have access to this tool, making it possible to collect all the information required to implement our proposed solution.

Under RCPA, the firm uses information about the baseline purchase probability, v_i, and the ad effectiveness, θ_i, for each individual consumer to technically restrict[23] the ad platform in which consumers are eligible to be targeted.[24] To identify the consumers in region 2 of Figure 1, that is, the consumers that are profitable to target for the firm, the firm needs to take both the cost and benefit of targeting consumers into account. The firm's cost of targeting a consumer *i* with advertising is

$$c_i = (v_i + \theta_i) \cdot f.$$

The firm's benefit, π_i, of targeting a consumer i, that is, the additional profit generated because a consumer is addressed with advertising, is

$$\pi_i = \theta_i \cdot r.$$

The firm should allow the ad platform to only target consumers for which

$$\pi_i > c_i.$$

To determine which consumers the firm should allow the ad platform to target, the firm needs to estimate consumers' baseline purchase probability, v_i, and how consumers are affected by advertising, θ_i. Using these parameter estimates, the firm can predict the cost of addressing a consumer, c_i, and the benefit of addressing a consumer, π_i.[25] The accuracy of these estimates determines to what extent the firm is able to successfully restrict the ad platform from targeting consumers in region 2 of Figure 1.

We empirically test RCPA by simulating how the ad allocation would play out under RCPA. First, we build machine learning models that predict consumers' baseline purchase probability and ad effectiveness based on information available to the firm.[26] To avoid overfitting, we train our machine learning models to predict the two decision parameters on a randomly selected training data set and assess the effectiveness of the restrictions on a separate evaluation data set. This approach resembles a firm using historical data to predict which consumers to target in the future. Next, we simulate the impact of RCPA by distinguishing between three restriction approaches. The distinction between three RCPA approaches allows us to get further insights into the impact of cost- and benefit-based restrictions and the impact of their combination.

More specifically, for the (1) *heterogeneous baseline purchase probability with average treatment effect* (HBAT) restriction, the firm assumes a constant average treatment effect for all consumers while estimating their profitability based on predicted, heterogeneous baseline purchase probabilities. For the (2) *average baseline purchase probability with heterogeneous treatment effect* (ABHT) restriction, the firm assumes to be targeting homogeneous consumers with equal baseline purchase probabilities while estimating consumers' profitability based on predicted, heterogeneous ad effectiveness. For the (3) *heterogeneous baseline purchase probability with heterogeneous treatment effect* (HBHT) restriction, the firm combines the two restriction strategies estimating based on both predicted, heterogeneous baseline purchase probability and ad effectiveness, which consumers are profitable and should be targeted.

Our data set allows us to assess the effectiveness of the different RCPA approaches for our empirical case. Table 5 compares the results of the currently implemented unrestricted CPA targeting with the results from the different RCPA approaches from our simulations.[27] We standardize all financial variables by dividing them by the standard deviation of the firm's revenue generated from purchases. For fair comparison with the unrestricted CPA targeting, we calculate the overall profit, cost, and return taking the data for all 166,895 consumers that were targeted with the firm's advertising in the field experiment into account. In contrast, we calculate the average profit and cost per consumer based only on the evaluation data set to avoid performance gains from overfitting.

Our results allow us to directly observe empirically whether the restrictions applied by the firm would move the targeting into the desired direction. We see that HBAT, as intended, restricts the ad platform to target consumers with low overall purchase probability, that is, consumers that are cheaper to target. Nonetheless, this RCPA approach is not profitable as consumers targeted under HBAT possess a low average advertising effect. Although the ROI of this strategy is lower than for unrestricted CPA targeting, the firm loses less money in absolute terms. When focusing on targeting consumers that are especially affected by advertising while assuming a homogeneous baseline probability (ABHT), we find that we target consumers that are especially affected by advertising ($ATE = 0.00963$). Given that these consumers possess a similar average overall purchase

Table 5. Overview and Comparison of RCPA Approaches

Targeting	Percent targeted[a]	ATE	Standardized AP	$Pr(y = 1 \mid ad = 1)$	Standardized AC	Standardized profit	Standardized cost	Standardized return	ROI
CPA	100.00%	0.00243	0.00066	0.05596	0.00124	110.88	206.96	−96.09	−46.43%
HBAT	74.17%	0.00008	0.00002	0.00744	0.00016	2.68	20.41	−17.73	−86.88%
ABHT	43.92%	0.00963	0.00263	0.05480	0.00121	192.82	89.01	103.80	116.61%
HBHT	44.59%	0.01066	0.00291	0.04790	0.00106	216.70	78.98	137.72	174.36%

Note. ATE, average treatment effect for consumers targeted with respective targeting strategy; standardized AP, standardized average incremental profit per consumer caused by advertising; standardized AC, standardized average cost per consumer, $Pr(y = 1 \mid ad = 1) \cdot f$.

