Beyond the Last Touch: Attribution in Online Advertising

Ron Berman

Please scroll down for article—it is on subsequent pages

With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

Beyond the Last Touch: Attribution in Online Advertising

Ron Berman[a]

[a] Marketing Department, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104
Contact: ronber@wharton.upenn.edu, http://orcid.org/0000-0002-8594-3627 (RB)

Abstract. Online advertisers often utilize multiple publishers to deliver ads to multi-homing consumers. These ads often generate externalities and their exposure is uncertain, impacting advertising effectiveness across publishers. We analyze the inefficiencies created by externalities and uncertainty when information is symmetric between advertisers and publishers, in contrast to most previous research that assumes information asymmetry. Although these inefficiencies cannot be resolved through publisher-side actions, attribution methods that measure campaign uncertainty can serve as alternative solutions to help advertisers adjust their strategies. Attribution creates a virtual competition between publishers, resulting in a team compensation problem. The equilibrium may potentially increase the aggressiveness of advertiser bidding, leading to increased advertiser profits. The popular last-touch method is shown to overincentivize ad exposures, often resulting in lower advertiser profits. The Shapley value achieves an increase in profits compared with the last-touch method. Popular publishers and those that appear early in the conversion funnel benefit the most from advertisers using last-touch attribution. The increase in advertiser profits comes at the expense of total publisher profits and often results in decreased ad allocation efficiency. We also find that the prices paid in the market will decrease when more sophisticated attribution methods are adopted.

1. Introduction

Advertisers in online campaigns often choose to deliver ads using multiple publishers that reach overlapping target populations. The ads displayed on Google's and Facebook's websites, for example, cover more than 50% of the online advertising market in the United States.[1] When consumers multihome and visit multiple publishers during these campaigns, the ads they see often have a cumulative effect that creates an externality between the effectiveness of ads across publishers.

Because the behavior of multihoming consumers is stochastic, when planning multipublisher campaigns, advertisers face uncertainty about the campaigns' performance. This uncertainty may be about the effectiveness of individual ad exposures, the reach the ads may have, or the effect from competition that targets the same customers. When buying these ads in an auction market, the advertiser needs to take these uncertainties into account and ex ante decide how to bid on ads, resulting in ad allocations and campaign performance that are less efficient than if the advertiser had complete information. These inefficiencies often result in severe degradation of the effectiveness of advertising campaigns, leading to frustration from advertisers and questions about the effectiveness of online advertising.

Procter & Gamble, for example, recently lowered their digital ad spending because they could not prove its effectiveness (Bruell and Terlep 2017).

The literature on online advertising markets has focused primarily on the inefficiency of advertising campaigns created by advertising channel discoordination and asymmetric information. This prior research takes the viewpoint of platforms and publishers, focusing on the actions they can take in the market. A primary assumption is that by aggregating their information, publishers can gain superior information compared with advertisers and can exploit this advantage to coordinate the market and improve its efficiency. Examples include changing the pricing mechanisms in the market to create incentive compatibility (Jordan et al. 2011, Johnson and Lewis 2015) and aggregating publishers into a network that can track and target individuals across different publishers (D'Annunzio and Russo 2017). These assumptions, however, can be limited in their applicability. One limitation is the fact that competing platforms such as Google and Facebook will be reluctant to share and coordinate information or may be forbidden to do so. A second limitation is that advertisers are left having to trust the publishers to tell them how effective their ads are in a market rife with moral hazard.

In this paper, we take the viewpoint of an advertiser and analyze the campaign inefficiency that arises when an advertiser buys ads on multiple publishers in a market with *structural uncertainties*. We call an uncertainty "structural" when the information is symmetric (i.e., all players have the same a priori knowledge about ad efficiency) and when the publishers cannot resolve it fully unless they change the market structure. Specifically, because consumer visits to websites are stochastic and because bids need to be submitted in advance and independently to each publisher, publishers are unable to discriminate between consumers who multihome and those who do not. Unless publishers collude and share information in real time or agree on a centralized pricing mechanism, they will be unable to exploit targeted pricing or allocation for multihoming consumers.

Unlike publishers, however, advertisers do have access to the complete information collected during a campaign, ex post. Because advertisers will not coordinate or cooperate with their competitors, their problem is viewed as a problem of measurement (of the uncertain state of the campaign) and optimization given this measurement. The assumption is that once proper measurement is in place, optimization of the advertising campaigns will improve the competitiveness of the advertisers, increase their profits, and make the market more efficient. Consequently, a common technique that advertisers use with these data is *multitouch attribution*, which is the focus of this paper. Multitouch attribution uses the information collected during the campaign to estimate the effectiveness of each publisher and the contribution of each ad exposure ("touch") to each consumer response. These estimates are then used to adjust the bids on different publishers, with a goal of optimizing the campaign.

Although multitouch attribution is commonly used to benchmark publisher performance (Econsultancy 2012), it is unclear what the value of the process is and how to best make use of the data. In a recent survey by Econsultancy (2012, p. 13) and Google, 42% indicated that being "unsure of how to choose the appropriate method/model of attribution" is an impediment to adopting an attribution method. The most common method used by advertisers is called the "last-touch" (LT) method, in which the last touchpoint in a sequence of ads prior to conversion gets full credit for the conversion. Another recently popular method is using the Shapley value (Sh) to assign the average marginal contribution to conversion as the credit for each touchpoint. Because these methods are in essence measurement methods of the uncertain state of the campaign, the result of the measurement process will impact the bidding strategies of the attributing advertiser.

The first research question we address is to understand how structural uncertainties impact the profits and behavior of advertisers when they run multipublisher campaigns. Because externalities exist between publishers, the bidding strategies of advertisers, the equilibrium bids by competitors, and allocation of ads by publishers may be different from when running single-publisher campaigns. Our model assumes that information is symmetric in the market, unlike most other research, and assesses the impact of uncertainty when neither publishers nor advertisers have an informational advantage. Although a significant body of research looks to understand the influences of asymmetric information, risk preferences, and monitoring capabilities online (Agarwal et al. 2009, Wilbur and Zhu 2009, Mayzlin et al. 2014, Hu et al. 2015), we show that even without such asymmetries, the market will still suffer from inefficiency, and attribution may help as a tool. Because we assume that publishers cannot coordinate through a centralized mechanism, and given the impact of structural uncertainty, we focus on the potential of multitouch attribution to increase advertiser profits. Specifically, we ask under what circumstances an advertiser should expect a benefit from applying an attribution method, and what the impact may be for competing advertisers and publishers. Answering these questions helps to explain why advertisers may be unhappy with the effectiveness of their campaigns once they start engaging in multipublisher advertising. As our results show, engaging in a multipublisher campaign changes the allocation of profits between advertisers and publishers, which also helps to explain why many publishers support the efforts of advertisers to introduce attribution solutions.

The second question we tackle touches on the effectiveness of two popular attribution methods (last touch and the Shapley value) when an advertiser competes with other strategic advertisers. As we will discuss in Section 2, the majority of research on advertising attribution is empirical in nature and focuses on developing more accurate measurement methods. A common issue with this analysis, however, is the assumption that ad prices and competitor responses will not change when an advertiser gains more information. These methods are benchmarked in a vacuum, without taking into account how the market will respond to an advertiser who modifies her bidding strategy. Analytical work on the topic also assumes a single advertiser that attempts to optimize her campaigns, disregarding the effects and responses of publishers and other advertisers. Our analysis, in contrast, shows that the indirect equilibrium effects of applying an attribution method are substantial. For example, we find cases where using multitouch attribution results in *lowered* profits for advertisers because of the competitive positioning it creates for the advertiser. The analysis in this case helps to shed light on and derive insights about the trade-offs that an advertiser faces when she deploys different attribution

methods. As we will show, although some methods are more accurate as measurement methods, they may result in undesired equilibrium consequences for advertisers, publishers, or the market as a whole.

The analysis we perform assists in gauging the potential benefit of improved market efficiency from the deployment of attribution methods under the current advertising market structure, in contrast to the approaches that call for structural market changes by platforms. When analyzing advertising market efficiencies, the prior literature has often focused on the potential of a method to increase ad allocation efficiency, with the goal of increasing social welfare. Such increase in welfare and allocation efficiency has often come at the expense of the profits of advertisers, through increased rent on the publisher side. Our focus on the division of profits between the advertisers and publishers contributes to this literature by emphasizing the trade-off between allocation efficiency and advertiser profits. Because advertisers today (and especially the large ones) are questioning the effectiveness of their campaigns, the results from our research sheds light on how the implementation of attribution methods may impact the incentives of advertisers to stay in the market. A second focus of the prior research on method optimality tries to asses which attribution method is "best," in the sense that it recovers the true uncertain state of the market. As we will show, a more accurate measurement method may not always be more profitable for an advertiser because of the indirect equilibrium effects it has. We thus question the assumption that more accurate measurement methods that lead to increased competition will increase the advertiser's profits or the ad allocation efficiency in the market. Our results show that there exists a complex interaction between ad allocation efficiency, measurement accuracy, and advertiser and publisher profits.

In Section 3, we introduce our two-period game theoretical model. The model consists of consumers, two publishers, and three competing advertisers engaged in an advertising campaign. The consumers are homogeneous and visit each publisher probabilistically, in which case they are exposed to ads. Because the visit patterns by consumers are stochastic, advertisers and publishers have uncertainty about realized ad exposures and their effects. When consumers are exposed to ads, they may take an action we call "conversion" (e.g., clicking an ad, pressing the "like" button, watching a video), which generates value for the advertiser. Because the consumer may see multiple ads, the returns to additional ads may be increasing or decreasing. We focus primarily on decreasing ad returns, but mention the impact of increasing returns where appropriate. All proofs in the paper appear in the appendix.

Advertisers bid in ad auctions and try to maximize their profits. One of the advertisers is a *global* advertiser, who runs the campaign on both publishers, and the other two advertisers are *local* advertisers, each running a campaign on one publisher only.[2]

The ad allocation mechanism we consider is a standard second-price auction with impression pricing (CPM).[3] Although we use the term "publishers," we have in mind large publishers such as Google or Facebook. The publishers in our model do not take the role of an ad network that allocates ads among different websites, which is why we interchangeably use the terms "publisher" and "platform."

Section 4 contains a benchmark analysis that compares a global advertiser who does not use attribution to an optimal advertiser who has complete market information. We find that when not using attribution, both the global and the local advertisers will benefit compared with the case when all advertisers are local, because of reduced competition resulting from ad externalities. In this case, the profits of popular publishers decrease, giving a plausible explanation for why large publishers (such as Google) provide free attribution solutions to large advertisers to increase their competitiveness. We also find that when advertisers value conversions highly, gaining full market information does not increase profits and is useless. This result shows that advertisers with high-value consumers should not expect to benefit from implementing attribution.

Section 5 looks at the impact of deploying an attribution solution for the global advertiser. We show that the attribution problem is an instance of the team problem identified by Holmstrom (1982), implying that a suboptimal number of ads will be purchased in the market and that attribution will never achieve the optimal allocation an advertiser could achieve with complete information. We then analyze two common attribution schemes: last-touch attribution and Shapley value attribution based on Shapley (1952).

In equilibrium, we find that last-touch attribution incentivizes an increase in ad exposures. This increase results in overexposure beyond the optimal level of advertising, limiting the benefit for the global advertiser. We find that there is no condition in which last-touch attribution yields higher profits for the global advertiser compared with both the Shapley value and to not using attribution. The Shapley value, however, will increase ad exposure at the unpopular publisher, but may decrease ad exposure at the popular publisher, resulting in increased profits compared with last-touch attribution when one of the publishers is unpopular. Counterintuitively, we also find that last-touch attribution benefits publishers that come earlier in the conversion funnel of a consumer. The intuition behind these results stems from the fact that last-touch attribution is extreme in allocating the full effect of the externality to the second publisher in the chain of visits. The Shapley value, in contrast, is order agnostic and generally does not overincentivize exposure at both publishers.

These properties raise two questions about the optimality of these attribution methods. Section 5.2 looks at the ad allocation efficiency in a market with attribution. As attribution is geared to increase advertiser profits, this increase may result in decreased allocation efficiency. This is because resolving uncertainties does not necessarily resolve externalities absent a centralized price coordination mechanism. We surprisingly find that last-touch attribution may increase efficiency beyond what is achievable when an advertiser has full market information. This is not the case for the Shapley value, however, which decreases market efficiency. Section 5.3 tries to determine how far each method is from reaching the maximal potential profit for an advertiser. We find that all methods perform well in the worst-case scenario, yielding more than 85% of potential profit to the advertiser. Altogether, these results show a trade-off between attribution methods that improve the overall market allocation efficiency and methods that allow individual advertisers to increase their profits.

Section 5.4 looks further into the question of the lingering stickiness and impact of last-touch attribution through a comparative statics analysis. When we fix the attribution method, changing the market parameters or changing ad effectiveness does not have an equal effect on the global advertiser's profit. If the advertiser can increase the effectiveness of her ad copy to elicit earlier conversions in the funnel, using last-touch attribution will increase the profit of the advertiser, whereas the Shapley value and no-attribution (NA) methods will decrease the advertiser's profit. This stability may be another reason for the popularity of last-touch attribution. We also find, in Section 5.5, conditions where the Shapley value not only decreases the profit of the global advertiser, but also decreases the revenue of both publishers, yielding a lose–lose–lose scenario. In contrast, last-touch attribution increases the profits of the global advertiser and the publishers. When we focus on ad prices, we find that the global advertiser will pay higher ad prices when using last-touch attribution and lower ad prices when using the Shapley value. This may partially explain the pricing trend observed in advertising markets today.

The discussion in Section 6 concludes by analyzing the limitations of the findings and implications for theory and practice. Our analysis shows that the effects of attribution are not trivial and that the benefits may be limited in many cases. In general, attribution will not yield the optimal (complete information) profit for an advertiser because of the team problem it creates. Moreover, there are cases when even gaining full information would not change the actions of the advertiser, yielding little gain from attribution. Analyzing these limitations has allowed us to give further guidance and understanding to advertisers considering the development and deployment of an attribution technology. Despite these limitations, we find that under a wide range of conditions, attribution does improve the profits of a global advertiser, may increase the ad allocation efficiency in the market, and, in certain cases, also increases the profits of publishers while at the same time increasing the profit of global advertisers. Thus, attribution is a viable advertiser-side alternative to price adjustment mechanisms or using information sharing on the publisher side.

