The Attention Economy of Online Advertising*

Alexander White† Kamal Jain‡

July 23, 2019

Abstract

Internet users often surf to multiple websites in order to accomplish a single task. When this happens, do these different sites face the right incentives when choosing their advertising policies? We build a model showing that websites face an interesting tradeoff: on the one hand, they are prone to over-advertise (similar to double marginalization); on the other hand, they tend to misallocate ads across sites (a distortion we call *misplacement*). Standard solutions to the double marginalization problem, such as adding competition among certain sites, make the misplacement problem more severe. This tradeoff is important for news aggregators and social networks, as it affects their decisions whether to link to external content providers or to expand the amount of content they offer by themselves. Understanding these incentives helps to inform the current debate regarding the concentration of influence among a small set of online platforms.

Keywords: Platforms, Advertising, Misplacement, Market Power

JEL Codes: D21, D40, L23, L42, L86

*We thank seminar audiences at the U.S. Federal Communications Commission, the University of Virginia, the University of Miami, Peking University and Hong Kong University as well as Simon Anderson, Gary Biglaiser, Marc Bourreau, Denis Charles, Jacques Crémer, Eric Horvitz, Bruno Jullien, Chris Meek, Martin Peitz, Paul Seabright, Sven Seuken, Wing Suen, Thomas Tregouet, Glen Weyl and Mike Whinston for comments/discussions, Yunhao Huang, Keyan Li and Lingxuan "Sean" Wu for exceptional research assistance. White acknowledges financial support for this research from the *Chaire Orange de l'innovation & régulation*. The authors are solely responsible for any errors.

†School of Economics and Management and National Institute for Fiscal Studies, Tsinghua University, Beijing, China; awhite@sem.tsinghua.edu.cn

‡Faira.com, Kirkland, Washington, U.S.A.

1 Introduction

Consider the top two online "US Media Publications by Pageviews, May 2017" as reported by SimilarWeb: MSN and the Drudge Report.[1] Both of these websites aggregate news and other content, attracting users to their main page and then directing them to other pages containing information of specific, current interest, generated by other parties. Interestingly, links found on these two websites differ in a crucial way. On the one hand, clicking on a Drudge Report link takes users to an external producer's website. On the other hand, MSN links takes users to another page containing externally produced content that is still hosted on the MSN.com server. As Microsoft's September 2014 blog post[2] announced, "the new MSN" features pages with content drawn from partners such as the *New York Times*, *Wall Street Journal*, *Washington Post*, *CNN*, *AOL*, and *Condé Nast*.

More generally, an aspect of Internet usage that seems almost too obvious to mention is that people *surf*. That is, they go from one page to another in a single browsing session. In doing so, they often seek to fulfill a single desire or "need". One such example is a user who starts off by visiting one of the sites mentioned above in order to read the latest news. Other examples include navigating first to a social media platform like Facebook or Twitter and using it to find a link to relevant content on a topic of interest. In other instances, even without navigating from one page to another, users can be simultaneously exposed to the content from multiple sites appearing within the same page: for instance, a blog may contain content written by its creator while also featuring an embedded video hosted by YouTube.

Such scenarios are all likely to share two crucial features. First, separate websites play complementary roles in fulfilling users' needs. In particular, some sites play the role of *organizers* to which users navigate directly. They, in turn, link to *producers* whose content users deem relevant but difficult to find directly.[3] Second, the websites in question are ad-funded and do not charge users money on a per-visit basis. This configuration raises a set of interesting questions regarding, in particular:

- the efficiency of sites' advertising practices (Should we expect the sites linked together into an

[1] See https://www.similarweb.com/blog/us-media-publications-ranking-may-2017.

[2] Available at https://blogs.microsoft.com/blog/2014/09/29/new-msn-arrives.

[3] For example, the *Guardian* reports that, of all the web traffic it received viewing a recent "bombshell" story, one quarter was referred by the Drudge Report (Pilkington, 2018).

organizer-producer "chain" to have well-coordinated advertising policies?);

- the effect of competition among potential rival producers in such a chain (Do the total profits earned by all sites in a chain increase when market power is consolidated into the hands of a single organizer? Would such an arrangement be optimal for the organizer in question?);

- the incentives for a given site to attempt to satisfy users' needs entirely "in house" (What underlying forces can explain the differences in the strategies adopted by MSN and the Drudge Report, described above? Under what circumstances is an organizer better off linking externally to producers and when should it acquire content that it hosts on its own site?).

This paper addresses these questions. We begin by analyzing a straightforward yet quite general model in which users must visit multiple websites in order to fulfill their particular need. Here, we compare the amount of advertising that each website shows, at equilibrium, with the amount that each would show at the websites' combined optimum. Equilibrium exhibits not only the advertising equivalent of double marginalization, whereby the total level of nuisance imposed on users exceeds the level that an integrated website would choose. Even holding fixed this total level of user distraction, the websites fail to attain maximum potential profits.

This latter inefficiency is what we call *misplacement*. It stems from the fact that, as the literature on online advertising has discussed (Evans, 2008, 2009; Goldfarb and Tucker, 2011; Athey et al., 2018), websites seem to differ significantly from one another in their ability to transform user attention into advertising dollars. In other words, websites are very heterogenous in the effectiveness of their advertising technologies. Consequently, at equilibrium, different complementary websites may differ from one another in the revenue they can earn by showing "one more ad." However, the websites' respective ad policies must equalize these two rates if their total profits are to be maximized. Thus, misplacement is a distinct coordination failure from double marginalization.

Moreover, to the best of our knowledge, misplacement has not been discussed in previous literature. We compare our setting, specifically, to those involving (i) platforms that show ads and charge money, (ii) vertical cooperative advertising, and (iii) Cournot substitute competition among asymmetric-cost firms. We show that, although these settings can have allocative distortions, none has the two key ingredients: critically, misplacement stems from the *combination* of sites' complementarity and the fact that they are offered to users for "free", in exchange for showing ads.

Importantly, misplacement leads to novel implications regarding the effects of competition. First, consider the question of whether, from the standpoint of an organizer-producer chain as a whole, introducing competition among different producers could overcome the two distortions. In a standard vertically-arranged market, in which two firms each *sell* necessary components that, together, constitute a final good, an effective and well-understood way to eliminate double marginalization is to introduce competition in the market for one of the two components (Rey and Tirole, 1986; Shleifer and Vishny, 1993; Lerner and Tirole, 2004, 2015; Dellarocas, 2012). For instance, suppose that the final good of "IT solutions" consisted of both "hardware" and "software." If a single hardware producer and a single software producer were each allowed to set prices independently, then the total price for the two would be higher than the level that maximizes total IT industry profits. On the other hand, if intense competition were introduced into the hardware industry, then the price for hardware will fall to its marginal production cost. In response, the single software seller then faces efficient pricing incentives: it will optimally set a markup equal to the one that a single, integrated firm would charge for the bundle, thereby maximizing total industry profits.

In contrast, with complementary, ad-funded websites, an interesting tradeoff arises under analogous circumstances. Returning to our initial example, suppose that, instead of featuring one prominent link to a story on the desired topic, the organizer consistently featured side-by-side links to several similar producers' version of roughly the same story. Faced with more competition, the producers would have an incentive to reduce advertisement on their sites.

However, although the organizer would become the only site with market power (here, that's the ability to show a marginally greater number of ads without prompting a steep drop-off in traffic) it would not be able to attain the industry's maximal level of potential profits. Either of two factors is sufficient to drive this result. First, it can arise if, unlike money spent on two goods in a bundle, distraction from advertisement is not perfectly fungible across different sites.[4] Second, it can arise if there is asymmetry in the demand among advertisers to reach users via one site rather than another. Indeed, as the market power of producers shrinks, misplacement tends to become worse. Thus,

[4]Whereas a consumer asked to pay $\alpha \cdot \$100$ for hardware and $(1-\alpha) \cdot \$100$ for software is indifferent regarding the value of α, the share of the total price paid to the hardware seller, a web user is *not* necessarily indifferent, in an analogous way, to the placement of ads across websites. Our basic assumption is that distraction from advertisement on a given site increases in a convex fashion.

as a whole, the organizer-producer chain faces a tradeoff between fostering competition among producers, in order to reduce double marginalization, and allowing them to have market power, in order to reduce misplacement.

Now consider the incentives of the organizer. In the most straightforward version of our setting, the organizer always benefits as competition among producers increases. Furthermore, it could earn even greater profits if it simply integrated vertically with a producer and took control of advertisement throughout the chain. However, these issue becomes more interesting when considered in the context of the MSN/Drudge comparison highlighted above, which suggests that, depending on the circumstances, integration of this form may or may not be optimal. To study this, we enrich the model to allow the organizer to make a choice of whether to operate in *linking* mode, as described thus far (similar to Drudge), or *acquisition* mode, in which it forms content partnerships with producers and hosts their content on its website (similar to MSN).