[a]Percentage of consumers targeted of overall targeted population ($n_{ad=1} = 166,895$).

probability as the consumers in the unrestricted CPA targeting, this approach is profitable. When combining the two restriction approaches in HBHT, we find that we can target consumers that are highly affected by advertising, generating more profit for the firm. Additionally, the consumers targeted under HBHT also have a lower overall purchase probability, being cheaper to target. This makes the combination of restriction approaches the most profitable type of RCPA. HBHT is also in absolute terms the most profitable strategy, as visible in the highest value for the standardized return.

An advantage of RCPA, compared with previously suggested remedies to incentive misalignment in programmatic advertising (Johnson and Lewis 2015, Xu et al. 2016, Lewis and Wong 2018), is that RCPA is implemented by the firm and does not require further intervention from the ad platform. This decreases issues of moral hazard that have been pointed out as significant problem in the relationship between firms, ad platforms, and publishers in related research (Berman 2018). We provide guidelines for firms on how to implement RCPA in Online Appendix N.

8. Discussion

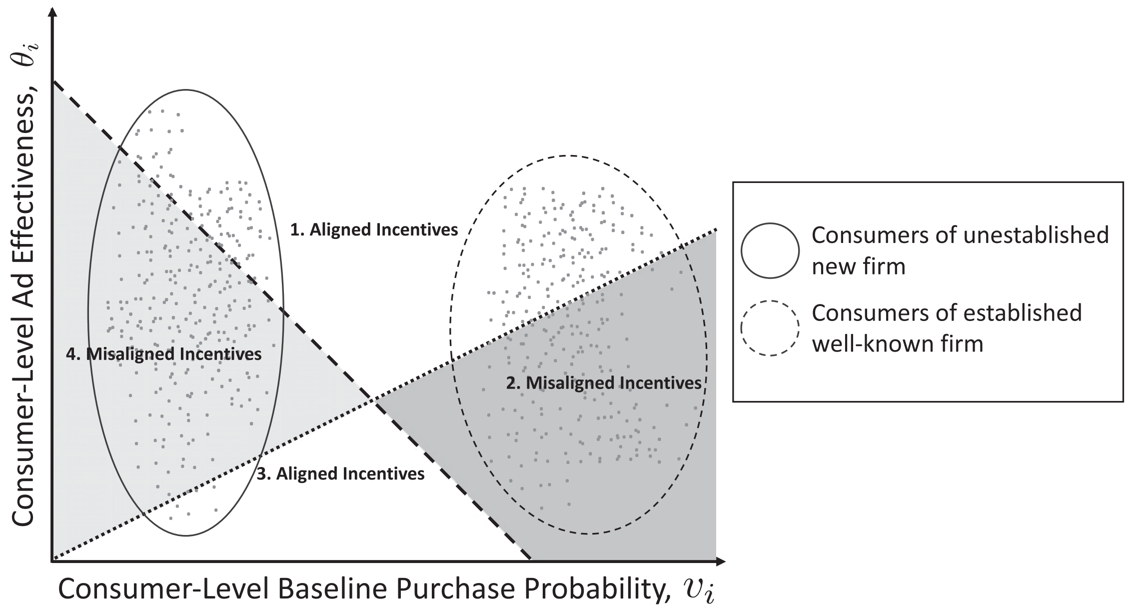
In this work, we provide a detailed description and assessment of the incentives specified in CPA contract designs for firms and ad platforms. In our theoretical model, we see that the presence and severity of an incentive misalignment in the context of CPA contracts is contingent on how consumers are distributed in the two-dimensional space of baseline purchase probability (v_i) and ad effectiveness (θ_i). In addition, our model uncovers aspects of the CPA contract design that have not been considered in previous studies. For example, from the perspective of the firm, common sense would tell us that only ad effectiveness matters when deciding which consumers to target: the firm should be interested in targeting all consumers for which ad effectiveness is above a certain threshold. However, our model reveals that is not the case: the higher the baseline purchase probability (v_i), the higher the ad effectiveness (θ_i) needs to be to make a specific consumer a profitable target for the firm. A higher baseline purchase probability leads to a higher expected fee the firm needs to pay to the ad platform ($v_i f$), even before considering the ad effectiveness. This fee increases with the baseline purchase probability, v_i, and can be thought of as a tax the firm needs to pay the ad platform to target a consumer, irrespective of ad effectiveness. Therefore, from the perspective of the firm, it is only worth paying such a tax if the extra revenue from ad effectiveness ($\theta_i(r-f)$) is large enough to cover this tax.