2. Related Work

Recent work employing large-scale field experiments by Lambrecht and Tucker (2013) on retargeting advertising, Blake et al. (2015) on search advertising, and Lewis and Rao (2015) on banner advertising found little evidence for the effectiveness of these campaigns when measured over a broad population. The main finding of these works is that the effects of advertising are moderate at best and require large sample sizes to properly identify.

The work on advertising ad auctions has generally focused on designing auctions with better equilibrium properties (e.g., Kempe and Mahdian 2008) or analyzing the impact of externalities from other advertisers when they appear on the same page. In these cases, the auctioneer can change their auction mechanism to adjust the efficiency of the ad allocation and correct for the externalities. We contribute to this research by analyzing the second-price auction when externalities exist among publishers, but not on the same publisher, preventing them from correcting for the externality.

In the past few years, the advertising industry has shown increased interest in improved attribution methods. In a recent survey by Econsultancy (2012, pp. 6, 13) and Google, 54% of advertisers indicated they used a last-touch method, whereas 42% indicated that being "unsure of how to choose the appropriate method/model of attribution" is an impediment to adopting an attribution method. Research focusing on the advertiser's problem of measuring and compensating multiple publishers is quite recent, however, with the majority focusing on empirical applications to specific campaign formats. Tucker (2012) and Li et al. (2016) analyze the impact of better attribution technology on campaign decisions by advertisers. The papers find that improved attribution technology lowered the cost per attributed converter or increased the return on investment (ROI) of the campaign.

Shao and Li (2011), Abhishek et al. (2012), Li and Kannan (2014), Anderl et al. (2016), and Kireyev et al. (2016) build specific attribution models for online campaign data using a conversion model of consumers and interaction between publishers. They find that publishers have strong interaction effects between one another that are typically not picked up by traditional measurements. In addition, Dalessandro et al. (2012) shows through simulation that using the Shapley value

can approximate the causal effects of different channels. Finally, Zantedeschi et al. (2017) and Barajas et al. (2016) show how challenging the attribution problem is even when collecting data using an experiment. The latter paper also finds that impressions with a high baseline of conversions will be allocated by a publisher to an advertiser even though this allocation may be inefficient. Our contribution to this literature is the analysis of the competitive equilibrium, often ignored by most empirical work. We are able, for example, to explain changes in ad prices and take them into account when analyzing the benefits of different schemes.

Theoretical analyses of the attribution problem are not common. The work by Jordan et al. (2011) analyzes a problem similar to ours, but has a goal of changing the auction mechanism to eliminate the need for attribution. In their model, the authors look at designing a bidding and pricing mechanism that is efficient and maximizes social welfare when there are externalities among publishers. In contrast, our work focuses on applying the advertiser's problem to the current common market mechanisms, when advertisers try to maximize their individual profits, and analyzes what conditions in the market call for attribution.

Finally, the results of Agarwal et al. (2009), Zhu and Wilbur (2011), and Hu et al. (2015) consider which payment schemes advertisers and publishers should use when different asymmetries exist in the market. The main conclusions are that an advertiser can take advantage of her private information to select different payment schemes. Our analysis, in contrast, tries to show that a significant challenge still faces the advertiser when she is limited to a specific payment scheme and when the information structure in the market is symmetric. These challenges can sometimes be mitigated by an attribution method.

3. Model of Consumers, Advertisers, and Publishers

Consider a two-period market with three types of players: three advertisers, two publishers, and a continuous mass of homogenous consumers with measure 1. We use the term "publisher" to denote a platform that facilitates the sale of advertising and its delivery. As such, publishers are websites that both sell ads directly and display them to consumers (such as Google's search engine or Facebook). We assume these large publishers will not share information between them because of competitive or regulatory restrictions. The case of platforms that can allocate ads between different publishers in their network in the face of attribution is left as a promising avenue for future work. We denote the two publishers as $j \in \{1, 2\}$.

Two advertisers in the market are *local* advertisers who advertise on only one of the publishers, and one is

global and advertises on both publishers. The global/local structure is common in the combinatorial auctions literature (when bidders can bid on more than one item; cf. Krishna and Rosenthal 1996, Ausubel and Baranov 2013) and allows us to fully characterize the equilibrium behavior of all players in the market. In addition, this assumption allows us to compare the effects of moving from running a single-publisher campaign to a multipublisher campaign. Our primary focus is on analyzing the competition of the global advertiser with the local advertisers, which we label $i \in \{G, 1, 2\}$, respectively. The online appendix analyzes the case of two global advertisers.

In the market, consumers generate revenue to the advertisers through ad exposures on publisher websites, which are bought in ad auctions conducted by each publisher.

3.1. Consumers
Consumers in the model may visit a publisher's website in each time period. If they visit a publisher, they are exposed to ads and may take an action that is valuable to the advertiser, called a conversion.[4]

Visiting publisher websites requires the consumer to decide whether to visit each publisher in each time period and in what order. We assume a consumer can visit each publisher up to once during the two periods, and in period t they visit a publisher with probability p_t when (p_1, p_2) is either (p_H, p_L) or (p_L, p_H) with $p_H > p_L$. This specification means that popular websites (p_H) may come earlier or later in the consumer's visit path. As a result, the world has four possible states, which we denote by $1H$, $1L$, $2H$, and $2L$. The first state, for example, means that the customer visits publisher 1 first with probability p_H and then publisher 2 with probability p_L. We assume these states occur with a priori equal probability.

Modeling the order of visits and decisions to visit in this way allows us to analyze the dynamics emerging from an advertising campaign and to disentangle the potential market power a publisher has from being later (or earlier) in a consumer visit path, or being popular (p_H) or unpopular (p_L) in the market. In addition, modeling order uncertainty and visit uncertainty in such a way captures advertisers' uncertainty about the response of consumers to advertising.

When a consumer visits either of the two publishers, she is exposed to an ad from one of the advertisers on the platform. The process that determines ad exposures is a second-price auction and is described in detail in the following section. Contingent on being exposed to ads, we assume that the consumer makes a conversion decision following every ad exposure. After being exposed to one ad, the consumer converts with probability γ. After being exposed to a second ad, the consumer converts with probability $(1+d)\gamma$; d determines the externality that the first ad may have on the second ad if

it is shown. When $-1 < d < 0$, advertising has positive but decreasing returns, and when $1 > d > 0$, advertising has positive and increasing returns.[5]

From the point of view of an advertiser, exposing a consumer to one ad yields a conversion with probability γ, and exposing a consumer to two ads yields a conversion with probability $\gamma + \gamma(1 + d)$. If, for example, consumers can convert only once, but the second ad will have a similar effect as the first ad on nonconverted consumers, then $d = -\gamma$.

3.2. Publishers

Publishers in the model sell ads directly to advertisers through a second-price auction. In this auction, advertiser i submits a bid b_{ij} to publisher j before the campaign.

If the advertiser wins the ad auction, her ad is being displayed and she is charged the minimal amount that would have allowed her to maintain winning the ad.[6]

We model the advertising process as a second-price pay-per-impression (CPM) auction. In this auction, if $b_{11} > b_{21}$, then advertiser 1 wins the auction and pays b_{21}. We assume that ties in bidding are broken randomly, though in our model they will appear with measure zero. We have chosen to focus on CPM pricing, because it allows analyzing a wide range of parameters in closed form. Moreover, this payment method is popular in the market, and especially within display advertising where attribution is the most relevant. The online appendix shows that the results qualitatively hold under cost-per-click (CPC) pricing as well.

3.3. Advertisers

Each advertiser submits a bid b_{ij} to each publisher, and, following a conversion by a consumer, an advertiser i receives revenue v_i. We assume each advertiser knows its own value for a consumer conversion v_i,[7] but does not know the valuation of the competing advertisers, which are a priori independently distributed, $U[0, 1]$.

A well-known result (Krishna 2009, proposition 2.1) is that in a single-item, second-price auction, it is weakly dominating for an advertiser to bid her value. Following this, publishers such as Facebook state that "There's no advantage to underbidding. In fact, doing so may lead to losing auctions you would've otherwise won at no additional cost" (Facebook 2017, p. 11). One of our first findings shows that this recommendation is incorrect for global advertisers facing externalities.

As mentioned before, advertisers can be either local or global advertisers. Local advertisers show ads only on one platform, and as a result submit only one bid. Moreover, because these advertisers do not compete for advertising on multiple publishers, they do not face an attribution problem—all of their sales will be attributed to the specific platform where they advertise. We will therefore label advertiser 1's bid b_1 and advertiser's 2's bid b_2, and assume that advertiser 1 bids at publisher 1, whereas advertiser 2 bids at publisher 2.

The global advertiser will submit bids b_{G1} and b_{G2} to publishers 1 and 2, respectively. Because the global advertiser does not know in advance the state of the world, she will have to maximize her expected profit when setting the bids. We will analyze how accounting for uncertainty using an attribution process impacts the profits of the global advertiser and analyze the effects of different attribution methods on the market as a whole.

3.4. Timing and Information Structure

The goal of the assumptions in our model is to minimize information asymmetry and preference asymmetry between publishers and advertisers that may drive our results and show that spillovers between advertising channels along with uncertainty about consumer visits are enough to generate substantial market inefficiencies.

We assume the values of p_H, p_L, γ, and d are public knowledge. The values of γ and d can be estimated by advertisers through running a simple experiment and assessing the behavior of consumers exposed to a single ad or a couple of ads. Publishers will have this information from historical data and from observing the experiments run by advertisers. Uncertainty about the conversion rate γ will cause advertisers to use their expected values. The results of this paper follow through in this case.

The state of the world (the order of consumer visits to publishers and whether a publisher is popular or not for this campaign) is unknown to both publishers and advertisers a priori. Each publisher knows whether a consumer visited their own site, but does not know whether a consumer visited or will visit the other publisher. Each publisher also knows the bids of both advertisers on its platform, but does not know the bids at the other publisher.

Each advertiser knows her own valuation for a consumer action v_i, but does not know the valuation of the other advertiser.

The timing of the game is as follows:

1. The global advertiser submits bids to both publishers. Local advertisers submit one bid.

2. Nature determines the order of visits of consumers to publishers and which publisher is popular. Assume, for example, it is first publisher 1 and then 2, and publisher i is visited with probability p_i.

3. In period 1, nature determines whether a consumer visits publisher 1 with probability p_1. If she does, an auction is run; the winning advertiser displays an ad to the consumer and pays the second highest bid for it.

4. If the consumer visits a publisher in period 1 and sees an ad from advertiser i,
  • she will convert with probability γ;
  • the advertiser will receive revenue v_i if the consumer converts.

5. In period 2, nature determines whether a consumer visits publisher 2 with probability p_2. If she does, an auction is run; the winning advertiser displays an ad to the consumer and pays the second highest bid for it.

6. If the consumer visits a publisher in period 2 and sees an ad from advertiser i,
  • she will convert with probability $(1 + d)\gamma$ if she has seen an ad from advertiser i in period 1, and otherwise she will convert with probability γ;
  • the advertiser will receive revenue v_i if the consumer converts.

4. Benchmark Equilibrium Without Attribution

To understand the motivation for the creation of attribution methods, we first assume that the advertisers do not know the state of customer visits, and each maximizes her expected profit.

If a local advertiser i faces a global advertiser that bids b_{Gi} with cumulative distribution function (CDF) $F_{Gi}(\cdot)$ at publisher i, then her expected profit is

$$\pi_i(b_i) = 1/2(p_H + p_L)\int_0^\infty \mathbb{I}[b_i > b_{Gi}](\gamma v_i - b_{Gi})dF_{Gi}(b_{Gi})$$
$$= \frac{p_H + p_L}{2}\int_0^{b_i}(\gamma v_i - b_{Gi})dF_{Gi}(b_{Gi}). \quad (1)$$

Suppose the state of the world is $1H$ (consumers first visit advertiser 1 with probability p_H). If the local advertisers bid according to bid distributions with CDFs F_1 and F_2, then the expected profit in this state of a global advertiser who bids B_{G1} and B_{G2} will be

$$\pi_G(b_{G1}, b_{G2}|1H)$$
$$= p_H(1 - p_L)\int_0^\infty \mathbb{I}[b_{G1} > b_1](\gamma v_G - b_1)dF_1(b_1)$$
$$+ (1 - p_H)p_L\int_0^\infty \mathbb{I}[b_{G2} > b_2](\gamma v_G - b_2)dF_2(b_2)$$
$$+ p_Hp_L\int_0^\infty\int_0^\infty \mathbb{I}[b_{G1} <= b_1]\mathbb{I}[b_{G2} > b_2]$$
$$\cdot (\gamma v_G - b_2)dF_1(b_1)dF_2(b_2)$$
$$+ p_Hp_L\int_0^\infty\int_0^\infty \mathbb{I}[b_{G1} > b_1]\mathbb{I}[b_{G2} <= b_2]$$
$$\cdot (\gamma v_G - b_1)dF_1(b_1)dF_2(b_2)$$
$$+ p_Hp_L\int_0^\infty\int_0^\infty \mathbb{I}[b_{G1} > b_1]\mathbb{I}[b_{G2} > b_2]$$
$$\cdot ((\gamma + \gamma(1 + d))v_G - b_1 - b_2)dF_1(b_1)dF_2(b_2). \quad (2)$$

Simplifying and taking the expectation over the four states, we receive

$$\pi_G(b_{G1}, b_{G2}) = \frac{p_H + p_L}{2}\int_0^{b_{G1}}(\gamma v_G - b_1)dF_1(b_1)$$
$$+ \frac{p_H + p_L}{2}\int_0^{b_{G2}}(\gamma v_G - b_2)dF_2(b_2)$$
$$+ p_Hp_Ld\gamma v_G F_1(b_{G1})F_2(b_{G2}). \quad (3)$$

The last additive shows the effect of the externality and uncertainty about the profit of a global advertiser compared with a local advertiser. Holding all else constant, the profit of the advertiser increases with increasing returns of advertising (positive d) and may decrease with decreasing advertising returns (negative d).

In second-price auctions, there are often multiple equilibria because the equilibrium strategies are weakly dominating. If a bidder bids above the highest bid of the other bidder, they can increase (or decrease) their bid without affecting their or the other bidder's profit or the allocation of the auction items. Therefore, when we report equilibrium bidding strategies, we will report the continuously differentiable strategies, but also indicate when the advertisers can lower their bids from this equilibrium while maintaining the same profit.

An analysis of the full equilibrium without attribution (called *NA*) shows the following results about the equilibrium bids of the global advertiser.