First, in the simplest version of this setting, we illustrate the basic tradeoff governing the organizer's choice between these two modes. The key point is that, although acquisition mode allows the organizer full control of advertisement, linking mode gives it greater flexibility to direct users to whichever producer happens to generate the most relevant content. This flexibility boosts users' enthusiasm for surfing and thus their willingness to tolerate ads shown by the organizer. Consequently, despite the fact that linking mode leads to misplacement and double marginalization, the organizer prefers linking mode when (i) the number of producers that could potentially obtain a "scoop" (e.g., exclusive, relevant content) is large and (ii) the transaction costs incurred and payments made when forming content partnerships are not too small. If, on the other hand, the organizer can, at a low cost, agree in advance to form content partnerships with enough of the potentially relevant producers, then it chooses acquisition mode.

Next, we uncover a subtler factor influencing this decision: the role of the correlation between different producers' obtaining scoops. Suppose there are two potentially relevant producers and that the organizer can form a content partnership with at most one of them. Now, consider the following two extreme cases. Under "full exclusivity", when a scoop is generated (that is, when some relevant piece of content comes into the picture), it falls into the hands of only one producer. On the other hand, under "no exclusivity", when a scoop arrives, both producers always obtain it.

Between these two extreme cases lies a continuum of correlation levels. When this correlation is low, conditional on one producer's obtaining a scoop, the other is unlikely to also obtain it. As the correlation grows, the other producer becomes more likely to also obtain the scoop. Intuitively, full exclusivity corresponds to the most monopolistic setting for the producer with the scoop. As the correlation grows, so does the intensity of competition.

We show that it is under *intermediate* levels of producer competition that linking mode is most favorable in the eyes of the organizer. This is a consequence of the tradeoff between misplacement and double marginalization, which, so far, we have described as a tension facing total profits, but which, here, has a direct impact on the organizer. On the one hand, when producer competition is very low, the loss stemming from double marginalization reaches its peak. On the other hand, when producer competition is very intense, the loss stemming from misplacement reaches its peak. Thus, the gain for the organizer from switching to acquisition mode and controlling advertisement throughout the chain is especially large in these two cases. At intermediate levels of competition, the producers' incentives to undercut each other endogenously come closer to solving the tradeoff, thereby lessening the potential gains to the organizer from taking control of advertisement throughout the chain.

Of course, this effect coexists with the simpler tradeoff identified above between flexibility and the cost of acquiring content. In the case of no exclusivity, linking mode delivers no flexibility benefit, because acquiring the content of just one producer is sufficient to guarantee that the organizer obtains the scoop. Meanwhile, it is under full exclusivity that the organizer benefits the most from linking mode's flexibility, because the chances that, under acquisition mode, its content partner would obtain the scoop are the lowest.

After exploring these issues, we enrich the model in two ways in order to show the robustness of our basic approach. In one extension, we use a two-sided markets approach, building on the classic paper by Anderson and Coate (2005), in order to explicitly model advertisers. Doing so allows us to make the further point that misplacement need not be caused by exogenous differences in the complementary websites' advertising technologies. Even when the complementary sites' advertising technologies are symmetric, misplacement arises when they face different demand from advertisers. Within the two-sided framework, we also address whether the complementary websites can mitigate the distortions by delegating control of their ad policies to an agency that acts on the collective behalf of advertisers. It turns out that ceding control in this way can either increase or decrease total website profits, but it does not systematically reduce the occurrence of either distortion.

In a final extension, we model, in a more detailed way, the process whereby the sites set ad levels and users decide whether to surf. Here, we take into account the sequential nature of our story and users' inability to observe how many ads a site has before they decide to visit it. We show that, under this specification, the two distortions on which we focus continue to be relevant.

Related Literature

The problem of double marginalization, first pointed out by Cournot (1838), is very well-studied in the context of complementary goods.[5] In contrast, we believe that misplacement, which leads to our central tradeoff, has not been discussed in any prior literature and is especially relevant in settings where multiple firms can demand attention from consumers who are trying to accomplish a single task. As we discuss in Subsection 2.2, the closest analog to misplacement of which we are aware arises in models, such as that of Schwartz (1989), where imperfectly competing sellers of substitutes have asymmetric marginal costs. Dellarocas's (2012) model resembles ours in that it also studies double marginalization in online advertising. However, the two works differ substantially, as that one studies the double markup problem in product pricing, whereas ours focus on the interaction between misplacement of ads and double marginalization in the amount of advertisement.

A strand of literature, mainly in the field of marketing, focuses on "vertical cooperative advertising." It stems from Berger (1972) continues through Cao and Ke (Forthcoming) and is surveyed by Jørgensen and Zaccour (2014). In such models, coordination problems arise between manufacturers and retailers of a good, each of which may choose to place ads in order to boost demand. In contrast, the current paper focuses on obstacles faced by multiple websites (or, more broadly, ad-funded platforms') in coordinating policies determining how much advertisement to show. In Subsection 2.2, we also provide a detailed comparison between this types of setting and ours.

A sizable literature has emerged in recent years focusing on situations where platforms such as

[5] Among many others, Spengler (1950), Rey and Tirole (1986), Shleifer and Vishny (1993), Lerner and Tirole (2004, 2015) study this phenomenon in a variety different situations.

search engines play an essential role in directing users to sellers of goods. Examples include Hagiu and Jullien (2011), Eliaz and Spiegler (2011), White (2013), Gomes (2014), and de Cornière (2016). In models of this sort, a central feature is the possibility for monetary transfers to be made from the seller(s) to the platform, which serves to eliminate the distortions that we study. Two articles that are more closely related are de Cornière and Taylor (2014) and Burguet et al. (2015). The structure of both of these models is like ours, in that users surf from an organizer to a content producer who cannot make transfers to one another. The focus of each of these papers, however, is to study the potential for bias on the part of the organizer in selecting the producer, and they do not address the websites' possible mis-coordination of advertisement, which is our main focus.

Given our focus on users' tendency to surf across the web, our model offers a new perspective on the issue of "multihoming," which has been studied by Anderson, Foros, and Kind (2017), Athey, Calvano, and Gans (2018), Liu (2018), as well as Gentzkow (2007), regarding online and print newspapers, and Katona and Sarvary (2008) studying the way such behavior influences the link structure of the web. Unlike our model, these works do not focus on situations where multiple sites play an essential role in satisfying a user's needs and thus do not focus on the same issues as we do. Our leading examples involves news aggregators, which Chiou and Tucker (2017) study empirically and Jeon and Nasr (2016) model in order to study equilibrium quality choices in the online news industry.

The two-sided market setup we adopt in Section 5 builds on a framework pioneered by Anderson and Coate (2005). Variants of this type of model appear in Gabszewicz, Laussel, and Sonnac (2004), Choi (2006), Peitz and Valletti (2008), and Crampes, Haritchabalet, and Jullien (2009). Finally, Evans (2008, 2009) and Choi, Mela, Balseiro, and Leary (2017) provide useful background about both the online advertising industry and research in this area. Like this paper, Van Zandt (2004) and Anderson and de Palma (2009; 2012), consider consumers with limited attention. They focus, however, on the excessive incentives that arise to send promotional messages, and they analyze such messages' impact on product markets with competing sellers.[6]

[6]See Eppler and Mengis (2004) for a multidisciplinary survey on the concept of "information overload".

2 Surfing Among Ad-Funded Websites

2.1 Model Setup

Suppose there are $n \geq 2$ websites, indexed by j, that together satisfy some web surfing "need" and are thus perfect complements for users. From visiting the entire set of sites, users derive payoff $v - \delta(\boldsymbol{a})$, where $v \sim [\underline{v}, \overline{v}]$, $\underline{v} \geq 0$ is their (heterogeneous) gross utility from satisfying the need, and $\delta(\boldsymbol{a}) \equiv \sum_{j=1}^{n} \delta_j(a_j)$ is the "distraction" or "nuisance" from being subjected to ads on the various sites. Without loss of generality, we normalize users' outside option to zero, and we assume they receive no benefit from visiting some but not all of the sites. The pdf of users' valuations, $f(\cdot)$, is assumed to be continuously differentiable and strictly positive over its entire support. Define $D(\delta) \equiv \int_{\delta}^{\infty} f(x)\,dx$, and note that it is twice differentiable and strictly decreasing over the interval $(\underline{v}, \overline{v})$. Let $h(x) \equiv \frac{D(x)}{-D'(x)}$ denote the inverse hazard rate, assumed to be strictly decreasing over this same interval. Demand for the bundle of sites, as a function of the vector of ad levels is thus $D(\delta(\boldsymbol{a}))$.

In the game, each website chooses an ad level, a_j, denoting its revenue per user that visits. Vector $\boldsymbol{a} \equiv (a_1, ..., a_j, ..., a_n)$ denotes the profile of the sites' ad levels, and $\overline{a} \equiv \sum_{j=1}^{n} a_j$ denotes the sum of ad levels. For all sites j, we assume distraction is zero when there are no ads, i.e., $\delta_j(0) = 0$, and that, as j chooses a higher ad level, the distraction from its ads grows at an accelerating rates, i.e., $\delta_j', \delta_j'' > 0$.[7] Finally, $c_j \geq 0$ denotes site j's marginal cost of serving users, and $\overline{c} \equiv \sum_{j=1}^{n} c_j$ denotes the sum of the sites' marginal costs.