Ultimately, assessing the incentive misalignment requires an empirical analysis with the aim to pinpoint consumers' position in the (v_i, θ_i) space. Firms face significant empirical challenges in their endeavor to assess the misalignment empirically: the experimental identification of ad effectiveness (Johnson et al. 2017a), limitations in the precision of estimates that require large sample sizes (Lewis and Rao 2015), and the need for symmetric targeting conducted by the ad platform to not introduce selection bias into the treatment group (Gordon et al. 2019). We overcome these challenges in the best possible way and assess the misalignment between firm and ad platform for our empirical case. Our findings suggest that the ad platform is currently able to appropriate profit that is not generated through advertising but would have been generated by the firm independently of addressing consumers with ads.

The reason why the ad platform can appropriate profit from the firm beyond the profit contributed by advertising is related to the discrepancy between overall purchase probability ($v_i + \theta_i$) and the incremental effect of advertising (θ_i). For the treated consumers, the firm observes 9,339 purchases. These purchases generate a profit of 2,549 (measured in the firm's revenue standard deviations). Importantly, the ad platform contributes only 4.35% of this profit generation, whereas the rest of the profit would have been generated without advertising. Nonetheless, the ad platform charges about 8.12% of the generated profit. The gap between overall profit generation from consumers addressed with ads and the actual contribution of profit from advertising allows the ad platform to appropriate profit without the firm—and maybe even the ad platform—directly noticing. In contrast, if the costs for the ad campaign would exceed the overall profit generated from 9,339 purchases, the misalignment would be directly evident to the firm.

How consumers are distributed in the (v_i, θ_i) space is likely to differ contingent on the degree to which a firm is established. Figure 5 showcases the distribution of two exemplary firms: an unestablished new firm and an established well-known firm. The unestablished new firm is likely to face mostly consumers with low baseline purchase probability v_i, as consumers are unaware of the firm's product before seeing an ad. At the same time, for this firm, advertising might prove to be especially effective as it creates an awareness of the firm's existence. The effectiveness of advertising for this firm is likely to be dependent on how consumers actually value the advertised product and how well the product value is communicated in advertising. Such a firm is suffering from a misalignment as the ad platform, based on the contracted incentives, does not target consumers in region 4 that the firm wants to target. In contrast, an established well-known firm is likely to face a lot of consumers with high baseline purchase probability v_i, especially when the firm has a large share of returning customers that value the firm's product. Established firms might not be able to advertise very effectively, as these returning customers are likely to conduct repeat

Figure 5. Example of Targeting Conflicts for Established and Unestablished Firm



purchases independent of advertising. The description of ad effectiveness for such an established firm is in line with findings from related research (Blake et al. 2015). The misalignment for an established firm originates in the ad platform targeting consumers in region 2 that the firm does not want to target. Given the difference in firms and their prospective customers, the experiment campaign conducted in collaboration with our partner firm is limited in its generalizability. Nonetheless, our analyses provide valuable insights and help to clarify the implications of CPA contracts for firms and ad platforms.

Recent studies have put forward solutions to address the issue of how to measure ad effectiveness in a precise and unbiased fashion (Johnson et al. 2017a), how to design contracts to align incentives (Johnson and Lewis 2015), and how to design and implement bidding based on the incremental effect of advertising (Xu et al. 2016, Lewis and Wong 2018). Although we see more ad platforms providing information on ad effectiveness, we observe little progress in the industry to move toward contracts with aligned incentives and ad bidding based on incremental ad effectiveness. Our work provides insights on why the transition to these more efficient solutions is challenging for the involved stakeholders. This transition is slowed down by both technical and empirical challenges but also by the risk for ad platforms to reduce their revenue stream when moving to a return on advertising-based incentive scheme. Moreover, we bring forward a remedy to the inventive misalignment in CPA-based contracts by describing and evaluating a solution to the incentive misalignment that can be implemented on the side of the firm, namely RCPA. We find that if the firm can successfully restrict which consumers the ad platform is targeting, the ROI on ad spend can increase, making advertising more profitable for the firm.