Lemma 1. *When a global advertiser competes with two local advertisers, we have the following:*
  • *The global advertiser bids $b^*_{G,NA} = \gamma v_G\frac{(p_H+p_L)}{p_H+p_L-2dp_Hp_Lv}$ and can reduce their bid when $b^*_{G,NA} > \gamma$. The local advertisers bid their expected value per impression γv_i, but can reduce their bid for high values of v_i when $d < 0$ and still maintain the same profit.*
  • *The global advertiser bids above the local advertisers when $d > 0$ and below the local advertisers when $d < 0$.*
  • *The global advertiser bids below the optimal bid at the popular publisher.*

Lemma [1](#) shows that when a global advertiser is facing two local advertisers, the advertiser will deviate from bidding their expected value for one impression because of the externality between the two publishers. It is interesting to see that when the returns to advertising are decreasing (but still positive), the global advertiser will lower her bids, whereas she will increase them when $d > 0$. When $d < 0$, the ability to track the same customer over multiple publishers increases the potential profit the advertiser can achieve, but decreases the average potential profit per publisher the advertiser will achieve. Thus, maximizing her profit requires the global advertiser to lower her bid. When $d > 0$, however, the externality is strong enough that even increasing her bid will compensate the global advertiser in her revenue

more than the increased cost paid because of stronger competition with the local advertiser.

The third item begins to shed light on which of the publishers may benefit more from efficiency in the market when advertisers become more global. As can be seen, the global advertiser will underestimate the potential income from the popular publisher and will bid less than she could have bid if the market had no uncertainty in it.

Using the bidding strategies from Lemma 1, we can find the expected profits of the global and local advertisers and the revenues of the publishers. We can then compare them to the case when there are no global advertisers. The following proposition summarizes the results.

Proposition 1.

- *The profit of the global advertiser is higher than if it was local. The profit of a local advertiser is higher than if it was competing with a local advertiser when $d < 0$ and lower otherwise.*
- *When $d > 0$ and $v > \frac{\sqrt{4d^2 p_H p_L + 4 d p_L + 1} - 1}{2(d^2 p_H p_L + d p_L)}$, there is no value in using attribution.*
- *When $d < 0$, the expected revenue of each publisher decreases when one advertiser is global compared with when all advertisers are local.*

Proposition 1 shows that when advertising returns are positive, the profit of the global advertiser will increase even without a sophisticated and efficient allocation of advertising budgets. A somewhat surprising result is also that the profits of the local advertisers are increasing when global advertisers can target their customers better and $d < 0$. This phenomenon is a consequence of the weaker bidding strategy the global advertiser is employing, which allows the local advertiser to pay less than she would have paid if she were competing with another local advertiser. Essentially, we see that in this case decreasing advertising returns and the ability to track users across publishers softens the competition in the market.

The second item of the proposition reveals another interesting result: if advertising returns are increasing and the global advertiser's valuation for consumers is high enough, the profit when not using any attribution method will be as high as the optimal profit when the state of the world is known. The intuition behind this result is that when d and v are large enough, the optimal bid of the global advertiser at both publishers is higher than the highest bid of a local advertiser. In this case, the second-price auction serves as insurance for the global advertiser against paying too much per impression. Thus, as long as the bid without attribution is set higher than the bid of the local advertiser, the advertiser will reach maximal profit even without attribution.

Finally, we see that when $d < 0$, publishers will have lower revenues compared with when all advertisers are local. The publishers in this market suffer from two effects. Because one advertiser now faces decreasing returns, her potential extracted value from consumers is lowered, which lowers the rent that the publishers can extract. In addition to that, because the global advertiser also bids with a weaker strategy, the local advertiser pays less for the same ads than she would have paid if she were competing with a local advertiser.

Adding insult to injury, the popular publisher will suffer the biggest loss of revenue as a corollary to Lemma 1. This may partially explain why large publishers (such as Google and Facebook) are eager to provide advertisers with analytics technology (including attribution), which is oftentimes provided for free. If the publishers believe they are more frequented by the target customers of the advertisers, they may benefit from the advertisers using such technology.

These results emphasize how externalities and uncertainty change ad allocation and profits in the market when consumers multihome and an advertiser switches to running a multipublisher campaign. The profits of publishers decrease because the uncertainty about visits lowers the efficiency of the second-price auction in allocating ads and in extracting rent, but also decreases the level of competition between the advertisers. Many solutions for this problem focus on what publishers and platforms can do to improve the allocation. For example, data are aggregated across a network to create an information asymmetry that can be used to price ads differently, and consequentially increase the profits of publishers. In our setup, publishers will not be able to know whether a consumer is multihoming or not and respond accordingly. However, advertisers are able to respond to these effects by using the information they collect in the campaign. In the next sections, we focus on the effect of attribution—the process by which advertisers try to resolve the uncertainty they face and improve their competitive positioning in the market.

5. The Role of Attribution

The previous section found that the global advertiser resorts to submitting an equal bid to both publishers because of the uncertainty about the market state. The goal of an attribution function is to associate revenues with specific publishers. The advertiser can then change her bids until the associated attribution of each publisher is optimized.

Many advertisers define campaign attribution as "the process by which marketers quantify the contribution of various media in an ad campaign to its desired impact" (Bager and Laszlo 2015). The result of this process (typically) is an allocation of credit to the two publishers such that the sum of the allocation equals the total campaign revenue. We will therefore make the following definition.

Definition 1 (Attribution Function). An attribution function separates the revenue r of a campaign that uses

publishers 1 and 2 to two values, $A_1(r)$ and $A_2(r)$, s.t. $r = A_1(r) + A_2(r)$. The revenue A_j is attributed to publisher j.

The results of an attribution function are then typically used to estimate the ROI attributed to each publisher (see, e.g., Borgs et al. 2007, Geyik et al. 2014, Li et al. 2016) and bids are then changed to equalize the ROI between the different platforms. Because maximizing ROI does not necessarily maximize profit, we will analyze the impact of the attribution function when the advertiser sets the bids such that the attributed profit to each publisher is maximized:[8]

Definition 2 (Attribution Maximizing Bids).

$$b_{G1}^* = \arg\max_{b_{G1}} A_1(r(b_{G1}, b_{G2})) - c_1(b_{G1}, b_{G2}),$$

$$b_{G2}^* = \arg\max_{b_{G2}} A_2(r(b_{G1}, b_{G2})) - c_2(b_{G1}, b_{G2}).$$

In this definition, $r(\cdot, \cdot)$ is the total revenue generated by the campaign, and $c_j(\cdot, \cdot)$ is the cost associated with publisher j when bids b_{G1} and b_{G2} are submitted. Because attribution is calculated during a campaign or after a campaign is run, the revenue r and costs c are the actual revenue and costs observed following the campaign, and hence reflect the true state of the world facing the advertiser.

Definition 2 is equivalent to the definition of a Nash equilibrium where the two publishers are allocated a share of the revenue. As such, the attribution process can be seen as creating a virtual competition between the two publishers that is being simulated by the advertiser to find the equilibrium allocation. In this case, the result of Holmstrom (1982) holds, and the allocation of bids will be an inefficient advertising effort (ad exposure). Consequentially, some publishers will always be allocated bids (or budgets) that are suboptimal.

Because there are many functions for which $r = A_1(r) + A_2(r)$, the advertiser's goal is to pick an attribution function that maximizes her profit subject to the constraint that each publisher maximizes its individual attributed profit. When the first-order conditions are sufficient for Definition 2, the following first-order conditions will hold for the optimal bids defined in Definition 2:

$$\frac{\partial A_1(r(b_{G1}^*, b_{G2}^*))}{\partial b_{G1}} = \frac{\partial c_1(b_{G1}^*, b_{G2}^*)}{\partial b_{G1}},$$
$$\frac{\partial A_2(r(b_{G1}^*, b_{G2}^*))}{\partial b_{G2}} = \frac{\partial c_2(b_{G1}^*, b_{G2}^*)}{\partial b_{G2}}. \qquad (4)$$

Using this set of first-order conditions as a constraint, we can now write the full attribution problem as follows.

Definition 3 (Attribution Problem).

$$\max_{A_1, A_2} \pi(b_{G1}^*, b_{G1}^*) \qquad (5)$$

s.t.
$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad (6)$$
$$A_1(r) + A_2(r) = r,$$
$$\frac{\partial A_1(r(b_{G1}^*, b_{G2}^*))}{\partial b_{G1}} = \frac{\partial c_1(b_{G1}^*, b_{G2}^*)}{\partial b_{G1}},$$
$$\frac{\partial A_2(r(b_{G1}^*, b_{G2}^*))}{\partial b_{G2}} = \frac{\partial c_2(b_{G1}^*, b_{G2}^*)}{\partial b_{G2}}. \qquad (7)$$

Definition 3 shows that the goal of the advertiser is to maximize her profit over the space of available attribution functions within the constraints defining the problem. This complex optimal control problem does not necessarily have a closed-form solution. The solution also highly depends on the parameters of the profit function of the campaign, which is why we focus on analyzing two commonly used attribution mechanisms, namely, the LT method and the Sh method. The online appendix shows that these methods belong to a family of methods we call fixed externality share methods and analyzes their efficiency. In what follows, we compare these popular attribution functions to two benchmarks—the case of no attribution (NA) and the optimal allocation possible when there is no market uncertainty (Opt). Using this analysis, we are also able to calculate an upper bound on how optimal (or suboptimal) each attribution method is, and under what conditions. We perform the analysis in the case of decreasing returns ($d < 0$) and discuss the impact of increasing returns in the conclusion.

5.1. Last-Touch Attribution and Shapley Value Attribution

Last-touch attribution allocates credit to the last publisher that has shown an ad prior to conversion. In our model, because consumers may convert immediately after seeing an ad, the revenue from a consumer that was exposed to two ads may be attributed to both the first publisher and the second publisher depending on the timing of conversions.

Using the general profit function of the advertiser in Equation (2), the revenue attributed to publisher 1 assuming it is the first in order of visits using last-touch will be

$$A_1^{LT} = p_1(1 - p_2) \int_0^\infty \mathbb{I}[b_{G1} > b_1] \gamma v_G dF_1(b_1)$$

$$+ p_1 p_2 \int_0^\infty \int_0^\infty \mathbb{I}[b_{G1} > b_1] \mathbb{I}[b_{G2} <= b_2]$$
$$\cdot \gamma v_G dF_1(b_1) dF_2(b_2)$$
$$+ p_1 p_2 \int_0^\infty \int_0^\infty \mathbb{I}[b_{G1} > b_1] \mathbb{I}[b_{G2} > b_2]$$
$$\cdot \gamma v_G dF_1(b_1) dF_2(b_2)$$
$$= p_1 \gamma v_G \int_0^\infty \mathbb{I}[b_{G1} > b_1] dF_1(b_1) = p_1 \gamma v_G F_1(b_{G1}).$$
$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad (8)$$

The revenue attributed to publisher 2 under last-touch assuming it is visited second by consumers will be

$$
\begin{aligned}
A_2^{LT} = {} & (1-p_1)p_2 \int_0^\infty \mathbb{I}[b_{G2} > b_2]\gamma v_G dF_2(b_2) \\
& + p_1 p_2 \int_0^\infty \int_0^\infty \mathbb{I}[b_{G1} <= b_1]\mathbb{I}[b_{G2} > b_2] \\
& \quad \cdot \gamma v_G dF_1(b_1)dF_2(b_2) \\
& + p_1 p_2 \int_0^\infty \int_0^\infty \mathbb{I}[b_{G1} > b_1]\mathbb{I}[b_{G2} > b_2] \\
& \quad \cdot \gamma(1+d)v_G dF_1(b_1)dF_2(b_2) \\
= {} & p_2\gamma v_G \int_0^\infty \mathbb{I}[b_{G2} > b_2]dF_2(b_2) \\
& + p_1 p_2 d\gamma v_G \int_0^\infty \int_0^\infty \mathbb{I}[b_{G1} > b_1]\mathbb{I}[b_{G2} > b_2] \\
& \quad \cdot dF_1(b_1)dF_2(b_2) \\
= {} & p_2\gamma v_G F_2(b_{G1}) + p_1 p_2 d\gamma v_G F_1(b_{G1})F_2(b_{G2}). \quad (9)
\end{aligned}
$$

Inspecting the two attribution functions, we see that increasing the bid at publisher 1 will decrease the attributed revenue to publisher 2 when $d < 0$ and will increase it when $d > 0$ because of the externality term. In other words, publisher 2 receives (or suffers) the impact of the externality between the publishers because it comes later in the visit process of the consumer.

It is easy to show that when using the last-touch attribution function, the first-order conditions for the optimal bids are also sufficient to fulfill the conditions in Definition 2 on the optimal bids. Hence, we can calculate the new equilibrium bidding strategies of the local and global bidders and analyze the resulting profit of each bidder.

Performing the analysis yields the following result:

Proposition 2. *Using last-touch attribution, we have the following:*
- *The global advertiser will bid $b_{Gj_1}^* = \gamma v_G$ if j_1 is the first publisher visited and $b_{Gj_2}^* = \gamma v_G + p_{j_1}d\gamma v_G^2$ if j_2 is the second publisher visited. The local advertisers will bid their expected values per impression.*
- *The number of ads allocated to the global advertiser will increase beyond the optimal number when $d < 0$.*
- *The expected profit of the global advertiser will be higher than without using attribution when $d < 0$ only if p_L is small.*

Proposition 2 illustrates how last-touch attribution distorts the incentives of an advertiser causing her to overbid and as a result overspend and receive lowered profit in an advertising campaign.

The intuition behind the results is that last-touch attribution causes the elasticity of the bid b_{G1} to be estimated from the revenue generated at publisher 1 and disregards the externality imposed on the revenue at publisher 2.

A surprising result from this analysis is that, counterintuitively, last-touch attribution causes an overestimate of the returns of the publishers that are *earlier* in the conversion funnel, and not necessarily later. This is because, in a concave revenue function, earlier ad impressions have a negative externality on later impressions, and not taking them into account overestimates the impact of earlier impressions. In addition to the earlier publisher benefitting, we find that the bid at the second publisher is lower when $d < 0$, and even lower when the first publisher is popular. That means that niche publishers that come later in a conversion funnel will suffer more when last-touch attribution is applied.

Finally, we see that applying last-touch attribution actually reduces the profit of the advertiser below the profit she would make without using any attribution function at all, unless p_L is small (has to be less than $1/3$, and usually smaller). When p_L is small, the market is one where most consumers visit only one publisher. In that case, the timing of visits of customers is indicative of both which has the higher probability of visit and which is visited first. Last-touch attribution is well suited to measure order effects, if they are indicative of the actual process that determines conversions, and when most consumers visit only one publisher.