Before proceeding to the analysis, we highlight two key extensions developed in later sections. Section 5 models websites explicitly as two-sided platforms that connect advertisers and users and shows that the main effects identified under this setup persist in that richer environment. In doing so, it considers an alternative specification of user preferences under which distraction from advertising is *fungible* across sites rather than convex with respect to the number of ads on each individual site. Section 6 relaxes the assumption that users make a fully-informed, one-shot decision about whether to surf to each site. There, instead, they decide sequentially whether to visit each site, they learn sites' ad levels only after visiting them for the first time, and they may decide to revisit.

[7] This assumption matches, in a stylized way, the empirical finding of Goldfarb and Tucker (2011) that, as the obtrusiveness of ads on a webpage increases, so does their total value to advertisers.

2.2 Miscoordination: Double Marginalization and Misplacement

As a benchmark, we consider the profit maximization problem of the entire set of websites, which we refer to as the "industry", assuming that second-order conditions hold, as specified in Appendix A. We then compare the solution to this problem to the outcome when sites set their advertising levels non-cooperatively. The industry profit maximization problem is given by $\max_{\boldsymbol{a}} (\overline{a} - \overline{c})\, D(\delta(\boldsymbol{a}))$, the solution to which, $\boldsymbol{a}^{\boldsymbol{\Pi}}$, is characterized by the set of first-order conditions

$$\overline{a}^{\Pi} - \overline{c} = \frac{h\left(\delta\left(\boldsymbol{a}^{\boldsymbol{\Pi}}\right)\right)}{\delta_j'\left(a_j^{\Pi}\right)}, \quad \forall j, \tag{1}$$

where $\overline{a}^{\Pi}$ denotes the sum of the elements in $\boldsymbol{a}^{\boldsymbol{\Pi}}$. Note that (1) implies that $\delta_j'\left(a_j^{\Pi}\right) = \delta_k'\left(a_k^{\Pi}\right)$, $\forall j, k$. Thus, in order for industry profits to be maximized, the marginal distraction from each site's advertising must be equalized.

We now examine sites' equilibrium behavior, using simultaneous Nash Equilibrium as our solution concept,[8] and assuming that technology allows all sites to make strictly positive profits at equilibrium. Taking as given other sites' equilibrium advertising levels, $\boldsymbol{a}_{-j}^{*}$, site j solves $\max_{a_j} \left(a_j - c_j\right) D\left(\delta\left(a_j, \boldsymbol{a}_{-j}^{*}\right)\right)$, yielding first-order condition

$$a_j^{*} - c_j = \frac{h\left(\delta\left(\boldsymbol{a}^{*}\right)\right)}{\delta_j'\left(a_j^{*}\right)}. \tag{2}$$

Summing over all sites' first-order conditions gives the following expression for the total margin per user the sites receive in equilibrium,

$$\overline{a}^{*} - \overline{c} = h\left(\delta\left(\boldsymbol{a}^{*}\right)\right) \sum_{j=1}^{n} \frac{1}{\delta_j'\left(a_j^{*}\right)}. \tag{3}$$

Proposition 1 interprets these first-order conditions, using the following two definitions.

Definition 1. *An equilibrium features* Double Marginalization *if its total level of user distraction exceeds the total level of user distraction at the industry optimum.*

[8] Assuming sequential timing leads to the same two distortions on which we focus but is more cumbersome.

Definition 2. *An equilibrium features* Misplacement *if industry profits can be increased by adjusting sites' ad levels while holding fixed the total level of user distraction.*

Proposition 1. *(a) Double marginalization occurs at equilibrium.*

(b) Misplacement occurs at equilibrium unless, for all $j \neq k$, $\delta'_j\left(a^*_j\right) = \delta'_k\left(a^*_k\right)$.

Whereas double marginalization results from the sites' failure to set the proper overall *level* of distraction, misplacement is a separate coordination problem that results from sites' failure to *allocate their demands for user attention*, in the most efficient way. We are unaware of any previous models that study, in a complements setting, a distortion like misplacement, which is driven by the combination of (a) the lack of monetary payments by users and (b) the complements setting. To see why misplacement arises and how it differs from potentially similar-seeming distortions that appear in other models, consider the comparisons between the following two alternative settings and ours.

Alternative Setting (a): Users Pay to Access Websites/Vertical Cooperative Advertising. Suppose that, in addition to showing ads, each of the sites charges users a "micropayment", $t_j \geq 0$, whenever they visit.[9] Then, site j's profits become $\Pi^j = \left(a_j + t_j - c_j\right) D\left(\delta_j\left(a_j\right) + t_j + K\right)$, where t_j is the micropayment that j charges and K is the total disutility caused by all other sites $k \neq j$, including both the distraction from their advertising and any micropayments they charge. As is shown in the literature studying free-to-air versus subscription television (Choi, 2006; Crampes, Haritchabalet, and Jullien, 2009; Peitz and Valletti, 2008),[10] when site j's profits take this form, its profit maximization implies that $\delta'_j\left(a_j\right) = 1$.

This can be interpreted as an "internal no misplacement" condition for j, between ads and money. To understand why this condition holds, suppose, for instance, that j chose a_j such that $\delta'_j\left(a_j\right) < 1$. Then it could increase its advertising level and decrease it price while holding fixed

[9]Note that the micropayment scenario we consider here differs from a subscription business model, in which users pay, for instance, a monthly fee in exchange for "all you can eat" access to a website. Such arrangements are not the primary cases we have in mind in this article. This is because, when a user makes a deliberate decision to subscribe to a given site (e.g., the *Wall Street Journal*), the form of complementarity that we concentrate on (which supposes that she depends on a homepage, an aggregator such as the Drudge Report, or a social media platform such as Facebook to locate the producer's website) is much less likely to arise. See, however, our discussion in the conclusion of Apple News Plus, which may prove to be an exception to this claim.

[10]These papers consider symmetric, substitute platforms that set ad levels and charge monetary prices.

total user disutility and strictly increasing revenue per user. This implies that, if all sites charge micropayments, then, $\forall j$, $\delta_j'\left(a_j\right) = 1$, and, therefore, there is no misplacement. (Meanwhile, double marginalization persists in its classic form.) If, however, one or more sites in the complementary set cannot charge micropayments, then misplacement of ads across sites can occur, because there is no force pushing the site(s) that cannot charge to adhere to a common "exchange rate" between revenue per ad and user distraction.

Note, further, that with minor modification, this model with monetary payments can alternatively be interpreted as one of vertical cooperative advertising (see, e.g., Berger (1972), Cao and Ke (Forthcoming) and Jørgensen and Zaccour's (2014) survey). In such settings, multiple players in a supply chain (e.g., a manufacturer and a retailer) both take partial responsibility for advertising a good that they jointly bring to market. To capture this configuration, let A_j denote firm j's total expenditure on ads promoting the good and let $\Delta_j\left(A_j\right)$ denote the boost in customers' willingness to pay generated by j's advertising. For simplicity (but without loss of generality with respect to our current point) assume there are two firms, $j = \mathbf{m}, \mathbf{r}$, so that $t_{\mathbf{m}}$ denotes the "wholesale price". Using P to denote the good's final price charged to consumers, let $t_{\mathbf{r}} \equiv P - t_{\mathbf{m}}$ denote the retailer's cut. Firm j's profits are then $\left(t_j - c_j\right) D\left(P - \sum_{j=\mathbf{m},\mathbf{r}} \Delta_j\left(A_j\right)\right) - A_j = \left(t_j - c_j\right) D\left(t_j - \Delta_j\left(A_j\right) + K\right) - A_j$, where K depends only on the rival's actions.

Firm j's first-order condition with respect to t_j requires that $t_j - c_j = h$. Plugging this pricing equation into j's first-order condition with respect to A_j then yields $\Delta_{\mathbf{m}}'\left(A_{\mathbf{m}}^*\right) = \Delta_{\mathbf{r}}'\left(A_{\mathbf{r}}^*\right) = 1/D$. Therefore, as in the case of user micropayments, the fact that the firms charge money prevents them from engaging in the equivalent of misplacement in promoting their joint product. To be clear, this argument by no means implies that other coordination problems are absent when firms cooperatively advertising. It does show, however, that a crucial driver of misplacement is the lack of money, transferred from the audience of ads to the parties setting ad levels, which is a particularly salient feature of ad-funded websites.