8.1. Limitations

Our work does not come without limitations. In the data set used for our empirical analysis, we only observe bids for ad auctions won by the focal firm. As lower bids are less likely to win an auction for an ad impression, our data could be truncated at the lower end of bids. In theory, such a truncation could keep us from finding that consumers that receive higher bids are more receptive to ads compared with consumers that receive lower bids, as these data might not be in our data set. Nonetheless, we believe this issue is of limited nature in our case. First, the coefficient of variation (ratio of standard deviation to mean) for the bid variable is higher than one (CV = 1.332), indicating sufficient variation in the variable that can be exploited for our analysis. Second, we investigate the effectiveness of advertising for different bids in the context of retargeting. Consumers need to have visited our partner firm's website to be eligible to participate in the experiment. These consumers have higher purchase probabilities compared with consumers that have not visited the partner firm's website. The ad platform should have little interest in bidding systematically low for this type of consumers. Last, when looking at the distribution of bids placed by the ad platform (Figure 2), we see that the bids we observe are clustered relatively close to zero, whereas we observe fewer high bids. This suggests that our data are not severely affected by truncation at the lower end of bids.

Another potential limitation of this paper is that our experimental design does not take the so called "defensive effect" of ads into account. The defensive effect of advertising describes that ads draw effectiveness not only from their direct impact on consumers' purchase probabilities but also from the fact that they prevent competitors from displaying their ads (Johnson et al. 2017a). Using PSAs to technically identify consumers in the control group does not represent the true counterfactual that encompasses the possibility that consumers are confronted with ads of competitors. There are two empirical challenges related to this limitation. First, because we do not observe the defensive effect with our design, we might underestimate the overall impact of ads on consumers' purchase probabilities. This issue seems to have limited impact on our study as we do find evidence that ads have a positive impact on both consumers' purchase and visit probability. Second, in case the defensive effect is not symmetric for different heights of bids, for example, smaller for lower bids and bigger for larger bids, we might not be able to detect a significant increase in ad effectiveness with an increase in the bid. We conduct an additional analysis in which we control for the competition for the opportunity to serve advertising to consumers (see Online Appendix O). When controlling for absolute and relative competition in our analysis, we find results in line with our main analysis. This makes us confident that not being able to consider the defensive effect of ads in our analysis is not affecting the contributions of our work.

Given the difficulty to receive data with the right characteristics to analyze the implications of incentives specified in CPA-based contracts, we can analyze these implications only for our focal firm that conducts a retargeting marketing campaign. The specific implications of CPA-based contracts are likely to differ when investigating them for a different firm and context. We encourage future research, if access to such data can be achieved, to investigate the presence and degree of an incentive misalignment for other firms and contexts.

9. Contributions and Managerial Implications

This study contributes to research in the area of economics of advertising. We investigate the implications of commonly used incentive schemes in programmatic advertising contracts. In our stylized model, we show that some consumers are profitable to target from the ad platform's point of view but not profitable for the firm. In practice, the magnitude of this potential incentive misalignment depends on both the actual targeting (i.e., bidding) behavior of the ad platform and the distribution of consumers in the two-dimensional space of baseline purchase probability (v_i) and ad effectiveness (θ_i). Our field experiment allows us to identify the causal impact of digital advertising on consumers' purchase probabilities while simultaneously exploiting the variation introduced in the bids for ad impressions by the ad platform. We find evidence for the presence of an incentive misalignment. Although ads do generally increase consumers' purchase probabilities, the ad platform targets consumers with higher baseline purchase probability and not those that are more receptive to ads. This finding renders the ad allocation process suboptimal for the firm. A welfare analysis for our empirical case provides further insights on the efficiency loss of the commonly implemented CPA-contract design. This work is the first to provide actual empirical evidence for the presence of an incentive misalignment between firms and ad platforms in programmatic advertising. With new advances in machine learning technologies, ad platforms will likely continue to improve their capability to identify consumers with high baseline purchase probability. Although this improvement helps ad platforms to generate more revenue under the currently contracted incentives, this will likely contribute to a stronger incentive misalignment between firm and ad platform.