This analysis gives a potential explanation for why advertisers have chosen to use last-touch attribution beyond the simplicity of its implementation. In a market where externalities are small in magnitude because of low co-occurrence of visits of customers, last-touch attribution can be used to properly model the order of visits and improve the profits of advertisers compared with not using attribution.

Another attribution method that has been proposed in the literature is using the Shapley value to attribute profit to publishers. This method has become popular recently as the "data-driven" method implemented in Google's attribution solution.[9]

The Shapley (1952) value is a cooperative game theory solution concept that allocates value created among players in a cooperative game. The value is defined as follows for our model:

$$
\begin{aligned}
A_1^{Sh} &= \frac{r(b_1, b_2) - r(0, b_2) + r(b_1, 0) - r(0,0)}{2} \\
A_2^{Sh} &= \frac{r(b_1, b_2) - r(b_1, 0) + r(0, b_2) - r(0,0)}{2}.
\end{aligned}
\quad (10)
$$

The value takes the average marginal contribution of the publishers at the points (b_1, b_2) and $(b_1, 0)$ for publisher 1 and (b_1, b_2) and $(0, b_2)$ for publisher 2.

The reason to focus on the Shapley value is that it has two desirable properties in an attribution function. The first is called the efficiency property, which means that all value created by the players is allocated, and the sum of the values equals the total revenue generated.

The second property, described in Young (1988), is called the marginality property, which says that if two players have the same marginal contribution in all possible order of players, their attribution will be equal. Moreover, the Shapley value is agnostic of the actual function that generates conversions and makes no assumptions on the shape of the function. Hence, it is quite robust to model misspecification.

The analysis of the equilibrium bids and allocation when using Shapley value attribution shows the following result:

Proposition 3. *When using the Shapley value to optimize the bids of the advertiser and when $d < 0$, we have the following:*

• *When v is low and p_L is low, the global advertiser bids less than the optimal value at the popular publisher, and bids above the optimal value otherwise. The global advertiser always bids above the optimal value at the unpopular publisher.*

• *When $v < \frac{\sqrt{3}-1}{-dp_H}$, the profit using Shapley value attribution is higher than the profit with no attribution. When $v \geq \frac{\sqrt{3}-1}{-dp_H}$, the profit is higher than no attribution if p_L is low.*

• *There exists a p_L^c such that if $p_L < p_L^c$, the profit using Shapley value attribution is higher than both the last-touch and the no-attribution profits.*

• *There is no case when $d < 0$ in which last-touch attribution generates more profit than both no attribution and Shapley value attribution.*

Proposition 3 shows that using the Shapley value as an attribution mechanism is beneficial when the value of a conversion is below some threshold, or for high values of v when p_L is small enough. The upper bound on the value of v at the worst case is $\sqrt{3} - 1 \approx 0.732$, and increases with lower p_H and d closer to zero. Consequentially, the Shapley value increases the profit of a global advertiser for a wide range of parameter values.

The second item shows that when p_L is low enough, implying there is enough asymmetry between the two publishers, the Shapley value yields the highest profit among the three attribution methods we are comparing. The intuition is that in these cases, similarly to the last-touch method, the Shapley value takes into account the asymmetric effect on conversion, which comes mostly from the popular publisher. Unlike last-touch attribution, however, the Shapley value is order agnostic and thus is better able to take into account the externality imposed by the first visited publisher on the second publisher.

The last item of the proposition shows the surprising result that last-touch attribution is not beneficial to advertisers on the entire range of the parameters when the other two methods are available. We basically see that for low values of p_L, when last-touch attribution is more effective than not using attribution, the Shapley value is even more effective, whereas for high values of p_L, either no attribution or the Shapley value is more effective.

Another interesting conclusion from this analysis is that when visit probabilities are high and when valuations are high, there is little benefit to using any attribution process. This result can explain why advertisers may be dissatisfied after they employ an attribution method and discover no benefit (or even decreased profits). The intuition behind this result is that when p_H and v are high enough, the optimal bids of the advertisers when they have complete information will be to bid the same value at both publishers. In this case, there will be no difference in bidding strategies between not having any information (the NA case) and having complete information (the Opt case). If the global advertiser decides to apply an attribution method in this setting, she will find out that she cannot improve her profit beyond the optimal profit, and will be surprised to learn that attribution has little value for her.

5.2. Impact on Ad Allocation Efficiency

Advertisers usually see the goal of attribution as helping them to compete in the market against better informed advertisers. Publishers may also prefer that advertisers use specific attribution methods versus others, as it may cause overbidding and enhance their profits. The research on uncertainty in advertising markets that takes the publisher perspective often focuses on methods that will increase the ad allocation efficiency in the market. It is unclear, however, whether allocation efficiency should be the goal of an advertising network or platform, or of an advertiser. Efficient allocation mechanisms maximize welfare, but not necessarily the profit of the platform, which is the market designer in this case. For example, the commonly used second-price auction is efficient for one item, but does not maximize the profit of the platform in many cases.

To analyze the efficiency of ad allocation, we will focus on the welfare implications of uncertainty and attribution. Focusing on welfare allows us to perform an analysis without too many assumptions about the allocation mechanism in the market, which may impact ad allocation but not total welfare. The total welfare in our model is the sum of profits of advertisers and publishers. Because the income of publishers comprises only transfers from advertisers, the total welfare is equal to the sum of revenues of advertisers. This simplifying observation allows us to find the maximum potential welfare without any assumption about the specific ad mechanism the publishers may use. Because the welfare is equal to the sum of advertiser revenues, the most efficient ad allocation is the one that generates the most revenue from visiting consumers, conditional on their visit path.

After we find this revenue, we compare it to the revenues achieved by the advertisers when they use the different attribution methods we analyze (NA, Sh, LT, and Opt) and when the publishers use the second-price auction we described before.

Performing the analysis yields the following result for the optimal outcome:

Lemma 2. *When d < 0, the maximum expected revenue possible in the market is*

$$\frac{1}{12}\gamma\left(p_H(d^3 p_L + 3d^2 p_L + 3d p_L + 8) + 8p_L\right).$$

The revenue increases with p_H, p_L, and d.

Although it is quite straightforward to find the welfare maximizing allocation,[10] calculating the total market revenue is surprisingly complex for the different attribution methods. The difficulty comes from the fact that the global advertiser, who is exposed to externalities, usually uses nonlinear bidding strategies in equilibrium. This, in turn, makes it difficult to calculate the revenues of the local advertisers. We are able to find closed-form results for the market revenues in the NA, LT, and Sh cases. We are unable to find a closed form for the total revenue in the Opt case, and we therefore use numerical computation.

Figure 1 shows the results of this analysis for two cases. Both panels display the ratio of the expected revenue of advertisers using the methods we analyzed to the expected revenue of advertisers in the optimal allocation scenario. In the left panel, the worst-case scenario for the Shapley value is displayed, which occurs when $p_H = 1$, $d = -1$, and $p_L = 1$. In this extreme case, the advertisers would like to show only the first ad in a sequence (as the second is worthless). Even in this extreme case, the Shapley value is able to achieve 94.33%

of the potential maximum welfare. Numerical analysis shows that last-touch attribution achieves 94.5% and no attribution reaches 95.4% in their respective worst-case scenarios. These results generally illustrate that market inefficiency is not high even when advertisers have complete certainty about the market state.

The right panel presents the case of $p_H = 1/2$, $p_L = 1/4$, and $d = -1/4$. All of the methods in this case reach over 99% of the maximum potential welfare, and the differences between them are very small. We see that last-touch attribution is quite efficient for low values of p_L. These are the cases where most visits go to one publisher and the externality is very small. As in the case of the global advertiser's profit before, such markets are ideal for using a method like last-touch attribution. When p_L is higher, we see that the Shapley value improves efficiency beyond last-touch attribution, but is unable to improve it above the cases of Opt and no attribution. Because the Shapley value is designed as a method to give advantage to advertisers and not to generate market efficiency, this result is not surprising.

The conclusion from the analysis is that attribution methods may have a nontrivial impact on welfare. A method like last-touch attribution, that encourages an advertiser to overbid, may increase welfare by transferring revenue to the publishers and to the local advertisers. A method like the Shapley value, that may create less aggressive bidding, may actually lower welfare. In all cases, however, the decrease in welfare is not dramatic compared with having complete information. The majority of the impact of attribution

Figure 1. (Color online) Expected Relative Welfare of the Sh, LT, NA, and Opt Cases to the Maximum Potential
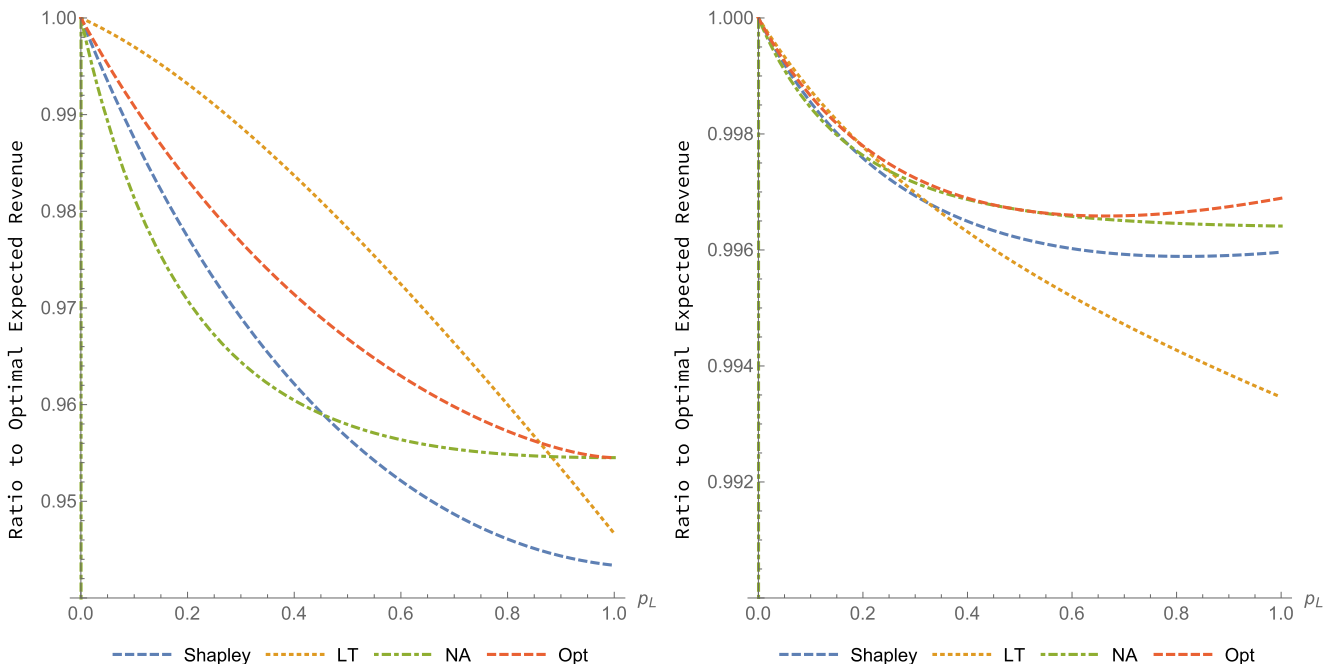
Notes. In the left panel, $\gamma = 1/2$, $p_H = 1$ and $d = -1$. In the right panel, $\gamma = 1/2$, $p_H = 1/2$ and $d = -1/4$.

methods will be from transferring profits from publishers to advertisers, but not by decreasing allocation efficiency in the market.

5.3. Optimality for Advertisers

Questions that naturally arise from the results of the previous section are how well the different attribution methods perform and how far they are from the optimal profit that an advertiser can achieve with full information (Opt). To perform this analysis, we calculate the ratio of the expected profits π^*_{NA}, π^*_{LT}, and π^*_{Sh} for no attribution, last-touch attribution, and Shapley value attribution, respectively, to the optimal expected profit π^*_{Opt} when $d < 0$. We call these ratios the conditional ratios on v. We further integrate the expressions for the expected profits over v and then take their ratios. We call these ratios the unconditional ratios. For example, for the Shapley value, the conditional ratio is

$\frac{\pi^*_{Sh}}{\pi^*_{Opt}}$, and the unconditional ratio is $\frac{\int_0^1 \pi^*_{Sh} dv}{\int_0^1 \pi^*_{Opt} dv}$.

For each of the ratios, we calculate the parameter values where they achieve their minimum and maximum. Because the expressions do not have closed forms, we perform a numeric analysis to find these values.[11] For completeness, the expressions are detailed in the appendix.

The conditional ratio indicates how each method performs for specific environment values as well as consumer values for an action, whereas the unconditional ratio gives the advertiser an indication of how their profit will deviate from optimum over multiple campaigns on average. This analysis can naturally be extended to include distributions for values of p_H, p_L,

and d, though these parameters should fluctuate less often than the values of conversions.

The results of the analysis show that all methods achieve a maximum ratio of 1 for different parameter values. As such, there are multiple conditions under which each method may be optimal, which is why we focus on the worst-case outcomes. When analyzing the worst and average worst-case scenarios, we find that in the conditional case, the ratios are all minimized when $p_H = 1$, $v = 1$, and $d = -1$. In other words, all methods perform the worst in the extreme case when one publisher is visited all the time and where only the first impression has impact on the consumer (because $d = -1$). The minimal ratios are 0.875 for the Shapley and last-touch methods, and ~0.888 for the no-attribution method. Figure 2 illustrates the value of the ratio when $p_H = 1$, $d = -1$, and $v = 1$ for different values of p_L. We also see in this figure that in this extreme case the Shapley value performs quite well when $p_L < 0.5$, with profits achieving 90% of the potential optimal profit.

The results from the analysis of the unconditional ratios also show that the minima are achieved at the extreme cases when $p_H = 1$ and $d = -1$. In this case, not using attribution achieves 93.9% of the optimal profit, the Shapley value achieves 92.8%, and last-touch attribution achieves 92%. We also see that when $p_L < 0.4$, the Shapley value achieves at least 95% of the optimal profit, whereas when $p_L > 0.5$, using no attribution achieves at least 95% of the profit.