Alternative Setting (b): Sellers of Substitute Goods with Asymmetric Costs. Consider a model of Cournot competition among two sellers, each of which produce the same substitute good but with different cost functions. One "low cost" firm, $\mathcal{L}$, has cost function $C_{\mathcal{L}}(\cdot)$, while its "high cost" rival, $\mathcal{H}$, has cost function $C_{\mathcal{H}}(\cdot)$, where both of these functions are strictly increasing and strictly convex and, for any output level, $Q \geq 0$, set by an individual firm, it holds that $C'_{\mathcal{L}}(Q) < C'_{\mathcal{H}}(Q)$. Let $Q_{\mathcal{L}}$ and $Q_{\mathcal{H}}$ denote the firms' respective output levels, and let $\mathcal{P}(Q_{\mathcal{L}} + Q_{\mathcal{H}})$ denote the market's inverse demand curve, where $\mathcal{P}' < 0$ throughout the relevant domain. Assume that the two firms simultaneously choose their output levels. At equilibrium, each firm, $\mathcal{J} = \mathcal{L}, \mathcal{H}$, chooses its quantity to satisfy $\mathcal{P} + Q_{\mathcal{J}}\mathcal{P}' = C'_{\mathcal{J}}$. Since $\mathcal{P}$ and $\mathcal{P}'$ do not differ across firms, these first-order conditions imply both $Q_{\mathcal{L}} > Q_{\mathcal{H}}$ and $C'_{\mathcal{L}} < C'_{\mathcal{H}}$.[11] This means that, using the term of Daughety (2008), there is *maldistribution*. That is, even though $\mathcal{L}$'s equilibrium output is greater than $\mathcal{H}$'s, it would be possible to hold fixed the aggregate level of output and increase industry profits by reallocating production from $\mathcal{H}$ to $\mathcal{L}$ until $C'_{\mathcal{L}} = C'_{\mathcal{H}}$.

Maldistribution, however, is a distinct phenomenon from misplacement, as indicated by the following three points.

1. They stem from different sources. Whereas maldistribution is supply-driven, arising from asymmetry in the firms' cost functions, misplacement is demand-driven, arising from differences in users' tolerance for ads shown on different sites (and, as Section 5 shows, differences in advertisers' willingness to pay to advertise on different sites).

2. They differ in their key metrics. In the substitute model featuring maldistribution, the two firms have asymmetric output levels that are insufficiently dispersed. On the other hand, in our complement model featuring misplacement, each firm serves the same number of users, and the revenue they earn, per user, from showing ads is insufficiently dispersed.

3. Most importantly, the economics and managerial implications of misplacement lie in opposition to those of maldistribution. In this substitutes model, the industry as a whole suffers from both a sub-optimally low price and maldistribution, both of which can be remedied by reducing competition, e.g., by eliminating the inefficient firm from the market.[12] In contrast, in our complements model, as Section 3 shows, misplacement and double marginalization do not give rise to such an alignment of incentives. Instead, competition has opposite effects

[11]These first-order conditions also imply that in this example, unlike in our model, decentralization of control leads to markups that are lower than would be optimal for the industry as whole, not to double marginalization. To see this, note that, at equilibrium of this substitutes model, $\mathcal{P} - \frac{C'_{\mathcal{L}}+C'_{\mathcal{H}}}{2} = -Q\mathcal{P}'/2$, whereas joint optimization would require $\mathcal{P} - \frac{C'_{\mathcal{L}}+C'_{\mathcal{H}}}{2} = -Q\mathcal{P}'$.

[12]See Schwartz (1989) for an example following this pattern.

on the two distortions. It tends to reduce double marginalization but increase misplacement. Thus, the optimal level must take both effects into account.

3 Competition

How can websites overcome these two coordination problems? The first, double marginalization, is a widely studied phenomenon. In cases where a final good is comprised of multiple, complementary components, a well-known solution to the double marginalization problem, that can be used to maximize industry profits, is to introduce competition within the markets for all but one of the individual components. For instance, suppose "hardware" and "software" were two perfectly complementary products, originally produced by two separate firms, the sum of whose prices is inefficiently high. If the market for hardware became perfectly competitive, then the single software maker would be in a position to charge a price that implements the profit maximizing outcome for the industry as a whole.[13]

That logic no longer holds, however, in our setting. Instead, when a set of complementary websites, whose revenue comes exclusively from advertising, engage in both misplacement and double marginalization, a tradeoff arises. As this section shows, although an increase in competition among websites within one category reduces double marginalization, it tends, at the same time, to make the misplacement problem more severe. Indeed, in some cases, the overall effect of such an increase in competition can be a *decrease* in total profits.

To illustrate this, we specialize the model introduced earlier in this section to the case where there are two sites with distinctly complementary roles. The first site is an *organizer* of content and the second site is a *producer* of content. We denote the organizer (which "sorts" content) by s and the producer by w.

We now examine the effects of increasing the competition faced by one of the two sites, which we will assume to be the producer.[14] Advertising technology is denoted by $\delta(\boldsymbol{a}) = \delta_s(a_s) + \delta_w(a_w)$.

[13]Casadesus-Masanell, Nalebuff, and Yoffie (2007) and Cheng and Nahm (2007) study variations of such models, applied to hardware and software, focusing on the case where the producers in the sectors with competition are vertically differentiated. Many other articles covering a wide range of topics also consider the idea of using competition among firms in a particular "category" of a complementary bundle as a solution to the double marginalization problem. In the topic of performance-based fees in online advertising, see Dellarocas (2012). More broadly, also see, for instance, Rey and Tirole (1986), Shleifer and Vishny (1993), Lerner and Tirole (2004), and Lerner and Tirole (2015).

[14]In the subsequent analysis, we consider variation in the competitiveness of producers; symmetric results hold if one

Define the function $\overline{\Pi}(a_w)$, denoting total industry profits as a function of the level of advertising chosen by w, assuming s best-responds to w's choice. Formally,

$$\overline{\Pi}(a_w) \equiv \Pi^s\left(a_s^*\left(\delta_w\left(a_w\right)\right), \delta_w\left(a_w\right)\right) + \Pi^w\left(a_w, \delta_s\left(a_s^*\left(\delta_w\left(a_w\right)\right)\right)\right), \tag{4}$$

where $a_s^*\left(\delta_w\left(a_w\right)\right) \equiv \arg\max_{a_s} \Pi^s\left(a_s, \delta_w\left(a_w\right)\right)$. Taking the derivative of (4) gives[15]

$$\begin{aligned}\overline{\Pi}'(a_w) &= \underbrace{\frac{\partial \Pi^s}{\partial a_s}}_{0} a_s^{*\prime}\delta_w' + \frac{\partial \Pi^s}{\partial \delta_w}\delta_w' + \frac{\partial \Pi^w}{\partial a_w} + \frac{\partial \Pi^w}{\partial \delta_s}\delta_s' a_s^{*\prime}\delta_w' \\ &= \underbrace{\left(a_w - c_w\right) D'\delta_w'\left(1 + \delta_s' a_s^{*\prime}\right)}_{\text{Loss from users dropping out}} + \underbrace{D\left(1 - \frac{\delta_w'}{\delta_s'}\right)}_{\text{Change in rev. from existing users}}. \end{aligned} \tag{5}$$

The right-hand side of equation (5) collects the impact of a change in a_w into two interpretable effects. Regarding the former, when a_w increases, the propensity of users to stop visiting the combination of sites is mitigated by the corresponding reduction in distraction caused by s, captured by the term $\delta_s' a_s^{*\prime}$, which is negative due to its second factor. Regarding the latter effect, when a_w increases, w's revenue per user increases proportionally to the change in a_w; however, when s responds optimally, its revenue per user decreases by the ratio of w's marginal distraction to s's marginal distraction.

Stemming from this analysis, we now state Proposition 2. It shows the sub-optimality, to the industry as a whole, of both the configuration with one organizer and one producer and the configuration with one organizer and perfectly competitive producers.

Proposition 2. *(a) Consider the Nash Equilibrium outcome when there is a single organizer, s, and a single producer, w. If the producer's advertising level is exogenously decreased, and the organizer responds optimally, then total industry profits increase, if and only if, locally,*

$$\underbrace{h'}_{<0} \cdot \left(\delta_s' - \delta_w'\right) + \underbrace{\left(1 + h\frac{\delta_s''}{\left(\delta_s'\right)^2}\right)\delta_w'}_{>0} > 0. \tag{6}$$

varies the competitiveness of organizers.

[15] To derive the right-hand side of (5), note that $\frac{\partial \Pi^s}{\partial \delta_w} = \left(a_s^* - c_s\right) D'$, $\frac{\partial \Pi^w}{\partial a_w} = D + \left(a_w - c_w\right) D'\delta_w'$, and $\frac{\partial \Pi^w}{\partial \delta_s} = \left(a_w - c_w\right) D'$. Moreover, the first-order condition for s implies that $a_s^* - c_s = -\frac{D}{D'\delta_s'}$.

(b) *Starting from an outcome featuring the perfectly competitive level of producer advertising ($a_w = c_w$) and monopolistic behavior by the organizer, if a_w is exogenously increased, then total industry profits increase if and only if, locally, $\delta_w'(c_w) < \delta_s'(a_s^*(\delta_w(c_w)))$.*

Proposition 2 lends itself to the following interpretation. Part (a) supposes that the *status quo* involves one producer and then contemplates an increase in competition from other producers, leading to a decrease in a_w, starting from the initial, "bilateral monopoly" level. Part (b) supposes that the *status quo* involves perfectly competitive producers and then imagines that this competition softens, leading to an increase in a_w, starting from c_w. For our current purposes, the particular form of competition one might consider, among producers, is beside the point, and thus we directly consider exogenous changes in a_w. (In Section 4, we explicitly model such competition.)