Beyond a clearer description and assessment of the presented incentive misalignment, our work contributes to the literature by clarifying that the incentive misalignment is not caused by a lack of information on the ad platform's side. Although it is very difficult for ad platforms to target consumers based on ad effectiveness, our model shows that the incentive misalignment is contingent on the distribution of consumers in the two-dimensional space of baseline purchase probability (v_i) and ad effectiveness (θ_i). This is the case independent of the ad platform's information about the distribution of consumers in such a space, that is, even if the ad platform has perfect information and can predict consumers' ad effectiveness. Thus, the difficulty to predict ad effectiveness is not the reason for the presence of a misalignment. In a contract design with aligned incentives, if the ad platform is not able to detect a difference in ad effectiveness for different consumers, it should target all consumers equally, that is, bid the same for the opportunity to serve ads to consumers.

To remedy the incentive misalignment between firm and ad platform, we propose a novel solution to the incentive misalignment in CPA contracts, namely RCPA. For this solution, firms restrict ad platforms in which consumers can be targeted by the ad platform, based on the profitability to target these consumers for the firm. The advantage of this solution is, that in contrast to previously suggested solutions (Johnson and Lewis 2015, Xu et al. 2016, Lewis and Wong 2018), our solution can be implemented on the firm side without intervention from the ad platform. Our empirical evaluation of the solution shows that firms can drastically improve their ROI on ad spend with RCPA.

Our research clearly points to the fact the firms need better data and better contracts to align their incentives with the incentives of ad platforms. Incentivizing ad platforms based on the number of absolute purchases does not lead to serving effective advertising.

Acknowledgments

The authors would like to thank the editors; three anonymous referees; the participants of the NBER Summer Institute 2017, Economics of Information Technology and Digitization Workshop, Boston; the participants of the Workshop on Information Systems and Economics 2017 (WISE), Seoul, South Korea; and participants of the Economics of Information and Communication Technologies (ZEW) Conference 2018, Centre for European Economics Research, Mannheim, Germany, for their valuable suggestions.

Endnotes

[1] *CPM* stands for cost per mille (a thousand impressions), *CPC* for cost per click, and *CPA* for cost-per-acquisition and cost-per-action, which are used interchangeably in the ad industry. This evolution of ad contract structures was possible because of technological developments, most notably cookie technologies, that allow ad platforms to track clicks and target actions on firms' websites.

[2] Although this approach has been criticized for being expensive, making firms spend money on ads for consumers in the control group that do not contribute to the firm's success, it is still a reliable option to assess the causal effect of ads (Johnson et al. 2017b).

[3] Hinging on advertisers knowing their true valuation of an ad impression.

[4] In our model, f is assumed fixed in the short run, that is, exogenous. Nonetheless, in the long run, the ad platform could adjust this fee, leading to potentially interesting dynamics between the ad platform and the firm. Those analyses are out of the scope of this paper.

[5] This approximation is valid if the purchase probabilities are relatively small and close to each other as it seems to be the case. In Online Appendix A, we present our model estimating the purchase probability with a logistic link function. The results and insights are qualitatively similar.

[6] For simplicity, we assume this profit, r, to be the same for every purchase instance.

[7] Related research has referred to contracts that incentivize ad platforms to target consumers based on ad effectiveness as cost per incremental action (CPIA) contracts (Johnson and Lewis 2015). In Online Appendix B, we present a more detailed theoretical description of how CPIA contracts would solve the incentive misalignment.

[8] This simplification seems plausible as the firm chooses one f for all purchases reported by the ad platform. In our notation, f does therefore represent the denominator in a linear transformation.

[9] We assume that the ad platform can consistently estimate the values of β_2 and β_3.