These results taken together illustrate the implications of the propositions from the previous sections. When the asymmetry between publishers or the effect

Figure 2. (Color online) Conditional Expected Profit Ratios When $v = 1$, $p_H = 1$, and $d = -1$
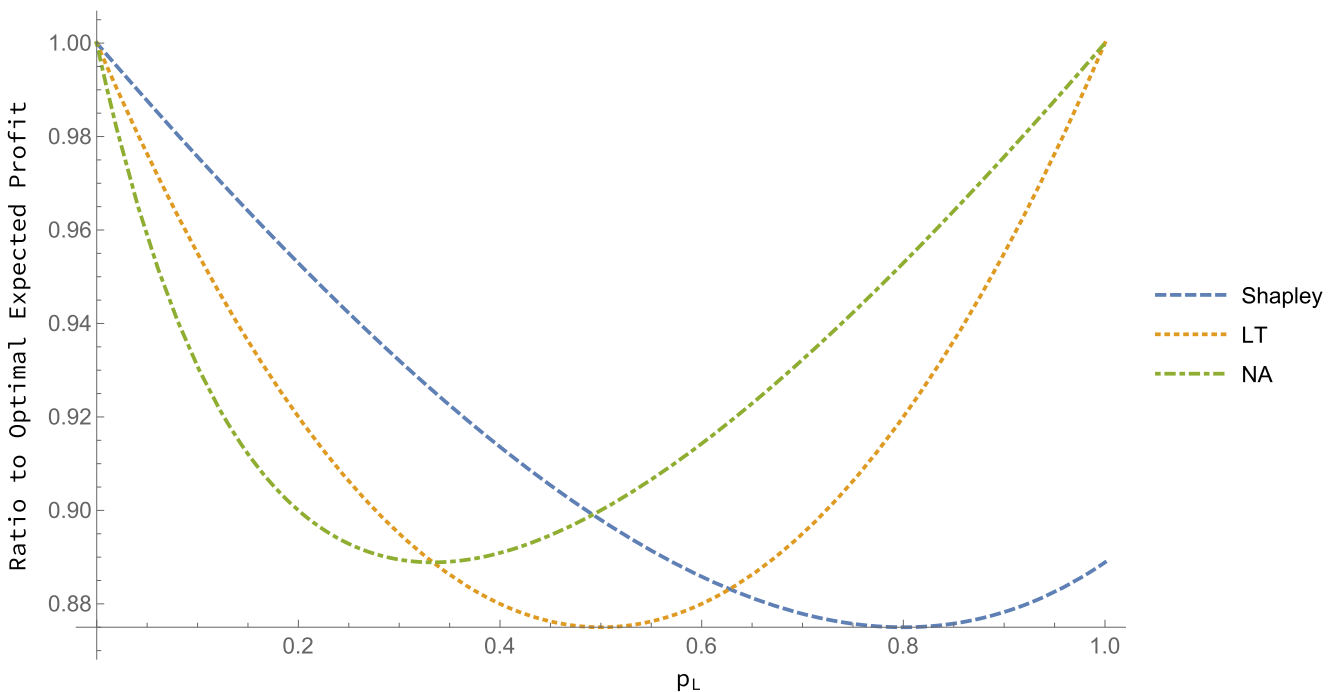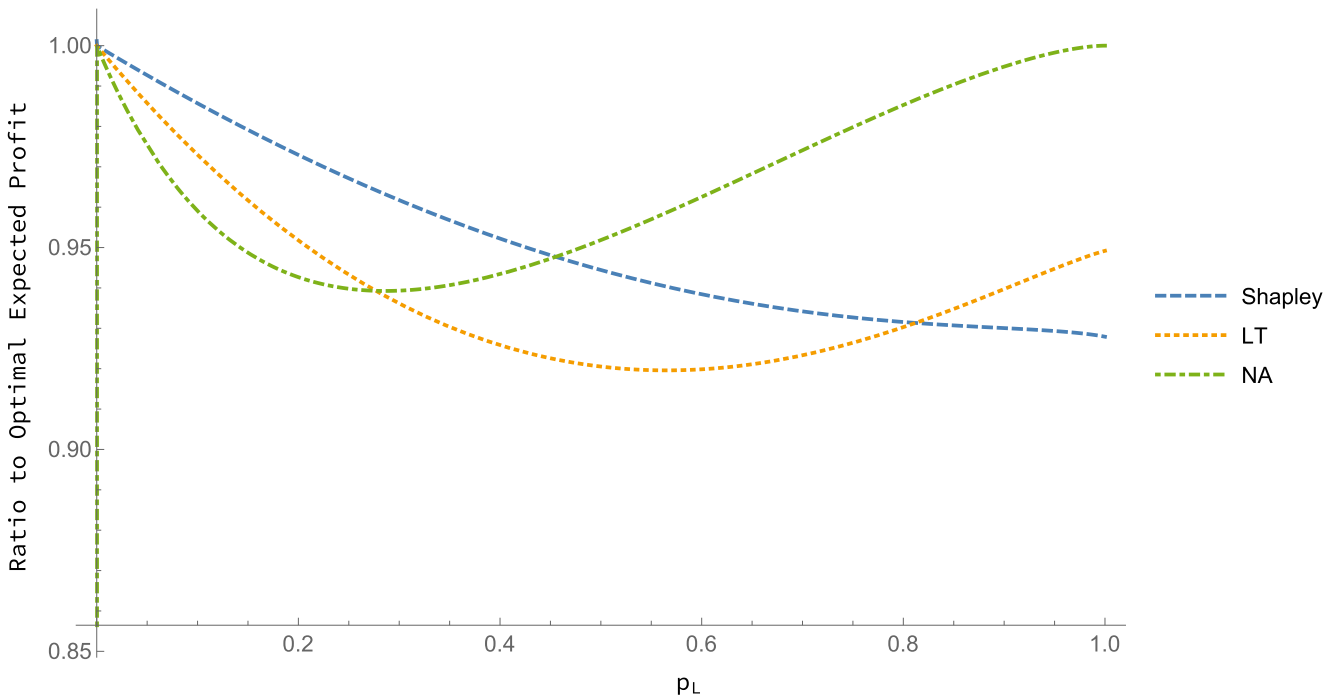
Figure 3. (Color online) Unconditional Expected Profit Ratios When $p_H = 1$ and $d = -1$

of the second ad is not substantial, using attribution methods may be countereffective for an advertiser, as it will reduce the effectiveness of advertising by creating a competition between publishers when one is not needed. When there is enough asymmetry, however, or when the externality is strong, using attribution, and particularly the Shapley value may yield advertiser profits that are close to optimal.

An additional point in the interpretation of these results is that because our model only has two publishers, we expect the effect of full information to not yield as much improvement in allocation efficiency as in the case of competition between a larger number of publishers. Similarly to the case of a centralized planner vs an oligopoly, increasing the number of competing firms will make the equilibrium result closer to the Pareto-efficient outcomes. As campaigns become more complex and with more publishers, we expect the effect of attribution to be more noticeable when compared with not using attribution and for these results to be sharpened.

5.4. Comparative Statics

The conclusion of the previous section was that an attribution scheme may not necessarily benefit advertisers compared with not using attribution at all. In this section, we analyze the impact of changes in the environment on the profits of advertisers using different schemes to understand the potential robustness of each attribution method.

Integrating an attribution method into an advertiser's campaign requires data integration from multiple sources and fine tuning the algorithm parameters to match the

market state. When markets change (for example, when one publisher becomes more popular), it is desirable that an attribution scheme remains as effective and not require replacement or reconfiguration.

We perform a comparative statics analysis on the outcomes from the previous sections to check the robustness of the different attribution schemes to market changes. The analysis reveals the following:

Proposition 4. *When market parameters change, we have the following:*
- *The profit of the global advertiser increases under all three methods (LT, NA, Sh) with an increase in γ, d, v, or p_H.*
- *The profit of the global advertiser increases with p_L only when v is small enough.*

Proposition 4 shows that the attribution methods we analyze are robust to market changes unless p_L changes and the asymmetry between the publishers decreases for high values of v. When p_L increases, the effect of the externality increases, which may lower the benefit of the attribution schemes (and increase the bias from the no-attribution method), because of the increased possibility of seeing the same consumer twice in the campaign and the increased bid at the publisher with lower visit popularity. This effect increases competition and may decrease the profit of the advertisers if they mostly win impressions, which happen when v is high. If v is low enough, the increased visit probability p_L contributes to increased revenues, but the increase in cost does not offset it and the profits increase.

In contrast to these results, the expected optimal profit (when there is no market uncertainty) increases

for an increase of any parameter. One conclusion is that for an advertiser with a high value for consumers, attribution may be a double edged-sword if she advertises on many websites, including nonpopular ones.

Another action an advertiser may take is to try and influence consumers by changing her ads so they have a stronger effect in the first impression by increasing γ to $\hat{\gamma} > \gamma$ while maintaining the overall conversion rate after seeing two ads at $\gamma + \gamma(1 + d)$. In this case, the advertiser controls how conversions are allocated between the first and second impressions, for examples by changing the ad copy.

The analysis shows the following results:

Proposition 5. *When the advertiser increases the impact of the first impression, we have the following:*
- *The distance from the optimal profit will increase without attribution or when using the Shapley value.*
- *There is a range of parameters of d, p_H, and p_L such that if $v > \frac{\sqrt{10}-1}{3}$, the distance from the optimal profit will decrease using last-touch attribution.*
- *Under last-touch attribution, the attributed revenue to the first publisher will always increase, and will increase for the second publisher only when v is small.*
- *Using the Shapley value, there is a range of values of p_H such that when $p_L \cdot v$ is high enough, the attribution to both publishers decreases.*

The analysis in Proposition 5 sheds more light on the effects of the different attribution methods and which types of publishers should benefit from which methods. The first two items show that when the conversions shift toward the first impression, last-touch attribution is advantageous for advertisers because it gets closer to the optimal profit. The intuition is similar to that in the case of low p_L values found before: when there is high asymmetry between the contribution of each impression, last-touch attribution can benefit the first advertiser by overemphasizing their effect on conversions.

The last two items compare the effect of increasing γ to $\hat{\gamma}$ on the attributed revenue in equilibrium. The results for last-touch attribution are expected and mimic the decrease in distance from optimal profit. We also see that there are cases where both publishers benefit from such a reallocation of conversion rates, which creates a unique case where last-touch attribution increases the profit of the advertiser and the attributed revenues of the publishers. The Shapley value, however, has a range of parameters were all players suffer. Both the distance from optimal profit increases, and in addition, the revenue attributed to both publishers decreases. The difference stems from the fact that the Shapley value allocates the externality d equally between both publishers, and in such a case, if γ is increased to $\hat{\gamma}$, it means that d has decreased and that the attributed revenue to both publishers may decrease.

5.5. Ad Prices and Publisher Profits
Finally, we compare the revenues of the two publishers and the expected market price of advertising when using the different schemes. Our goal is to understand how moving from a simple attribution method like last-touch attribution to a more sophisticated method like the Shapley value will influence the ad price dynamics in the market.

Proposition 6. *When $d < 0$, we have the following:*
- *The first publisher will have a higher revenue under last-touch attribution than under the Shapley value. The second publisher will have a higher revenue under the Shapley value than under last-touch attribution.*
- *The expected price of an impression the global advertiser will pay is higher under last-touch attribution than using either the Shapley value or not using attribution.*

Proposition 6 shows that publishers that arrive earlier in a conversion funnel will prefer last-touch attribution over the Shapley value attribution. The intuition comes from the previous results, where the revenue attributed to the first publisher increases under last-touch attribution. This yields more competitive bidding by the global advertiser, which translates to higher publisher profit.

The second item shows an interesting result that may explain the current market trends in online advertising where ad prices are decreasing after a period of increase for mature technologies such as Google AdWords. When firms moved from not using attribution at all to using last-touch attribution, the price of advertising was expected to increase for the large and global advertisers because of overbidding and increased competition. When they move to more sophisticated (and efficient) methods, the result is a decrease in the price paid per impression because of softening competition. Our model thus provides a prediction for trends in advertising prices that will arise when more advanced measurement technologies are introduced.

6. Discussion, Limitations, and Conclusion
As multipublisher campaigns become common and new advertising technologies appear in the market, advertisers find more opportunities to interact with consumers over multiple touch points. The cumulative effect of advertising creates externalities between individual ad exposures for multihoming consumers, and because these exposures are on different publishers, the measured effectiveness of each publisher will depend on past exposures of each consumer on other publishers.

As we have shown using our model of a global advertiser competing with two local advertisers, externalities between exposures coupled with uncertainty about consumer visits to publishers lower the bidding incentives of global advertisers and create inefficiency in the market that decreases the revenues of global advertisers and publishers. Past attempts to analyze this problem have shown that if publishers have an

information advantage, for example, by being able to track consumers over a network, then they can exploit this asymmetry to increase the efficiency of ad allocation in the market through changing the pricing in the market. In contrast to these analyses, our focus on the case of structural uncertainties, in which publishers cannot gain superior information, allows us to analyze the impact and potential for using advertiser-side multitouch attribution. More often than not, publishers will not be able to identify multihoming consumers or estimate the externalities from previous exposures correctly. The results we derive assist in understanding how attribution may serve to change the incentives in an advertising market and potentially benefit advertisers. Unlike the focus on advertising optimality in terms of allocation efficiency and social welfare, we try to understand how advertisers respond to these uncertainties.

Multitouch attribution uses the information collected by the advertiser during a campaign to estimate the contribution of each ad exposure (touch) to consumer conversions. Our analysis of the optimization problem the advertiser faces when using multitouch attribution shows that this problem is an instance of the team problem identified in Holmstrom (1982). Because of this analogy, attribution methods are not expected to fully help the advertiser reverse the effects of uncertainty and externalities. Furthermore, when taking the equilibrium response of competing advertisers into account, we found that in many cases a simple attribution method such as last-touch attribution may lower the profits of a global advertiser compared with not using attribution at all. Using a method such as the Shapley value, which is less extreme in credit allocation, was found to improve the profits of advertisers. In both of these cases, however, we are able to show that attribution may hold limited potential when an advertiser values consumers highly in the market. In these scenarios, uncertainty about consumer visits would not impact the bidding strategy of an advertiser, and in this case, attribution may not yield any improvement over not using it at all. These findings contribute to the empirical literature on attribution modeling that assumes that more accurate measurement of the uncertain state of the campaign will always benefit advertisers. Our results quite clearly show that attribution might be counterproductive because of its competitive effect.

Several interesting questions arise from our analysis and hold potential for further research. The first is analyzing attribution as a fully strategic choice by advertisers and publishers, and not just as a measurement technique. The online appendix uses numerical analysis to look at the case of two global advertisers who compete in the market using CPC pricing. We find that our results hold qualitatively, but we also find that different amounts of information held by the global advertisers may cause a softening of the competition in the market. A second constraint in our analysis restricts the publishers to not respond to the actions of the advertisers by changing their auction structure or the market structure. It is possible, however, that when faced with attributing advertisers, publishers may decide to invest in attracting different segments of customers, or develop technology that will allow them to offer behavior-based segmentation to advertisers. Merging the two research areas on behavioral targeting on the publisher side with attribution on advertiser side would be a logical next step in this area.