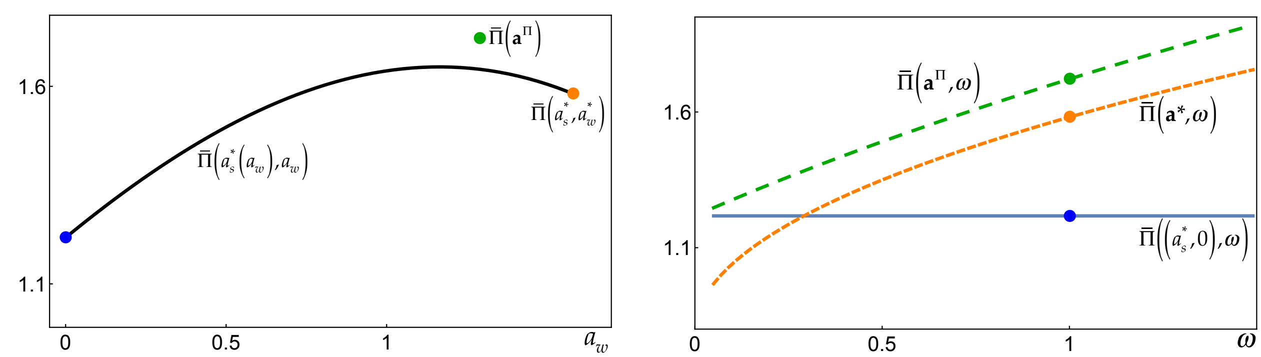
The result of part (a) matches the standard behavior of a bilateral monopoly model, so long as the condition in (6) is satisfied. That is, when the *status quo* involves two firms offering perfectly complementary items, industry profits increase when one of them is subjected to more competition.

Part (b), however, goes against this intuition. According to the typical argument, decreasing competition among producers should lead to more double marginalization and, therefore, lower industry profits. Here, though, because of the separate misplacement distortion, such an argument is no longer valid. To see this, consider the condition given in part (b): $\delta_w'(c_w) < \delta_s'(a_s^*(\delta_w(c_w)))$. Roughly speaking, the left-hand side of this inequality can be thought of as the marginal distraction to users from the producer's *first* advertisement. Meanwhile, the right-hand side is the marginal distraction of an additional ad shown by the organizer, starting from its monopoly level. If the producer's advertising technology allows it to show its "first ad" in a way that creates less nuisance than the organizer's "last ad", then the misplacement distortion dominates, and the industry benefits from the decrease in producer competition, even though this increases double marginalization.

Broadly speaking, the important point of Proposition 2 is the following. With ad-funded sites, unlike in situations where consumers pay with money, an interesting new tradeoff arises as competition increases among producers. In a standard model in which consumers spend money, such an increase in competition, when taken far enough, would deliver industry-optimal profits. Here, however, competition fails to do so, because it leads to increased misplacement. As the following example illustrates, the industry-optimal intensity of competition among producers is an intermediate level that takes the misplacement-vs.-double marginalization tradeoff into account.

Example. Suppose that $D(\delta) = 1 - \delta/10$, that $c_s = c_w = 0$, and that $\delta(\boldsymbol{a}) = a_s^2 + a_w^2/\omega$. Here, the parameter ω captures the relative efficiency of the producer's advertising technology. First we allow the two sites to be fully symmetric ($\omega = 1$) and supplement the exercise considered in Proposition 2. To do so, we examine total industry profits while exogenously varying the organizer's advertising level, a_w, from zero (perfect producer competition) to its equilibrium level under bilateral monopoly. As Figure 1(a) shows, this curve takes an inverted "U" shape: the dominant effect of the initial increase in a_w is to reduce misplacement, but, to the right of the peak, the harm from double marginalization takes over.

Next, we allow ω to vary and compare industry profits under the two polar cases of producer competition. Figure 1(b) reports profits under three regimes, perfect producer competition, bilateral monopoly, and joint profit maximization, as a function of ω. The crucial point is that, for low values of ω, when the producer's ad technology is inefficient, the first of these regimes yields greater profits, because solving the misplacement problem is of little value. In contrast, for sufficiently large values of ω, sacrificing the use of the producer's technology is very costly. Thus, bilateral monopoly becomes more profitable for the industry, despite the double marginalization that it invites.



(a) Inverted "U" curve: industry profits as a_w varies from 0 (perfect competition) to a_w^* (bilateral monopoly). Dot lying above is industry optimum.

(b) Industry profits as producer's technology varies. Bilateral monopoly exceeds perfect competition for sufficiently high ω. Joint maximization lies on top.

Figure 1: Industry profit comparisons in the numerical example.

4 Linking to Content versus Acquiring It

This section turns to the question suggested by our initial, motivating example comparing the strategies of MSN and the Drudge Report. It asks under what circumstances it is better for an organizer to act as a link in a complementary chain and under what circumstances it is better off acquiring content for itself. Subsections 4.1 and 4.2 set up and analyze the basic framework, and Subsection 4.3 extends this to show how the tradeoff between misplacement and double marginalization identified in Section 3 directly affects this choice.

4.1 Setup

Now assume there is one organizer, s, with advertising technology $\delta_s(\cdot)$, and there are l producers, $j = w1, ..., wl$, each with an identical advertising technology, $\delta_w(\cdot)$. In an initial stage, the organizer has the choice of two modes of operation. In one mode, which corresponds directly to the model considered so far, the organizer's site contains links to other sites operated by producers, as well as ads. We call this mode *linking*. In the other mode, the organizer contracts in advance with a specific subset of producers in order to acquire their content and thus control the ads that appear alongside it. We call this mode *acquision.* The timing in these two modes is as follows.

- **Linking**

  1. The organizer and each of the l producers simultaneously set their ads levels. These are publicly observed.
  2. Nature selects, with equal probability across all producers, exactly one of them to receive a *scoop* (i.e., an exclusive news story or other piece of content that interests users). The organizer observes which of the producers has received the scoop but users do not. The organizer links to the producer that has the scoop in a way that clearly signals its relevance (e.g., it places this link in the most prominent slot at the top of the page).[16]
  3. Users decide whether or not to surf. That is, if they surf, they visit the organizer and click through to the site of the producer that has the scoop, as signaled by the organizer's

[16]In order to avoid unhelpful game-theoretic complexity regarding users' beliefs, we do not model the organizer's policy of clearly displaying the site with the scoop as a strategic choice. However, when this decision is modeled as part of a game of incomplete information, it is straightforward to support this outcome as a Perfect Bayesian Equilibrium.

link. In doing so, they receive benefit v but are subjected to ads on both sites, causing distraction of $\delta_s(a_s) + \delta_w\left(a_j\right)$, where j denotes the producer with the scoop.

- **Acquision**

  1. The organizer offers contracts to form *content partnerships* with as many as $m \leq l$ producers.[17] If signed, these contracts give the organizer the right to host these producers' content on pages that it controls. The organizer sets ad levels for both its main page and the pages on which the acquired content appears. Users observe both the number of content partnerships formed by the organizer, and they learn its two ad levels.
  2. Nature selects, with equal probability across all producers, exactly one of them to receive a scoop. The organizer observes which of the producers this is, but users do not.
     (a) If the organizer has acquired this producer's content, then the organizer places a link in its most prominent slot to a page that it hosts, containing the scoop.
     (b) If the organizer has not acquired this producer's content, then the organizer places a link in its most prominent slot to a page that it hosts containing material from an arbitrary producer whose content it has acquired.
  3. Users decide whether or not to surf. That is, if they surf, they visit the organizer and click through to its most prominently linked content page. They derive utility v from the episode if and only if this content page contains the scoop, but they are subjected to the ads on both pages in either case, causing distraction of $\delta_s(a_s) + \delta_w(a_{ws})$, where a_{ws} denotes the ad level on the content page, as chosen by the organizer.

4.2 Equilibrium

We first briefly discuss the outcomes arising under each of these two modes. Later, we compare their profitability for the organizer in order to determine which one it selects at equilibrium.

Under linking mode, the analysis of Sections 2 and 3 applies directly, as though there were just two sites – one organizer and one producer. This is because any given producer receives user traffic only if it gets the scoop. Thus, its choice of ad level affects its payoff only in this state of

[17] This upper limit of m partnerships captures, in the simplest possible way, the transaction costs involved in making such arrangements, which are a separate cost from the payments made to producers, discussed below.

nature. Consequently, each of the l producers sets its ad level as in the case of "bilateral monopoly" described in Section 3. This reasoning leads to the following remark.