[10] There is an additional reason why the ad platform's estimates of ad effectiveness may be attenuated. The ad platform optimizes the bid for all consumers, in both treatment and control group, at the same time, that is, the ad platform is "blind" to which consumer is in the treatment or control group. This joint optimization can bias the ad platform's (post hoc) estimate of the ad effectiveness and consequently lead to a lower bid. However, the estimate that gets biased is the ad platform's estimate and not the researchers' estimate. From Equation (8), this means that $\hat{\beta}_2$ and $\hat{\beta}_3$ are not consistent. More specifically, these estimates are attenuated by a factor γ corresponding to the proportion of consumers served the actual ad ($\gamma = 0.8$ in our case): $\hat{\beta}_2 \xrightarrow{p} \gamma\beta_2$; $\hat{\beta}_3 \xrightarrow{p} \gamma\beta_3$. As γ approaches zero, we get closer to our ideal situation in which we would be estimating Equation (9).

[11] Online Appendix C shows that the allocation to treatment and control group does indeed follow a 80%–20% split, independent of the bid, as expected.

[12] We follow this approach of serving PSAs to the control group for both technical and methodological reasons. Technically, we need to place a cookie on consumers' devices in the control group to identify them as consumers in the control group. Importantly, just identifying consumers by the cookie placed on their computers as consumers allocated to the control group is insufficient. As for the treatment group, not all consumers assigned to the intent-to-treat group are treated. This can be the case because some consumers have no opportunity to display an ad or the ad auction is not won and no ad is displayed. To ensure that our control group is comparable to the treatment group (and not representing an intent-to-control group), the ad platform needs to win impressions for these consumers in auctions and record these consumers and their subsequent purchase decisions in the system. This way, we avoid introducing a bias in our estimates by comparing our treatment with an intent-to-control group. The treatment with the PSAs essentially represents a placebo treatment.

[13] In all our analyses in the main section of this work, we use the first bid of the ad platform per consumer as an operationalization of the bid variable. We do this to overcome potential endogeneity arising from within-consumer learning throughout the experiment by the ad platform.

[14] Whereas our data includes only winning bids, this is not a problem in the context of our paper. First, we focus on a retargeting campaign—that is, all the relevant consumers have visited the firm's website within 14 days before the first bid. In essence, these consumers are more attractive to be served ads from the focal firm than from a competing campaign, decreasing the concern that there is a sizeable number of missing observations because of systematically low bids. Second, the observed first bids vary over a wide range of values, which provides enough variation with respect to how much the platform values consumers (we can assess the value of advertising over a wide range of values for the bid). These two factors taken together ensure us we can properly estimate our parameters of interest—average ad effectiveness and the respective relationship between ad effectiveness and consumer baseline purchase probability—for the population of consumers that would win the bids.

[15] An F test on the direct effect of advertising (β_2) and the interaction effect (β_3) for our reduced-form model (column 2; $F(2, 208534) = 3.41$; $p = 0.033$) indicates that the joint significance of including the direct and conditional effect of the advertising treatment is similar to the significance of including the main effect of the advertising treatment (column 1), with a p value of 0.051. This confirms the expectation that the aggregate effect of ad is statistically significant.

[16] Our estimates' precision is not qualitatively affected by adding covariates controlling for ad characteristics, that is, format and placement, and time controls, that is, hour of the day, day of the month, and day of the week (see Online Appendix F).

[17] Despite the coefficient of *bid* being endogenous, its interaction with an exogenous variable—in this case, *ad*—is identified by OLS under mild assumptions and can be interpreted causally (Bun and Harrison 2019).

[18] According to our model (Equation (8)), the baseline purchase probability is a linear function of the ad platform's first bid. The 95% confidence intervals of the difference between treatment and control are calculated using the formula for the difference between two independent normal variables: $\sigma_d = \sqrt{\sigma_t^2 + \sigma_c^2}$.

[19] We perform a sensitivity analysis for the opportunity cost q by varying it between the minimum and the maximum observed values for the cost of the first impression in our data. We find that the estimated welfare loss does not change substantially with this parameter (see Online Appendix M).

[20] We perform a sensitivity analysis by varying this parameter between zero and 0.04, that is, four times the standard deviation observed for the baseline purchase probability. For very low levels of dispersion of ad effectiveness, the ad platform regime is beneficial to the ad platform, at the expense of the firm's profit but with a smaller loss on total welfare. (see Online Appendix M).

[21] As a result of the random draws for θ, some consumers end up having a total purchase probability outside the interval $[0,1]$ ($v_i + \theta_i \notin [0,1]$). We adjust the total purchase probability in these cases so that remains within the $[0,1]$ interval. Alternative approaches—such as removing these consumers or keeping their total purchase probability outside the $[0, 1]$ interval—result in qualitatively similar outcomes.