A second question is how optimal are the specific attribution methods we analyze in comparison with alternative methods? The online appendix shows that both last-touch attribution and the Shapley value are part of a family of methods we call fixed externality share methods, and that in some cases the Shapley value is optimal, whereas last-touch attribution is never optimal. In a second analysis, we compare the attribution methods to the maximal gain from using media-mix modeling (MMM). In that analysis, which utilizes results from the theory of team compensation, we are able to show that attribution always gains more than MMM because of its access to individual level exposure information. These analyses and our formulation of the attribution problem as an optimal control problem open an interesting area for further research that focuses on finding optimal attribution methods.

Our paper concludes with an analysis of the welfare implications and stability of attribution methods in changing markets. Because attribution methods are employed by advertisers that are not the market designer, we find that they often create a trade-off between market optimality, measured in allocation efficiency and social welfare, and advertiser optimality, which focuses on increasing an advertiser's profits. The focus on allocation efficiency in prior research often leads to a similar conclusion: increased efficiency comes at the expense of advertiser profits. However, if advertisers are unable to field profitable advertising campaigns in the market, they may decide to leave the market altogether, as recently evidenced by some large advertisers. Focusing on the development of attribution methods and their impact on market outcomes may serve to counteract this trend, and our research takes a step forward in that direction.

Acknowledgments

Appendix. Proofs
Proof of Lemma 1

Suppose the equilibrium bidding strategies of the local players have the CDFs $F_1(\cdot)$ and $F_2(\cdot)$ and those of the global advertiser have the CDFs $F_{G1}(\cdot)$ and $F_{G2}(\cdot)$ when observed from viewpoint of the competing advertiser.

We first need to determine whether the local and global advertisers have full support of their bidding strategies in the range $[0, \gamma]$. Suppose the highest bid a global advertiser will submit is $\bar{b}_i$ at publisher i. We assume that F_i's for $i \in \{G1, G2\}$ have probability density functions denoted $f_i(\cdot)$, and that $f_i(b) > 0$ for $0 \leq b \leq \bar{b}_i$ and $f_i(b) = 0$ when $b > \bar{b}_i$.

The expected profit of a local advertiser i is

$$\pi_i(b_i) = 1/2(p_H + p_L) \int_0^\infty \mathbb{I}[b_i > b_{Gi}](\gamma v_i - b_{Gi}) dF_{Gi}(b_{Gi}). \quad (A.1)$$

The local advertiser will never bid above γ. If they bid above γ and $b_{Gi} > \gamma$, the advertiser can increase her profit on that bid by lowering it.

Suppose $\bar{b}_i \geq \gamma$. In that case, the local advertiser's expected profit can be written as

$$\pi_i(b_i) = 1/2(p_H + p_L) \int_0^{b_i} (\gamma v_i - b_{Gi}) f_i(b_{Gi}) db_{Gi}. \quad (A.2)$$

Solving for the first-order condition yields $b_i^* = \gamma v_i$, and $F_i(b) = b/\gamma$.

If $\bar{b}_i < \gamma$, then when the local advertiser bids above $\bar{b}_i$, she will have the same profit for any value of her bid. Solving the maximization problem the advertiser's bid will be the same as before, but can be lower than γv_i if $v_i > \bar{b}_i/\gamma$.

The first-order conditions for the profit of the global advertiser are

$$\frac{p_H + p_L}{2}(\gamma v_G - b_{G1}) f_1(b_{G1}) + p_H p_L d \gamma v_G f_1(b_{G1}) F_2(b_{G2}) = 0, \quad (A.3)$$

$$\frac{p_H + p_L}{2}(\gamma v_G - b_{G2}) f_2(b_{G2}) + p_H p_L d \gamma v_G f_2(b_{G2}) F_1(b_{G2}) = 0. \quad (A.4)$$

We can see that when the global advertiser bids above the local advertiser's highest bid, any bid will generate the same profit.

Suppose the highest bid of the local advertisers is γ. Then the solution to the FOCs is

$$b_{G1,NA}^* = b_{G2,NA}^* = \gamma v_G \frac{p_H + p_L}{p_H + p_L - 2dp_H p_L v_G}. \quad (A.5)$$

When $v_G > \frac{p_H + p_L}{p_H + p_L + 2dp_H p_L}$, the global advertiser can reduce her bid (as long as it is higher than γ) and still maintain her profit. When $d > 0$, the highest bid of the global advertiser is higher than γ, and thus these two bidding strategies constitute an equilibrium, but the global advertiser can also reduce her bids for high values of v_G.

When $d < 0$, the highest bid of the global advertiser will be lower than γ, and the bidding strategies are also an equilibrium. In this case, the local advertisers can reduce their bids for high values of v_i.

Comparing the bidding strategies of the local and the global advertisers, we can see that the global advertiser bids above the local advertiser when they have the same valuations if $d > 0$, whereas the global advertiser bids below the local advertiser if they have the same valuation when $d < 0$ and $\bar{b} > 0$, proving the second item of the lemma.

To compare the global advertiser's bid with uncertainty to the optimal bid without uncertainty, we solve for the same equilibrium but using the profit function in Equation (2) and not the expected profit.

The equilibrium solution (using the same method as above) is that the local advertisers maintain bidding γv_i, whereas the global advertiser bids $b_{G1,Opt}^* = \gamma v \frac{1 + dp_L v}{1 - d^2 p_H p_L v^2}$, $b_{G2,Opt}^* = \gamma v \frac{1 + dp_H v}{1 - d^2 p_H p_L v^2}$. Comparing these values to $b_{G,NA}^*$, we find the result of the third item.

Proof of Proposition 1

To prove the items in the proposition, we calculate the expected profits of the global and local advertisers using the bidding strategies from Lemma 1, as well as the costs both advertisers face, which constitute the revenue of the publishers. In addition, we calculate the expected profit and costs of a local advertiser facing a local advertiser. Throughout the proof, we will make repeated use of the logarithm inequality $\log(1 + x) \geq \frac{2x}{2+x}$ when $x > 0$ and the reverse otherwise. The calculations result in

$$\pi_{Local}^* = \frac{1}{4} \gamma v^2 (p_H + p_L); \quad (A.6)$$

$\pi_{Local,NA}^*$

$$= \begin{cases} \left[\gamma(p_H + p_L)^2 \left(2dp_H p_L v + (p_H + p_L) \log\left(\frac{p_H + p_L}{2dp_H p_L v + p_H + p_L}\right)\right)\right] \Big/ \\ \quad (8d^2 p_H^2 p_L^2) \qquad v \leq \frac{p_H + p_L}{p_H + p_L - 2dp_L p_H}, \\ \left[\gamma(p_H + p_L)\left(2dp_H p_L (2dp_H p_L v + p_H + p_L)\right. \right. \\ \quad \left. \left. + (p_H + p_L)^2 \log\left(\frac{-2dp_H p_L + p_H + p_L}{p_H + p_L}\right)\right)\right] \Big/ (8d^2 p_H^2 p_L^2) \\ \quad d \leq 0 \text{ and } v > \frac{p_H + p_L}{p_H + p_L - 2dp_L p_H}; \end{cases}$$

$$(A.7)$$

$\pi_{NA,Local}^*$

$$= \begin{cases} [\gamma v^2 (p_H + p_L)^2] / [2(p_H + p_L) - 4dp_H p_L v] \\ \qquad v < (p_H + p_L)/(2dp_H p_L + p_H + p_L), \\ \frac{1}{2}\gamma(2v(dp_H p_L + p_H + p_L) - p_H - p_L) \\ \qquad v \geq (p_H + p_L)/(2dp_H p_L + p_H + p_L). \end{cases}$$

$$(A.8)$$

In the above notation, π_{Local}^* is the profit of a local advertiser competing with a local advertiser; $\pi_{Local,NA}^*$ is the profit of a local advertiser competing with a global advertiser, and $\pi_{NA,Local}^*$ is the profit of a global advertiser competing with a local one.

Proving the first item is a simple comparison of the profit $\pi^*_{NA,Local}$ and π^*_{Local}. To prove the second part of the first item, the difference $\pi^*_{Local,NA} - \pi^*_{Local}$ when $d < 0$ equals

$$\pi^*_{Local,NA} - \pi^*_{Local}$$

$$= \begin{cases} \frac{1}{8}\gamma(p_H + p_L)\left(\left(\left[(p_H + p_L)\left(2dp_Hp_Lv + (p_H + p_L)\right)\right.\right.\right. \\ \left.\left.\left. \cdot \log\left(1 + \frac{-2dp_Hp_Lv}{p_H + p_L}\right)\right)\right] \middle/ (d^2p_H^2p_L^2)\right) - 2v^2\right) \\ \qquad\qquad\qquad\qquad v \le \frac{p_H + p_L}{p_H + p_L - 2dp_Lp_H}, \\ \frac{1}{8}\gamma(p_H + p_L)\left(\left(\left[2dp_Hp_L(2dp_Hp_Lv + p_H + p_L) + (p_H + p_L)^2\right.\right.\right. \\ \left.\left.\left. \cdot \log\left(1 + \frac{-2dp_Hp_L}{p_H + p_L}\right)\right] \middle/ (d^2p_H^2p_L^2)\right) - 2v^2\right) \\ \qquad\qquad\qquad\qquad v > \frac{p_H + p_L}{p_H + p_L - 2dp_Lp_H}. \end{cases}$$
(A.9)

Because $\frac{-2dp_Hp_Lv}{p_H+p_L}$ (including for $v = 1$) is positive when $d < 0$, we can replace the logarithm term with $x = \frac{-2dp_Hp_Lv}{p_H+p_L}$ in the logarithm inequality (and with $v = 1$ were appropriate) and show using standard algebra that the difference is positive for the entire range of our parameters. When $d > 0$, a similar exercise shows that the difference is always negative, thus proving the first item.

To prove the second item, we find the expected profit of the global advertiser when she knows the state of the world in advance using her optimal bids found in Lemma 1. The resulting expected profit is

$$\pi^*_{Opt}$$

$$= \begin{cases} [\gamma v^2(2dp_Hp_Lv + p_H + p_L)]/(2 - 2d^2p_Hp_Lv^2) \\ \qquad v < \left(\sqrt{4d^2p_Hp_L + 4dp_H + 1} - 1\right)/[2(d^2p_Hp_L + dp_H)], \\ -\left(\left[\gamma(d^4p_H^2p_L^3v^4 + d^2p_Hp_L^2v^2(v(2dp_H(v+1)v + v + 2) - 2)\right.\right. \\ \qquad + 2p_Lv(dp_Hv(dp_Hv^2 - 1) - 1) - p_Hv^2 + p_L)]/ \\ \qquad [2(d^2p_Hp_Lv^2 - 1)^2]\right) \quad \left[\left(\sqrt{4d^2p_Hp_L + 4dp_H + 1} - 1\right)\middle/ \right. \\ \qquad [2(d^2p_Hp_L + dp_H)]\right] \le v < \left(\left[\sqrt{4d^2p_Hp_L + 4dp_L + 1} - 1\right]\middle/ \right. \\ \qquad [2(d^2p_Hp_L + dp_L)]\right], \\ \frac{1}{2}\gamma(2v(dp_Hp_L + p_H + p_L) - p_H - p_L) \\ \qquad v \ge \left(\sqrt{4d^2p_Hp_L + 4dp_L + 1} - 1\right)/[2(d^2p_Hp_L + dp_L)]. \end{cases}$$
(A.10)

Comparing the expected optimal profit to the expected profit with no attribution, we receive the condition in the proposition, thus proving the second item.

To prove the third item, we calculate the expected costs of each advertiser. We sum the costs of the local and global advertisers when they compete against one another and take the difference from the cost of two local advertisers that are competing against one another. We then integrate the difference over $v \in [0, 1]$ to calculate the expected difference in revenues of the publisher.

We perform the calculation only for the case of the publisher having visit probability p_H, because for the other advertiser the difference is just a multiple by $\frac{p_L}{p_H}$.

The resulting difference is

$$\left[\gamma\left(\left(\left[2dp_Hp_L(dp_Hp_L + p_H + p_L)(p_H^2(8d^2p_L^2 - 3)\right.\right.\right.\right. \\ \left.\left. - 6p_Hp_L - 3p_L^2\right]\middle/(-2dp_Hp_L + p_H + p_L)\right) \\ \left. - \left[3(p_H + p_L)^2(2dp_Hp_L + p_H + p_L)\right.\right. \\ \left.\left.\left. \cdot \log((-2dp_Hp_L + p_H + p_L)/(p_H + p_L))\right]\right)\right]/(24d^3p_H^2p_L^3).$$
(A.11)

When $d < 0$, the denominator is negative, and hence we can solve the inequality

$$\frac{2dp_Hp_L(dp_Hp_L + p_H + p_L)(p_H^2(8d^2p_L^2 - 3) - 6p_Hp_L - 3p_L^2)}{-2dp_Hp_L + p_H + p_L} \\ - 3(p_H + p_L)^2(2dp_Hp_L + p_H + p_L)\log\left(1 + \frac{-2dp_Hp_L}{p_H + p_L}\right) > 0.$$

Because $d < 0$, $\frac{-2dp_Hp_L}{p_H+p_L}$ is positive, and we can apply the logarithm inequality, which shows the inequality is always true.

As a result, the revenue of the publishers is lower when they sell to one global and one local advertiser compared with when they serve two local advertisers.

Proof of Proposition 2

Proving the first item is a straightforward application of the technique from Lemma 1. The results are the bidding strategies defined in the proposition.

To prove the second item, we use the equilibrium bidding strategies of the global advertiser and calculate the expected number of impressions allocated to the advertiser when $d < 0$. We denote by Imp^*_{Opt} the expected number of impressions under optimal allocation, and by Imp^*_{LT} the expected number of impressions under last-touch attribution.