Remark 1. *Under linking mode, Proposition 1 applies. That is, at equilibrium, double marginalization occurs, and misplacement occurs unless* $\delta_s'(a_s^*) = \delta_w'(a_w^*)$ *, where, here,* a_w^* *denotes the ad level chosen by each of the producers.*

Under acquisition mode, users rationally anticipate that the organizer acquires the scoop with probability m/l. Thus, they choose to surf if and only if $vm/l \geq \delta_s(a_s) + \delta_w(a_{ws})$, prompting the following remark.

Remark 2. *Under acquisition mode, when setting the two ads levels, the organizer faces demand of* $D\left(\frac{l}{m}(\delta_s(a_s) + \delta_w(a_{ws}))\right)$.

In order to compute the organizer's profits under acquisition mode, we must characterize the equilibrium of the contracting game that it plays with producers. In particular, we wish to determine the size of the upfront fee that the organizer must pay a given producer in order to acquire its content. We do so using the following logic. The lower bound of this fee is zero, which would prevail when the organizer can make take-it-or-leave-it offers and the outcome for any given producer if it fails to enter into an agreement is to receive no user traffic. The upper bound of this fee is the level of expected profits earned by a producer under linking mode, which we denote by Π_w^L. This fee would prevail if each producer (or a coalition thereof), by refusing to enter into an agreement, could cause all negotiations to break down and leave the organizer with no choice other than to revert to linking mode.

To capture this range of possibilities, we introduce parameter $\rho \in [0, 1]$ representing the bargaining power of the producers, and we write the equilibrium fee paid by the organizer to each producer as $\rho\Pi_w^L$. As our prior expectation in the linking-verus-acquisition comparison would likely be for acquisition to be more profitable for the organizer, except where explicitly noted, we focus on the most conservative assumption (that is, the most favorable to acquisition) of $\rho = 0$, and then we briefly discuss the case of $\rho > 0$.

We now compare the organizer's profits under these two modes. To do so, we readopt the functional form assumptions used in the example above: $D(\delta) = 1 - \delta/10$, $c_s = c_j = 0$, and $\delta(\boldsymbol{a}) = a_s^2 + a_w^2/\omega$.

Proposition 3. *Assume $\rho = 0$. The organizer prefers linking mode to acquisition mode if and only if getting the scoop under the latter mode is sufficiently uncertain, and the advertising technology available on the organizer's main page, $\delta_s(\cdot)$, is relatively efficient compared to the one available on content pages, $\delta_w(\cdot)$. Formally, linking mode arises at equilibrium if and only if $\frac{m}{l} \leq \frac{27}{64(1+\omega)}$.*

The crucial factor driving this result is the limit on the organizer's ability to acquire, upfront, the content of all producers that could potentially get the scoop. The key advantage of linking mode is, thus, the flexibility that it allows the organizer in choosing the most relevant source of content. Under linking mode, users are more confident that the organizer will link to relevant content and are thus willing to tolerate higher levels of distraction from advertisement. As the ratio grows of maximum feasible content partners to total number of producers, the additional flexibility under linking mode diminishes until, eventually, acquisition mode dominates.

Furthermore, linking mode tends to be preferable if the technology available for advertising on content pages is relatively inefficient. This is because, under linking mode, as ω gets smaller, the equilibrium ad level set by producers decreases. As this occurs, the value to the organizer of taking control over the web advertisement level (and thus eliminating the two distortions identified in Proposition 1) decreases.

Under the most extreme assumption in favor of acquisition mode, $\rho = 0$, the organizer can acquire, for free, the content of as many as m producers, and this constraint is binding. If, as is more likely, the organizer must pay a strictly positive amount in order to acquire each producer's content ($\rho > 0$), this directly increases the cost of forming m content partnerships, thereby increasing the relative profitability of linking (see Figure 2).[18]

4.3 Non-Exclusive Scoops

So far in this section, we have made the simplifying assumption that exactly one of the producers receives the scoop exclusively. We now generalize the model to allow for relevant news to be obtained non-exclusively, by multiple producers. For technical simplicity, we assume that there are 2 producers and that, if the organizer chooses acquisition mode, it can form a content partnership

[18]Note that, if ρ, ω and m all become sufficiently large, under acquisition, the organizer's optimal number of content partnerships becomes strictly less than m, the maximum feasible number. When this occurs, acquisition is always more profitable than linking.

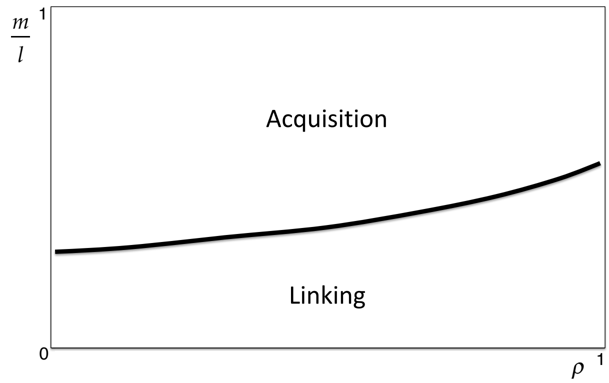


Figure 2: The organizer's preferred mode as a function of ρ and m/l, when $\omega = 1/2$.

with only of them. Note that, in the version with exclusive scoops discussed above, when $m = 1$ and $l = 2$, the organizer always chooses acquisition mode.

Here, $w1$ and $w2$ each receives the scoop exclusively with probability $(1 - q)/2$, while, with probability q, they both obtain content that is relevant to users. When only one of the producers receives the scoop, the situation is comparable to the one studied above; however, when both producers are relevant, in flexibility mode, the organizer prefers to link to the one with fewer ads. This gives each producer the incentive to undercut its rival in one state of the world but not in another. As a result, they adopt mixed strategies. We now describe the updated timing.

- **Linking**

  1. The organizer and the 2 producers set their advertising strategies, which may be mixed and are denoted by CDFs $F_s(\cdot)$ and $F_j(\cdot)$, $j = w1, w2$. All ad levels are then realized and publicly observed.
  2. Nature selects which pruducer(s) receive the scoop, with the probabilities mentioned above. The organizer observes which of the producers has received the scoop but users do not. If only one producer receives the scoop, the organizer links to that one. If both receive the scoop, the organizer links to the one with the lower ad level.[19]
  3. Users decide whether or not to surf, as described in Subsection 4.1.

[19] In case of a tie, the organizer chooses each organizer with probability 1/2. At equilibrium, ties will occur with zero probability.

- **Acquisition**: unchanged compared to Subsection 4.1. Here, $m = 1$, $l = 2$, and we continue to focus on the case of $\rho = 0$.

Given the behavior described in stages 2 and 3, we solve for mixed-strategy Nash equilibrium in stage 1 that is symmetric across the two producers, and we refer to such an equilibrium as "producer-symmetric." Proposition 4 characterizes this. In doing so, it refers to a_w^*, the bilateral monopoly ad level discussed in Remark 1 and best-response function $a_s^*(\delta_w) \equiv \arg\max_{a_s} \Pi^s(a_s, \delta_w)$, defined in Section 3, in a context where the producer played a pure strategy.

Proposition 4. *Assume $q \in (0,1)$. Under linking mode, the game played in stage 1 has exactly one producer-symmetric mixed-strategy equilibrium. It has the following features.*

(a) The producers' equilibrium advertising strategy is an atomless distribution with strictly positive support over an interval $[\underline{a}, \overline{a}]$, where $0 < \underline{a} < \overline{a} < a_w^*$.

(b) When the two producers play a symmetric mixed-strategy, the organizer's best response is $a_s^*(\delta_0)$, where $\delta_0 \equiv q\mathbb{E}\left[\min_{\{w1,w2\}}\left\{\delta_j\right\}\right] + (1-q)\,\mathbb{E}\left[\delta_j\right]$ and, thus, at equilibrium, $\delta_0^* \in \left(\delta_w(\underline{a}), \delta_w(\overline{a})\right)$.

This result says that when scoops that are sometimes exclusive to one producer and sometimes offered by both (i.e., $q \in (0,1)$), producers' competition with one another is of intermediate intensity.

Now consider the relative profits for the organizer under the two modes. Proposition 3 showed that, when $q = 0$, the organizer may prefer either mode, but that, in the specific case we are currently studying featuring a total of $l = 2$ producers and a maximum of $m = 1$ content partnerships, acquisition always dominates. Next, suppose $q = 1$, so that there are no exclusive scoops (and maintain the assumption that $\rho = 0$. Here, it is straightforward to see that acquisition dominates, because the organizer obtains, at no cost, content that is relevant with probability 1, and, in this mode, but not under linking, it can set both ad levels without any distortions away from industry-optimal levels. It may thus seem reasonable to expect that, as q increases from 0 to 1, acquisition becomes better, compared to linking. However, as Proposition 5 shows, this is not the case.