[22] We consider only the profits and costs for the 166,895 treated consumers. Although targeting control consumers with PSAs also had costs, these correspond to the costs of implementing the randomized experiment. For completeness—and under the assumption that PSAs have no effect on the consumers' purchase likelihood—targeting 20% of the consumers with PSAs cost the firm 1.19 standard deviations per 1,000 consumers, or a total of 49.40 standard deviations of the firm's revenue.

[23] The firm is technically able to exclude an individual consumer from being targeted by setting a flag that is stored in a tracking database and linked to a consumer identifier, that is, cookie.

[24] Although this is not the ideal targeting strategy from the firm's perspective—the firm would want to target consumers in both regions 1 and 4—it is a compromise in which only consumers that are valuable for both ad platform and firm are targeted. This does not mean this is a suboptimal solution. To determine whether a solution is optimal, we need to identify the efficient outcome. The reason why consumers in region 4 are not targeted with the ad from the focal firm, is because the ad platform prefers to serve other ads to these consumers. This means that these consumers are likely being served the ad they value the most, leading to an efficient outcome.

[25] For simplicity we assume a constant profit contribution, r, for all consumers. This is for example the case for a firm that sells a single product. For firms with a heterogeneous product portfolio, the average profit contribution can serve a sensible proxy.

[26] Online Appendix I presents details on the machine learning models used for RCPA.

[27] These economic implications of the different RCPA strategies will differ from empirical case to empirical case as they depend on consumers distribution in the v_i-θ_i-space.

References

Adomavicius G, Curley SP, Gupta A, Sanyal P (2012) Effect of information feedback on bidder behavior in continuous combinatorial auctions. *Management Sci.* 58(4):811–830.

Agarwal A, Mukhopadhyay T (2016) The impact of competing ads on click performance in sponsored search. *Inform. Systems Res.* 27(3):538–557.

Arnosti N, Beck M, Milgrom P (2016) Adverse selection and auction design for internet display advertising. *Amer. Econom. Rev.* 106(10):2852–2866.

Arora N, Dreze X, Ghose A, Hess JD, Iyengar R, Jing B, Joshi Y, et al (2008) Putting one-to-one marketing to work: Personalization, customization, and choice. *Marketing Lett.* 19(3-4):305–321.

Asdemir K, Kumar N, Jacob VS (2012) Pricing models for online advertising: CPM vs. CPC. *Inform. Systems Res.* 23(3):804–822.

Athey S, Imbens G (2017) *The Econometrics of Randomized Experiments*, vol. 1 (Elsevier, North Holland), 73–140.

Balseiro SR, Feldman J, Mirrokni V, Muthukrishnan S (2014) Yield optimization of display advertising with ad exchange. *Management Sci.* 60(12):2886–2907.

Berman R (2018) Beyond the last touch: Attribution in online advertising. *Marketing Sci.* 37(5):771–792.

Blake T, Nosko C, Tadelis S (2015) Consumer heterogeneity and paid search effectiveness: A large-scale field experiment. *Econometrica* 83(1):155–174.

Bleier A, Eisenbeiss M (2015) Personalized online advertising effectiveness: The interplay of what, when, and where. *Marketing Sci.* 34(5):669–688.

Bun MJ, Harrison TD (2019) OLS and IV estimation of regression models including endogenous interaction terms. *Econometric Rev.* 38(7):814–827.

Choi H, Mela CF, Balseiro SR, Leary A (2020) Online display advertising markets: A literature review and future directions. *Inform. Systems Res.* 31(2):556–575.

Clarke EH (1971) Multipart pricing of public goods. *Public Choice.* 11(1):17–33.

Dalessandro B, Perlich C, Stitelman O, Provost F (2012) Causally motivated attribution for online advertising. *Proc. 6th Internat. Workshop on Data Mining for Online Advertising and Internet Econom.* (ACM, New York), 1–9.

Edelman BB, Ostrovsky M, Schwarz M (2007) Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *Amer. Econom. Rev.* 97(1):242–259.

Feng J, Xie J (2012) Research note: Performance-based advertising: Advertising as signals of product quality. *Inform. Systems Res.* 23(3):1030–1041.

Gordon B, Zettelmeyer F, Bhargava N, Chapsky D (2019) A comparison of approaches to advertising measurement: Evidence from big field experiments at Facebook. *Marketing Sci.* 38(2): 193–225.