The difference between the number of impressions is

$$Imp^*_{LT} - Imp^*_{Opt}$$

$$= \begin{cases} p_H(dp_Lv^2 + v - 1) + p_L(v - 1) & v(\gamma + dp_Lv(\gamma + dp_H)) \ge 1, \\ & v(\gamma + dp_Hv(\gamma + dp_L)) \ge 1, \\ [dp_Hp_Lv^2(dp_Hv + 1)(dp_Lv + 1)]/(d^2p_Hp_Lv^2 - 1) \\ & v(\gamma + dp_Lv(\gamma + dp_H)) < 1, \\ & v(\gamma + dp_Hv(\gamma + dp_L)) < 1, \\ [p_Hv(d^2p_Lv(p_H(dp_Lv^2 + v - 1) + p_Lv) - 1) + p_H]/ \\ (d^2p_Hp_Lv^2 - 1) & v(\gamma + dp_Lv(\gamma + dp_H)) \ge 1, \\ & v(\gamma + dp_Hv(\gamma + dp_L)) < 1, \\ [p_Lv(d^2p_Hv(p_L(dp_Hv^2 + v - 1) + p_Hv) - 1) + p_L]/ \\ (d^2p_Hp_Lv^2 - 1) & v(\gamma + dp_Lv(\gamma + dp_H)) < 1, \\ & v(\gamma + dp_Hv(\gamma + dp_L)) \ge 1. \end{cases}$$
(A.12)

When $d < 0$, the difference reduces to $\frac{dp_Hp_Lv^2(dp_Hv+1)(dp_Lv+1)}{d^2p_Hp_Lv^2-1}$, the denominator and numerator are both negative, and hence the difference is positive, proving the second item.

To prove the third item, we calculate the expected profit of the global advertiser under last-touch attribution:

π^*_{LT}

$$
= \begin{cases}
\frac{1}{4}\gamma v^2 (p_H(dp_L v(dv(p_H + p_L) + 4) + 2) + 2p_L) \\
\qquad\qquad v < (\sqrt{4dp_H + 1} - 1)/(2dp_H), \\[6pt]
\frac{1}{4}\gamma(p_H(v(4dp_L v + v + 2) - 1) + p_L(v(v + 2) - 1)) \\
\qquad\qquad [(\sqrt{4dp_H + 1} - 1)/(2dp_H)] \leq \\
\qquad\qquad v < [(\sqrt{4dp_L + 1} - 1)/(2dp_L)], \\[6pt]
\frac{1}{4}\gamma(d^2 p_H p_L^2 v^4 + p_L(v(2dp_H(v+1)v + v + 2) - 1) + 2p_H v^2) \\
\qquad\qquad v \geq [(\sqrt{4dp_L + 1} - 1)/(2dp_L)].
\end{cases}
$$
$$(A.13)$$

The difference $\pi^*_{LT} - \pi^*_{NA,Local}$ when $d < 0$ equals

$$
\frac{1}{4}\gamma v^2 \left(-\frac{2(p_H + p_L)^2}{-2dp_H p_L v + p_H + p_L} \right.
$$
$$
\left. + p_H(dp_L v(dv(p_H + p_L) + 4) + 2) + 2p_L \right).
$$

This difference is positive only when $p_L < \frac{p_H}{dp_H v + \sqrt{dp_H v(dp_H v + 8) + 8} + 3}$, proving the third item. The highest value of this upper bound is $1/3$, achieved at $d = -1$, $p_H = 1$, and $v = 1$.

Proof of Proposition 3

To prove the results, we first find the equilibrium bidding strategies of the local and global advertisers using a method similar to the one in Lemma 1. We find that the local advertiser will continue bidding the expected value, whereas the global advertiser will bid $-\frac{2\gamma(dp_L v^2 + 2v)}{d^2 p_H p_L v^2 - 4}$ at the popular publisher and $-\frac{2(\gamma dp_H v^2 + 2\gamma v)}{d^2 p_H p_L v^2 - 4}$ at the unpopular publisher.

Comparing these bids to the optimal bids found in Lemma 1, we find that the bid at the unpopular publisher is always above the optimal bid when using Shapley value attribution. The bid at the popular advertiser is lower than the optimal bid when $v > \frac{\sqrt{9 - \frac{8p_L}{p_H}} - 3}{2dp_L}$ and $p_L < 3 - \frac{2}{p_H}$, and higher otherwise.

The expected profit under Shapley value attribution is

$$
\pi^*_{Sh} = \begin{cases}
\frac{1}{2}\gamma(2v(dp_H p_L + p_H + p_L) - p_H - p_L) \\
\qquad v \leq (2(\sqrt{d^2 p_H p_L + 2dp_L + 1} - 1))/(d^2 p_H p_L + 2dp_L), \\[6pt]
\frac{1}{2}\gamma(-(4p_H v^2(dp_L + 1)(dp_L v + 2))/(d^2 p_H p_L v^2 - 4) \\
\qquad - (4p_H v^2(dp_L v + 2)^2)/((d^2 p_H p_L v^2 - 4)^2) + p_L(2v - 1)) \\
\qquad (2(\sqrt{d^2 p_H p_L + 2dp_L + 1} - 1))/(d^2 p_H p_L + 2dp_L) \\
\qquad < v \leq (2(\sqrt{d^2 p_H p_L + 2dp_H + 1} - 1))/(d^2 p_H p_L + 2dp_H), \\[6pt]
(2\gamma v^2(p_H(dp_L v(dv(p_H + p_L) + 8) + 4) + 4p_L))/ \\
\qquad (d^2 p_H p_L v^2 - 4)^2 \\
\qquad v > (2(\sqrt{d^2 p_H p_L + 2dp_H + 1} - 1))/(d^2 p_H p_L + 2dp_H).
\end{cases}
$$
$$(A.14)$$

To compare with the profits with no attribution and last-touch attribution, we limit to the cases of $-1 < d < 0$, and calculate

$$
\pi^*_{LT} - \pi^*_{NA} = \frac{1}{4}\gamma v^2 \left(\frac{4(p_H + p_L)^2}{4dp_H p_L v - 2(p_H + p_L)} \right.
$$
$$
\left. + p_H(dp_L v(dv(p_H + p_L) + 4) + 2) + 2p_L \right), \quad (A.15)
$$

$$
\pi^*_{LT} - \pi^*_{Sh} = \frac{1}{4}\gamma v^2 \left(-\frac{8(p_H(dp_L v(dv(p_H + p_L) + 8) + 4) + 4p_L)}{(d^2 p_H p_L v^2 - 4)^2} \right.
$$
$$
\left. + p_H(dp_L v(dv(p_H + p_L) + 4) + 2) + 2p_L \right), \quad (A.16)
$$

$$
\pi^*_{Sh} - \pi^*_{NA} = \gamma v^2 \left(\frac{2(p_H(dp_L v(dv(p_H + p_L) + 8) + 4) + 4p_L)}{(d^2 p_H p_L v^2 - 4)^2} \right.
$$
$$
\left. + \frac{(p_H + p_L)^2}{4dp_H p_L v - 2(p_H + p_L)} \right). \quad (A.17)
$$

To prove the second item, we solve the inequality $\pi^*_{Sh} - \pi^*_{NA} > 0$, which holds when $v < \sqrt{3} - 1$ or when $p_L < \tilde{p}_L$ and $\tilde{p}_L$ is the solution to

$$
\tilde{p}_L^3 d^4 p_H^2 v^4 + \tilde{p}_L^2 (d^4 p_H^3 v^4 + 4d^3 p_H^2 v^3 - 6d^2 p_H v^2)
$$
$$
+ \tilde{p}_L(-6d^2 p_H^2 v^2 - 32dp_H v - 8) - 8p_H = 0.
$$

The third item is solved in a similar way, solving the two inequalities $\pi^*_{Sh} - \pi^*_{NA} > 0$ and $\pi^*_{LT} - \pi^*_{Sh} < 0$. These inequalities hold when $p_L < p_L^c$ and p_L^c is the solution to

$$
(p_L^c)^3 d^2 p_H v^2 + (p_L^c)^2 (2d^2 p_H^2 v^2 + 8dp_H v - 12)
$$
$$
+ p_L^c (d^2 p_H^3 v^2 + 8dp_H^2 v + 40p_H) - 12p_H^2 = 0.
$$

The fourth item is a simple application of the same approach with the inequalities $\pi^*_{LT} - \pi^*_{NA} > 0$ and $\pi^*_{LT} - \pi^*_{Sh} > 0$.

Proof of Lemma 2

There are four possible allocations when the consumer visits both publishers, which we label GG, $1G$, $G2$, and 12. Outcome ij means the ad was allocated to advertiser i at publisher 1 and to advertiser j at publisher 2. There are also two possible outcomes when the consumer visits only one of the publishers, either 1 or G at publisher 1 or 2 or G at publisher 2.

Each of these outcomes generates specific revenue. For example, outcome GG generates $\gamma v_G + \gamma(1 + d)v_G$. We solved four sets of inequalities when the consumers visits both publishers to find the range in which each outcome generates the maximum revenue. Integrating over v_G, v_1, and v_2 when the consumer visits both publishers yields $\gamma \frac{1}{12}(d^3 + 3d^2 + 3d + 16)p_1 p_2$ when p_j is the probability of a visit at publisher j. Similarly, integrating when the consumer visits only one publisher yields $\frac{2}{3}(p_2 - p_1(2p_2 - 1))$. Summing these two expressions yields the result in the lemma. A simple comparative statics analysis shows the total revenue increases in all three parameters.

Ratios Used in Section 5.3

For completeness, we detail the ratios used for the numerical analysis of Section 5.3.

These ratios were all calculated using the expected profit expressions found in the previous proofs using the constraint that $d < 0$. It is interesting to note that all expected profits are linear in γ, and hence the value of γ does not affect these ratios. The ratios are:

$$\frac{\pi_{Sh}^*}{\pi_{Opt}^*} = \frac{2\,(2 - 2d^2 p_H p_L v^2)\,(p_H(dp_L v(dv(p_H + p_L) + 8) + 4) + 4p_L)}{(2dp_H p_L v + p_H + p_L)\,(d^2 p_H p_L v^2 - 4)^2},$$

(A.18)

$$\frac{\pi_{LT}^*}{\pi_{Opt}^*} = \frac{(2 - 2d^2 p_H p_L v^2)\,(p_H(dp_L v(dv(p_H + p_L) + 4) + 2) + 2p_L)}{4(2dp_H p_L v + p_H + p_L)},$$

(A.19)

$$\frac{\pi_{NA}^*}{\pi_{Opt}^*} = \frac{(p_H + p_L)^2\,(d^2 p_H p_L v^2 - 1)}{4d^2 p_H^2 p_L^2 v^2 - (p_H + p_L)^2}.$$

(A.20)

For the unconditional ratios, we find the following expressions:

$$\left(\int_0^1 \pi_{Sh}^* \, dv\right)\bigg/\left(\int_0^1 \pi_{Opt}^* \, dv\right) = (4\sqrt{p_H p_L}(d^3 p_H p_L(p_H + p_L)$$
$$- 4d^2 p_H p_L(1 + \log(4)) - 8d(p_H + p_L) + 16\log(4)))/$$
$$((d^2 p_H p_L - 4)((p_H + p_L)\tanh^{-1}(d\sqrt{p_H p_L})$$
$$- \sqrt{p_H p_L}(\log(1 - d^2 p_H p_L) + d(dp_H p_L + p_H + p_L)))) \quad \text{(A.21)}$$

$$- (16(d^2 p_H p_L - 4)((p_H + p_L)\tanh^{-1}(\tfrac{1}{2}d\sqrt{p_H p_L})$$
$$- \sqrt{p_H p_L}\log(4 - d^2 p_H p_L)))/((d^2 p_H p_L - 4)((p_H$$
$$+ p_L)\tanh^{-1}(d\sqrt{p_H p_L}) - \sqrt{p_H p_L}(\log(1 - d^2 p_H p_L)$$
$$+ d(dp_H p_L + p_H + p_L)))), \quad \text{(A.22)}$$

$$\left(\int_0^1 \pi_{LT}^* dv\right)\bigg/\left(\int_0^1 \pi_{Opt}^* dv\right)$$
$$= (d^3(p_H p_L)^{3/2}(p_H(3dp_L(d(p_H + p_L) + 5) + 10) + 10p_L))/$$
$$(30((p_H + p_L)\tanh^{-1}(d\sqrt{p_H p_L}) - \sqrt{p_H p_L}(\log(1 - d^2 p_H p_L)$$
$$+ d(dp_H p_L + p_H + p_L)))), \quad \text{(A.23)}$$

$$\left(\int_0^1 \pi_{NA}^* dv\right)\bigg/\left(\int_0^1 \pi_{Opt}^* dv\right)$$
$$= \Bigg(2dp_H p_L(p_H + p_L)^2(dp_H p_L + p_H + p_L) + (p_H + p_L)^4$$
$$\log\left(\frac{-2dp_H p_L + p_H + p_L}{p_H + p_L}\right)\Bigg)\bigg/\Big(8p_H^2 p_L^2(\log(1 - d^2 p_H p_L)$$
$$+ d(dp_H p_L + p_H + p_L)) - 8(p_H p_L)^{3/2}$$
$$(p_H + p_L)\tanh^{-1}(d\sqrt{p_H p_L})\Big). \quad \text{(A.24)}$$

Proof of Proposition 4

The analysis shows that the derivative of each expected profit function found in the previous proofs with respect γ, v, d, and p_H when $d < 0$ is positive, proving the first item.

The second item holds when

$$v < \frac{\sqrt{2 - \frac{4p_L}{p_H}} - 2}{d(p_H + 2p_L)}$$

if using last-touch attribution, when

$$v < -\frac{p_H + p_L}{2dp_H(p_H - p_L)}$$

if not using attribution, and when $v > v^c$ using the Shapley value when v^c solves

$$(v^c)^3 d^3 p_H^2 p_L + 6(v^c)^2 d^2 p_H p_L + 12(v^c)dp_H + 8 = 0.$$

Proof of Proposition 5

To prove all items, suppose that the total conversion rate $\gamma + \gamma(1 + d) = c$ when c is constant. Then using the implicit function theorem, $\frac{dd}{d\gamma} = -\frac{2+d}{\gamma}$.

If the difference of expected profit is Δ^*, then we can calculate

$$\frac{d\Delta^*}{d\gamma} = \frac{\partial\Delta^*}{\partial\gamma} + \frac{\partial\Delta^*}{\partial d}\frac{dd}{d\gamma}$$

using the envelope theorem.

Performing this analysis shows that $\frac{d}{d\gamma}(\pi_{Opt}^* - \pi_{Sh}^*) > 0$ and $\frac{d}{d\gamma}(\pi_{Opt}^* - \pi_{LT}^*) > 0$ for all parameters values.