Proposition 5. *The organizer prefers linking mode to acquisition mode when (i) producers sometimes but not always receive exclusive scoops, and (ii) ads on content pages are relatively inefficient. Formally, there exist $\overline{\omega} > 0$ and $0 < \underline{q} < \overline{q} < 1$ such that, for all $\omega \in (0, \overline{\omega})$, $q \in (\underline{q}, \overline{q})$, the organizer chooses linking mode. Figure 3 illustrates this result.*

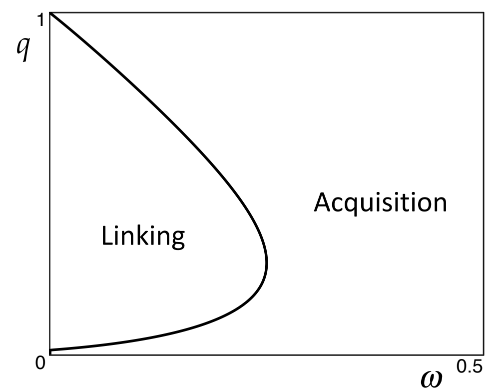


Figure 3: The organizer's preferred mode as a function of ω and q.

For a given level of ω, why do intermediate frequencies of exclusive scoops favor linking when either full exclusivity or no exclusivity favor acquisition? Misplacement plays a key role in the answer to this question. To see why, let us take stock of the shifts that occur, under each of the two modes, as q increases.

Under acquisition mode, as q increases, the relevance of the organizer's content becomes more reliable. This boosts its demand, pushing profits upward. In the meantime, its advertising efficiency remains unchanged, because, regardless of the value of q, it can choose the industry-optimal levels.

Under linking mode, as q increases, the relevance of the organizer's content remains unchanged, because it can always link to a relevant producer. In the meantime, as exclusivity becomes scarcer, producers compete harder, reducing the expected distraction caused by their ads. As q increases anywhere along the unit interval, this decrease in expected distraction leads to an increase in organizer profits. Crucially, however, when ω is sufficiently low, on the one hand, as q increases in a range just above zero, this also leads to a decrease in misplacement. On the other hand, when q increases in a range just below one, this leads to an increase in misplacement. Hence, in the former case, the relative benefit the organizer perceives from taking control over ad levels diminishes with q over the lower interval but grows with q over the higher one.

5 A Two-Sided Markets Approach

So far in the paper, in order to maximize expositional clarity, we have not explicitly modeled advertisers. Instead, we have taken the shortcut of assuming that each website, j, chooses its "advertising level", a_j, which represents the revenue it receives, per user, from advertising. We now specify the model more fully, by including advertisers, and show that the main themes identified above still hold. We first establish conditions under which double marginalization and misplacement persist. In particular, this two-sided framework shines a brighter light on the drivers of misplacement. These, it will be shown, include not only asymmetries in sites' advertising technologies but also in their demand from advertisers. Second, we address the possibility that, by relying on an ad agency to determine their advertisement policies, the websites may be able to overcome these distortions. The basic motivation for this exercise is the possibility that, although the ad agency does not intrinsically care about the profits of the websites, its more centralized perspective may be helpful to them. We show that while this is sometimes true, any such benefits do not prevail systematically.

5.1 The Model with Advertisers

Setup. As before, there is a continuum of users, of mass one, for whom two websites, an organizer, s, and a producer, w, are perfect complements. To model advertisers, we build closely on the setup introduced in Anderson and Coate's (2005) classic article. Assume that there is a continuum of advertisers, of mass $m_s + m_w$. Each advertiser is a monopolist of its own particular good, which costs 0 to produce. For a given advertiser's good, some users have valuation $\theta_j > 0$, while others have valuation 0. Thus, all advertisers on site j charge a price of θ_j for their respective goods.

Initially, we consider the simplest possible configuration in which there are two separate pools of advertisers. Advertisers in one pool may wish to appear on the organizer's site, while advertisers in the other pool may wish to appear on the producer's site. Later, we relax this assumption of separate pools. Within each pool, advertisers are heterogeneous in the number of users to whom they appeal. Specifically, a potential advertiser on site j features type $\sigma_j \in [0, 1]$. This can be interpreted as the fraction of users who, upon seeing the ad in question displayed on site j, click on it and then proceed to purchase the advertiser's good.

As a foundation for this assumption, one can suppose that $\sigma_j = \xi_j \mu$, where ξ_j is the probability that any user visiting site j notices that particular ad, and μ is the independent probability that any given user has valuation θ_j for that advertiser's good. We assume that σ_j is distributed according to density $g_j(\cdot)$ that is continuously differentiable and strictly positive on the open unit interval, and we denote the corresponding cdf by $G_j(\cdot)$. Finally, we assume that each advertiser may also choose an outside option that gives a payoff of 0.

We assume that the each site non-cooperatively picks the number of ad "slots" to offer, and that ad prices adjust in order for markets to clear.[20] Let p_j denote the price per user charged to advertise on site j. An advertiser of type σ_j will choose to advertise if and only if $\theta_j \sigma_j \geq p_j$. Consequently, site $j's$ demand for advertising can be written $q_j\left(p_j\right) = m_j\left(1 - G_j\left(p_j/\theta_j\right)\right)$, implying that inverse demand is $p_j\left(q_j\right) = \theta_j G_j^{-1}\left(1 - q_j/m_j\right)$. Note that, assuming G_j is differentiable, for all $q \in \left(0, m_j\right)$, $p_j'(q) < 0 < p_j(q)$.

User preferences take the same form as they have throughout the paper. However, we can now write distraction from advertising directly as a function of the number of ads appearing on each of the two sites. That is, $\delta(\boldsymbol{q}) = \delta_s(q_s) + \delta_w(q_w)$, where $\delta_j(0) = 0$ and $\delta_j', \delta_j'' > 0$.

Assume that site j's demand from advertisers gives rise to revenue per user, $q_j p_j\left(q_j\right)$, that is strictly concave and single-peaked in q_j, when the number of users is held fixed. Also, note that, because more ads means more distraction for users, the number of users that visit the pair of sites, $D(\delta(\boldsymbol{q}))$, strictly decreases with q_j. Therefore, in considering a site's optimization problem, we can restrict attention to ad levels that are "to the left of the peak" of $q_j p_j\left(q_j\right)$, that is, ad levels such that each site's revenue per user is increasing rather than decreasing. To this end, denote site j's revenue per user in the relevant interval by $r_j\left(q_j\right) \equiv q_j p_j\left(q_j\right)$, defined over the domain $\left\{q_j : \frac{d}{dq_j}\left\{q_j p_j\left(q_j\right)\right\} > 0\right\}$ and thus strictly increasing in q_j. Further assume that second-order conditions hold, as specified in Appendix A.

Miscoordination. First consider the benchmark of industry profit maximization. This problem is given by $\max_{q_j}\left(r_j\left(q_j\right) + r_k(q_k) - \bar{c}\right) D(\delta(\boldsymbol{q}))$, which gives rise to first-order conditions

$$r_j\left(q_j^\Pi\right) + r_k\left(q_k^\Pi\right) - \bar{c} = h\left(\delta\left(\boldsymbol{q}^\Pi\right)\right) \frac{r_j'\left(q_j^\Pi\right)}{\delta_j'\left(q_j^\Pi\right)}, \quad j = s, w. \tag{7}$$

[20] Under the current "separate pools" configuration, it is equivalent to assume each site sets an advertising price per user, but, with more general advertiser demand, such conduct assumptions have bite.

Equation (7) is the analog, in this fully specified two-sided model, of equation (1) in Subsection 2.2. It implies that, at the industry optimum, $r_j'\left(q_j^{\Pi}\right)\Big/\delta_j'\left(q_j^{\Pi}\right) = r_k'\left(q_k^{\Pi}\right)\Big/\delta_k'\left(q_k^{\Pi}\right)$. When this holds, it means that, on the two sites, the ratios of marginal revenue from showing an extra ad to marginal distraction from showing an extra ad are equalized. Below, we further explore the significance of this equation, relative to its reduced-form counterpart.

Turning to Nash equilibrium, when the sites operate independently, site j takes q_k as given and solves $\max_{q_j}\left(r_j\left(q_j\right)-c_j\right)D\left(\delta\left(\boldsymbol{q}\right)\right)$. This yields first-order condition

$$r_j\left(q_j^*\right)-c_j = h\left(\delta\left(\boldsymbol{q}^*\right)\right)\frac{r_j'\left(q_j^*\right)}{\delta_j'\left(q_j^*\right)}, \tag{8}$$

which, summed over both sites, implies the total equilibrium margin per user to be

$$r_s\left(q_s^*\right)+r_w\left(q_w^*\right)-\bar{c} = h\left(\delta\left(\boldsymbol{q}^*\right)\right)\left(\frac{r_s'\left(q_s^*\right)}{\delta_s'\left(q_s^*\right)}+\frac{r_w'\left(q_w^*\right)}{\delta_w'\left(q_w^*\right)}\right). \tag{9}$$

Equations (8) and (9) are the two-sided analogs of equations (2) and (3), respectively. Maintaining the definitions of double marginalization and misplacement established above, Proposition 6 shows that these distortions persist in the two-sided model.