Gordon BR, Jerath K, Katona Z, Narayanan S, Shin J, Wilbur KC (2021) Inefficiencies in digital advertising markets. *J. Marketing* 85(1):7–25.

Groves T (1973) Incentives in teams. *Econometrica* 41(4):617–631.

Hoban PR, Bucklin RE (2015) Effects of internet display advertising in the purchase funnel: Model-based insights from a randomized field experiment. *J. Marketing Res.* 52(3):375–393.

Hu YJ, Shin J, Tang Z (2016) Incentive problems in performance-based online advertising pricing: Cost per click vs. cost per action. *Management Sci.* 62(7):2022–2038.

Johnson GA, Lewis RA (2015) Cost per incremental action: Efficient pricing of advertising. Working paper, Simon Business School Working Paper No. FR 15-29, University of Rochester, Rochester, NY.

Johnson GA, Lewis RA, Nubbemeyer EI (2017a) Ghost ads: Improving the economics of measuring ad effectiveness. *J. Marketing Res.* 54(6):867–884.

Johnson G, Lewis RA, Reiley D (2016) Location, location, location: Repetition and proximity increase advertising effectiveness. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2268215.

Johnson GA, Lewis RA, Reiley DH (2017b) When less is more: Data and power in advertising experiments. *Marketing Sci.* 36(1):43–53.

Lambrecht A, Tucker C (2013) When does retargeting work? Information specificity in online advertising. *J. Marketing Res.* 50(5): 561–576.

Lewis RA, Rao JM (2015) The unfavorable economics of measuring the returns to advertising. *Quart. J. Econom.* 130(4):1941–1973.

Lewis RA, Reiley DH (2014a) Advertising effectively influences older users: How field experiments can improve measurement and targeting. *Rev. Industrial Organ.* 44(2):147–159.

Lewis RA, Reiley DH (2014b) Online ads and offline sales: Measuring the effect of retail advertising via a controlled experiment on Yahoo! *Quant. Marketing Econom.* 12(3):235–266.

Lewis RA, Wong J (2018) Incrementality bidding & attribution. https://dx.doi.org/10.2139/ssrn.3129350.

Lewis R, Rao JM, Reiley DH (2015) *Economic Analysis of the Digital Economy* (University of Chicago Press, Chicago).

Lu Y, Gupta A, Ketter W, Van Heck E (2016) Exploring bidder heterogeneity in multichannel sequential b2b auctions. *Management Inform. Systems Quart.* 40(3):645–662.

Lu Y, Gupta A, Ketter W, Van Heck E (2019) Information transparency in business-to-business auction markets: The role of winner identity disclosure. *Management Sci.* 65(9):4261–4279.

Sahni NS (2015) Effect of temporal spacing between advertising exposures: Evidence from online field experiments. *Quant. Marketing Econom.* 13(3):203–247.

Sahni NS, Narayanan S, Kalyanam K (2019) An experimental investigation of the effects of retargeted advertising: The role of frequency and timing. *J. Marketing Res.* 56(3):401–418.

Schumann J (2014) Targeted online advertising: Using reciprocity appeals to increase acceptance among users of free web services. *J. Marketing* 78(1):59–75.

Schwartz EM, Bradlow ET, Fader PS (2017) Customer acquisition via display advertising using multi-armed bandit experiments. *Marketing Sci.* 36(4):500–522.

Simester D, Timoshenko A, Zoumpoulis SI (2020a) Efficiently evaluating targeting policies: Improving on champion vs. challenger experiments. *Management Sci.* 66(8):3412–3424.

Simester D, Timoshenko A, Zoumpoulis SI (2020b) Targeting prospective customers: Robustness of machine-learning methods to typical data challenges. *Management Sci.* 66(6):2495–2522.

Vickrey W (1961) Counterspeculation, auctions, and competitive sealed tenders. *J. Finance* 16(1):8–37.

Xu J, Shao X, Ma J, Lee KC, Qi H, Lu Q (2016) Lift-based bidding in ad selection. *Proc. 30th Conf. on Artificial Intelligence* (Association for the Advancement of Artificial Intelligence, Palo Alto, CA), 651–657.

Zantedeschi D, Feit EM, Bradlow ET (2017) Measuring multichannel advertising response. *Management Sci.* 63(8):2706–2728.