We have $\frac{d}{d\gamma}(\pi_{Opt}^* - \pi_{LT}^*) < 0$ when $v > \frac{\sqrt{10}-1}{3}$, $d < \tilde{d}$, $p_L > \tilde{p}_L$, and $p_H > \tilde{p}_H$, and the cutoffs are defined as the solutions to

$$\tilde{d}^3 v^2 + \tilde{d}^2(4v^2 - 2v) + \tilde{d}(-4v - 1) - 4 = 0, \quad \text{(A.25)}$$

$$\tilde{p}_H = \frac{d + 2 - \sqrt{2}\sqrt{d^2 + 6d + 10}}{d^2 v + 4dv}, \quad \text{(A.26)}$$

$$\tilde{p}_L^3(d^5 p_H^2 v^4 + 4d^4 p_H^2 v^4)$$
$$+ \tilde{p}_L^2(d^5 p_H^3 v^4 + 4d^4 p_H^3 v^4 + 8d^3 p_H^2 v^3 - 4d^3 p_H v^2 - 8d^2 p_H v^2)$$
$$+ \tilde{p}_L(-4d^3 p_H^2 v^2 - 8d^2 p_H^2 v^2 - 8d^2 p_H v - 24dp_H v - d - 4)$$
$$- dp_H - 4p_H = 0. \quad \text{(A.27)}$$

This proves the first two items.

To prove the last two items, we use the same approach, but instead of applying it on the difference from the optimal profit, we apply it to the attributed revenue of each publisher in equilibrium.

We find that $\frac{d}{d\gamma}(A_{1,LT}^*) > 0$ always, whereas $\frac{d}{d\gamma}(A_{2,LT}^*) > 0$ only when $v < \frac{1}{p_H(d+4)}$.

Similarly, we find that $\frac{d}{d\gamma}(A_{1,Sh}^*) < 0$ and $\frac{d}{d\gamma}(A_{2,Sh}^*) < 0$ when $p_L v > \frac{2}{(3d+4)}$ and

$$8/(v(d((5d+8)p_L v + 2)$$
$$+ \sqrt{d(d(p_L v((5d+8)^2 p_L v + 44d + 144) + 4) + 32(4p_L v + 1)) + 64}$$
$$+ 8) < p_H < ((8 - 4(d+4)p_L v)/(dp_L v^2(d((3d+4)p_L v + 10) + 16))).$$

Proof of Proposition 6

To prove the first item, we prove that if the bids of an advertiser under one attribution scheme are always higher than those under another scheme in a specific state, then the revenue of the publishers is higher under the former scheme.

Suppose that the global advertiser bids $b_1(v_G)$ under one attribution scheme in a specific state with visit probability p and $b_2(v_G)$ under another scheme such that $b_1(v_G) > b_2(v_G)$, and that the local bidder bids her equilibrium strategy γv.

Then, the expected difference in revenue of a publisher in a specific state is

$$p\left(\int_0^1\int_0^1\mathbb{I}(\gamma v\geq b_1(v_G))b_1(v_G)dv_G dv\right.$$

$$+\int_0^1\int_0^1\mathbb{I}(\gamma v<b_1(v_G))\gamma v\,dv\,dv_G$$

$$-\int_0^1\int_0^1\mathbb{I}(\gamma v\geq b_2(v_G))b_2(v_G)dv_G dv$$

$$\left.+\int_0^1\int_0^1\mathbb{I}(\gamma v<b_2(v_G))\gamma v\,dv\,dv_G\right). \qquad (A.28)$$

We can change the order of integration and collect the integrals to

$$p\left(\int_0^1\int_0^1(\mathbb{I}(\gamma v\geq b_1(v_G))(b_1(v_G)-\gamma v)\right.$$

$$\left.-\mathbb{I}(\gamma v\geq b_2(v_G))(b_2(v_G)-\gamma v))\,dv_G dv\right). \qquad (A.29)$$

The integrand can be written as

$$\mathbb{I}(\gamma v\geq b_2(v_G))(\gamma v-b_2(v_G))-\mathbb{I}(\gamma v\geq b_1(v_G))(\gamma v-b_1(v_G)). \qquad (A.30)$$

Because $b_1(v_G)>b_2(v_G)$ for every v_G, the difference is always positive, proving that the expected revenue of the publisher is higher under the first scheme.

To prove the first item of the proposition, it is therefore enough to show that $b_{1,Sh}^*(1H)<b_{1,LT}^*(1H)$ and $b_{1,Sh}^*(1L)<b_{1,LT}^*(1L)$ for the first publisher, and $b_{2,Sh}^*(1H)>b_{2,LT}^*(1H)$ and $b_{2,Sh}^*(1L)>b_{2,LT}^*(1L)$ for the second. These inequalities hold for every value of v when $d<0$.

To prove the second item, we calculate the cost of advertising in each attribution scheme and state and the number of impressions in each scheme and state, and find the price as the ratio of those. The expected value over states yields

$$\mathbb{E}[price_{LT}]=(\gamma v(d^2p_H^2p_Lv^2+p_H(d^2p_L^2v^2+4dp_Lv+2)$$

$$+2p_L))/(8(dp_Hp_Lv+p_H+p_L)); \qquad (A.31)$$

$\mathbb{E}[price_{Sh}]$

$$=\begin{cases}-\left(\gamma v(d^2p_H^2p_Lv^2+p_H(d^2p_L^2v^2+8dp_Lv+4)+4p_L)\right)/\\ \quad(4(dp_Hp_Lv+p_H+p_L)(d^2p_Hp_Lv^2-4))\\ \qquad\qquad v\left(4\gamma+d^2p_Hp_Lv+2\gamma dp_Hv\right)<4,\\ \qquad\qquad v\left(4\gamma+d^2p_Hp_Lv+2\gamma dp_Lv\right)<4,\\ \left(\gamma v^2(d^2p_H^2p_Lv^2+p_H(d^2p_L^2v^2+8dp_Lv+4)+4p_L)\right)/\\ \quad((p_H+p_L)(d^2p_Hp_Lv^2-4)^2)\\ \qquad\qquad v\left(4\gamma+d^2p_Hp_Lv+2\gamma dp_Hv\right)\geq4,\\ \qquad\qquad v\left(4\gamma+d^2p_Hp_Lv+2\gamma dp_Lv\right)\geq4,\\ \left(\gamma v^2(d^2p_H^2p_Lv^2+p_H(d^2p_L^2v^2+8dp_Lv+4)+4p_L)\right)/\\ \quad((d^2p_Hp_Lv^2-4)(d^2p_H^2p_Lv^2-2p_H(dp_Lv^2+2)-4p_Lv))\\ \qquad\qquad v\left(4\gamma+d^2p_Hp_Lv+2\gamma dp_Hv\right)<4,\\ \qquad\qquad v\left(4\gamma+d^2p_Hp_Lv+2\gamma dp_Lv\right)\geq4,\\ \left(\gamma v^2(d^2p_H^2p_Lv^2+p_H(d^2p_L^2v^2+8dp_Lv+4)+4p_L)\right)/\\ \quad((d^2p_Hp_Lv^2-4)(d^2p_Hp_L^2v^2-2p_L(dp_Hv^2+2)-4p_Hv))\\ \qquad\qquad v\left(4\gamma+d^2p_Hp_Lv+2\gamma dp_Hv\right)\geq4,\\ \qquad\qquad v\left(4\gamma+d^2p_Hp_Lv+2\gamma dp_Lv\right)<4;\end{cases} \qquad (A.32)$$

$$\mathbb{E}[price_{NA}]=\begin{cases}(\gamma v^2(p_H+p_L)^2)/(4(-2dp_Hp_Lv+p_H+p_L)^2)\\ \qquad v(2dp_Hp_L+\gamma(p_H+p_L))\geq p_H+p_L,\\ (\gamma v(p_H+p_L))/(4(-2dp_Hp_Lv+p_H+p_L))\\ \qquad v(2dp_Hp_L+\gamma(p_H+p_L))<p_H+p_L.\end{cases} \qquad (A.33)$$

The complexity of the expressions requires numerical analysis. We calculated the maxima of the ratios $\frac{\mathbb{E}[price_{Sh}]}{\mathbb{E}[price_{LT}]}$ and $\frac{\mathbb{E}[price_{NA}]}{\mathbb{E}[price_{Sh}]}$. Both of these ratios equal 1 at the maximum, which proves the second item.

Endnotes

[1] Source: IAB Revenue Report 2015. Accessed July 18, 2018, https://www.iab.com/wp-content/uploads/2016/04/IAB_Internet_Advertising_Revenue_Report_FY_2015-final.pdf; Alphabet 2015 Annual Report. Accessed July 18, 2018, https://www.sec.gov/Archives/edgar/data/1288776/000165204416000012/goog10-k2015.htm; and Facebook 2015 Annual Report. Accessed July 18, 2018, https://www.sec.gov/Archives/edgar/data/1326801/000132680116000043/fb-12312015x10k.htm.

[2] The online appendix analyzes the case of two global advertisers.

[3] The online appendix analyzes the effect of CPC auctions with bid adjustment.

[4] Conversions may take many forms. Facebook allows for multiple campaign objectives such as accumulating page "likes," generating website visits through ad clicks, and generating video views. Google allows bidding on clicks.

[5] We limit $d<1$ to avoid extreme cases where the advertiser would like to bid infinitely high for an impression.

[6] Google states, "the most you'll pay is what's minimally required to hold your ad position" (http://support.google.com/adwords/answer/6297?hl=en, accessed July 17, 2018). Facebook states, "The amount you get charged is the minimum amount you would've needed to set your advertiser bid at to win the auction" (Facebook (2018). Accessed July 18, 2018, http://www.facebook.com/business/help/430291176997542).

[7] Advertisers can learn the value of a conversion through running an initial experiment prior to the campaign.

[8] If the advertiser is budget constrained, this would be equal to setting the marginal ROIs to being equal between the two publishers.

[9] See Google. Accessed July 18, 2018, https://support.google.com/analytics/answer/3191594?hl=en.

[10] The details appear in the proof of Lemma 2.

[11] We thank the editor and associate editor for suggesting this analysis.

References

Abhishek V, Fader P, Hosanagar K (2012) Media exposure through the funnel: A model of multi-stage attribution. Working paper, Carnegie Mellon University, Pittsburgh.

Agarwal N, Athey S, Yang D (2009) Skewed bidding in pay-per-action auctions for online advertising. *Amer. Econom. Rev.* 99(2): 441–447.

Anderl E, Becker I, Von Wangenheim F, Schumann JH (2016) Mapping the customer journey: Lessons learned from graph-based online attribution modeling. *Internat. J. Res. Marketing* 33(3):457–474.

Ausubel LM, Baranov OV (2013) Core-selecting auctions with incomplete information. Working paper, University of Maryland, College Park.

Bager A, Laszlo J (2015) 2016 is the year of attribution. Accessed July 18, 2018, https://www.clickz.com/2016-is-the-year-of-attribution/89517/.

Barajas J, Akella R, Holtan M, Flores A (2016) Experimental designs and estimation for online display advertising attribution in marketplaces. *Marketing Sci.* 35(3):465–483.

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

Borgs C, Chayes J, Immorlica N, Jain K, Etesami O, Mahdian M (2007) Dynamics of bid optimization in online advertisement auctions. *Proc 16th Internat. Conf. World Wide Web* (ACM, New York), 531–540.

Bruell A, Terlep S (2017) P&G cuts more than $100 million in 'largely ineffective' digital ads. *Wall Street Journal* (July 27).

Dalessandro B, Perlich C, Stitelman O, Provost F (2012) Causally motivated attribution for online advertising. *Proc. Sixth Intern. Workshop Data Mining Online Advertising Internet Econom.*, Vol. 7 (ACM, New York), 1–9.

D'Annunzio A, Russo A (2017) Ad networks, consumer tracking, and privacy. Working paper, University of Naples Federico II - CSEF, Naples, Italy.

Econsultancy (2012) Marketing attribution: Valuing the customer journey. Accessed February 1, 2012, https://www.thinkwithgoogle .com/_qs/documents/703/marketing-attribution-valuing-the -customer-journey_research-studies.pdf.

Facebook (2017) Your guide to Facebook bid strategy. Accessed July 17, 2018, https://go.fb.com/rs/267-PVB-941/images/ BiddingStrategyGuide_FINAL.pdf.

Geyik SC, Saxena A, Dasdan A (2014) Multi-touch attribution based budget allocation in online advertising. *Proc. Eighth Internat. Workshop Data Mining Online Advertising* (ACM, New York), 1–9.

Holmstrom B (1982) Moral hazard in teams. *Bell J. Econom.* 13(2):324–340.

Hu Y, Shin J, Tang Z (2015) 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, Boston University, Boston.

Jordan P, Mahdian M, Vassilvitskii S, Vee E (2011) The multiple attribution problem in pay-per-conversion advertising. *Algorithmic Game Theory* (Springer, Berlin), 31–43.

Kempe D, Mahdian M (2008) A cascade model for externalities in sponsored search. Papadimitriou C, Zhang S, eds. *Internet and Network Economics*. Lecture Notes in Computer Science, Vol. 5385 (Springer, Berlin), 585–596.

Kireyev P, Pauwels K, Gupta S (2016) Do display ads influence search? Attribution and dynamics in online advertising. *Internat. J. Res. Marketing* 33(3):475–490.

Krishna V (2009) *Auction Theory* (Academic Press, Burlington, MA).

Krishna V, Rosenthal RW (1996) Simultaneous auctions with synergies. *Games Econom. Behav.* 17(1):1–31.

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.

Li H, Kannan PK (2014) Attributing conversions in a multichannel online marketing environment: An empirical model and a field experiment. *J. Marketing Res.* 51(1):40–56.

Li HA, Kannan PK, Viswanathan S, Pani A (2016) Attribution strategies and return on keyword investment in paid search advertising. *Marketing Sci.*, 35(6):831–848.

Mayzlin D, Dover Y, Chevalier J (2014) Promotional reviews: An empirical investigation of online review manipulation. *Amer. Econom. Rev.* 104(8):2421–2455.

Shao X, Li L (2011) Data-driven multi-touch attribution models. *Proc. 17th ACM SIGKDD Internat. Conf. Knowledge Discovery Data Mining* (ACM, New York), 258–264.

Shapley LS (1952) *A Value for n-Person Games* (RAND Corporation, Santa Monica, CA).

Tucker C (2012) Implications of improved attribution and measurability for antitrust and privacy in online advertising markets. *George Mason Rev.* 20(4):1025–1054.

Wilbur KC, Zhu Yi (2009) Click fraud. *Marketing Sci.* 28(2):293–308.

Young HP (1988) Individual contribution and just compensation. *The Shapley Value: Essays in Honor of Lloyd S. Shapley* (Cambridge University Press, New York), 267–278.

Zantedeschi D, McDonnell Feit E, Bradlow E (2017) Measuring multi-channel advertising response. *Management Sci.* 63(8): 2706–2728.

Zhu Y, Wilbur KC (2011) Hybrid advertising auctions. *Marketing Sci.* 30(2):249–273.