Proposition 6. *(a) Double marginalization occurs at equilibrium.*

(b) Misplacement occurs at equilibrium unless $r_s'\left(q_s^*\right)/\delta_s'\left(q_s^*\right) = r_w'\left(q_w^*\right)/\delta_w'\left(q_w^*\right)$.

The Drivers of Misplacement. Prior to this section, our discussion of misplacement has focused on asymmetries in sites' advertising technologies, represented by $\delta\left(\cdot\right)$, as the key driver. Fully specifying this two-sided model reveals that, while such technological asymmetries may indeed cause misplacement, so may asymmetries in the demand for advertisement on the different sites.

To see this with full clarity, consider the following variation of the model. Assume that $\delta\left(\boldsymbol{q}\right) = \delta\left(q_s+q_w\right)$, meaning that, from the standpoint of the user distraction they cause, ads are fully *fungible* across the two sites. When advertising exhibits fungibility, no matter what the ad level is, the marginal distraction from a change in q_s is always the same as that from a change in q_w. Thus, under this setup, misplacement arises at equilibrium if and only if the two sites' respective marginal revenue levels on the advertiser side of the market differ from one another, that is, whenever $r'_s(q_s) \neq r'_w(q_w)$.

This suggests that, when multiple, complementary sites act in a decentralized way, misplacement may be even more likely to arise than the one-sided analysis indicates. Even if one were to argue, for instance, that different sites were likely to coalesce around certain advertising technologies with comparable marginal levels of distraction, this would not imply that they face the same demand as one another from advertisers. Moreover, even if advertising exhibited fungibility and the advertiser demand curves were symmetric across the two sites, misplacement would still arise if, as modeled in Proposition 2, the two organizer and the producer faced differing levels of competition from similar websites, as this would lead to asymetric ad levels at equilibrium.

More General Advertiser Demand. We now relax the assumption that there are two separate pools of advertisers, each of which is available to just one site. Suppose, instead, there is a single continuum of advertisers of mass m. Maintaining the key feature of Anderson and Coate's setup, we assume that advertisers extract all surplus arising from any transactions with users. We now let $\boldsymbol{\phi} = \left(\phi_s, \phi_w\right)$ denote a given advertiser's willingness to pay, per user, for its ad to appear on the respective sites. The fact that, for a given advertiser, ϕ_s and ϕ_w may differ from one another represents the idea that individual advertisers do not necessarily fit equally well on different types of websites. For example, one ad might perform relatively well in the text based form more common to a search engine, whereas another might perform better as a banner alongside a news story. We assume that $\boldsymbol{\phi}$ is distributed according to a joint density $g(\cdot)$ that is continuously differentiable and strictly positive on some compact two-dimensional interval with bounds $\underline{\boldsymbol{\phi}}$ and $\overline{\boldsymbol{\phi}}$. Finally, each advertiser wishes to advertise on at most one of the two sites and may also choose an outside option yielding 0.

The two sites non-cooperatively pick the number of ad slots, and prices adjust so that markets clear. An advertiser of type $\boldsymbol{\phi}$ will choose to advertise on site j if and only if $\phi_j - p_j \geq \max\left\{0, \phi_k - p_k\right\}$,

$k \neq j$. Consequently, as a function of $\boldsymbol{p} = (p_s, p_w)$, site j's demand for advertisement can be written

$$\begin{aligned} q_j(\boldsymbol{p}) &= m \cdot \Pr\left(\phi_j - p_j \geq \max\left\{0, \phi_k - p_k\right\}\right) \\ &= m \cdot \int_{p_j}^{\overline{\phi_j}} \int_{\underline{\phi_k}}^{p_k + \phi_j - p_j} g(\boldsymbol{\phi})\, d\phi_k d\phi_j. \end{aligned}$$

As $\frac{\partial q_j}{\partial p_j} < 0 < \frac{\partial q_j}{\partial p_k}$, this demand system exhibits the well-known "gross substitutes" property.[21] Therefore, there exists an inverse demand function, $\boldsymbol{p}(\boldsymbol{q}) : \mathbb{R}^2_{++} \to \mathbb{R}^2$, where $p_j(\boldsymbol{q})$ denotes the market-clearing price-per-user to show an ad on site j, given the number of ad slots offered by each of the two sites. Analogously to the version above, we define $r_j(\boldsymbol{q}) \equiv q_j p_j(\boldsymbol{q})$ over the domain $\left\{\boldsymbol{q} : \frac{d}{dq_j}\left\{q_j p_j(\boldsymbol{q})\right\} > 0\right\}$. User preferences remain unchanged. Proposition 7 deals with misplacement under this setup.

Proposition 7. *Misplacement occurs at equilibrium unless* $\left(\frac{\partial r_s}{\partial q_s} + \frac{\partial r_w}{\partial q_s}\right) \Big/ \delta_s'(q_s^*) = \left(\frac{\partial r_w}{\partial q_w} + \frac{\partial r_s}{\partial q_w}\right) \Big/ \delta_w'(q_w^*)$.

This proposition show that the above discussion on the drivers of misplacement does not depend on the simplifying "separate pools" assumption. In this more general environment, advertiser demand may still cause misplacement independently of asymmetry in the sites' advertising technologies.

Note that Proposition 7 makes no claim about double marginalization. This is because, interestingly, under this more general advertiser demand system, we are unable to rule out the possibility that equilibrium distraction is lower than the industry optimum level. If the functions $\frac{\partial^2 r_j}{\partial q_j \partial q_k} = \frac{\partial p_j}{\partial q_j} + \frac{\partial^2 p_j}{\partial q_j \partial q_k}$ are unrestricted and can thus be arbitrarily negative, the argument we make to establish part (a) of Proposition 6 no longer goes through. As this issue of cross-partial derivatives in the advertiser demand system is quite apart from the paper's main focus, we leave it for future research.

5.2 Delegating Control to an Ad Agency

We now consider the potential impact of an advertising agency with significant say over the websites' joint advertising policies. If, rather than choosing their advertising strategies independently, the websites were to delegate ad placement decisions to a single agency acting on the collective behalf of

[21] See, for instance, Proposition 17.F.3 of Mas-Colell, Whinston, and Green (1995).

advertisers, could this be beneficial to websites? Seen from a high level, this question is not obvious, because, on the one hand, such an agency would share websites' broad incentive to draw in users. However, on the other hand, the agency and the websites have opposing incentives regarding the distribution of revenue that is generated from users' transactions with advertisers.

To model this, we readopt the simplifying "separate pools" assumption on advertiser demand, allowing us to write $p_j = p_j\left(q_j\right)$. We assume that the ad agency seeks to maximize total advertiser profits, capturing, in a broad-brush way, the idea that the agency acts on behalf of advertisers' aggregate interests, potentially because it receives a percentage of their profits. (Our discussion below, however, sheds light on the issue of what would happen if the ad agency favored some advertisers over others.) It thus has objective function

$$U(\boldsymbol{q}) = \left(\int_0^{q_s} p_s(x)\,dx - p_s(q_s)\,q_s + \int_0^{q_w} p_w(x)\,dx - p_w(q_w)\,q_w\right) D\left(\delta(\boldsymbol{q})\right). \tag{10}$$

Proposition 8 regards the outcome that would occur if the ad agency were allowed to set the number of ad slots on each website, potentially subject to a constraint on the total distraction level.

Proposition 8. *Holding fixed the total distraction level, if the ad agency is allowed to choose the allocation of slots across sites, this allocation, $\tilde{\boldsymbol{q}}^{\boldsymbol{U}}$, satisfies $-\tilde{q}_s^U p_s'\left(\tilde{q}_s^U\right)\big/\,\delta_s'\left(\tilde{q}_s^U\right) = -\tilde{q}_w^U p_w'\left(\tilde{q}_w^U\right)\big/\,\delta_w'\left(\tilde{q}_w^U\right)$.*

The following special cases help illustrate the economic meaning of Proposition 8.

Example: Linear Demand. Suppose that $p_j\left(q_j\right) = \theta\left(1 - q_j/m_j\right)$. That is, advertiser demand across the two sites share the same intercept with the vertical axis but may have different slopes. Then, the ad agency's solution to the constrained problem described in Proposition 8, $\tilde{\boldsymbol{q}}^{\boldsymbol{U}}$, features no misplacement if and only if $\delta_s'\left(\tilde{q}_s^U\right) = \delta_w'\left(\tilde{q}_w^U\right)$. This follows from the fact that, in this setting, $q_s p_s'(q_s)/q_w p_w'(q_w) = m_w q_s/m_s q_w$, whereas $r_s'(q_s)/r_w'(q_w) = (m_w q_s - 1/2)/(m_s q_w - 1/2)$. The key feature of this setting is, thus, as follows. As discussed above, at equilibrium of the game in which the websites act non-cooperatively, misplacement may be driven by asymmetries in websites' advertising technology and/or by asymmetries in advertiser demand. Under this form of advertiser demand, if asymmetries in advertising technology are minimal, then delegating control to the ad agency limits the damage that could arise from asymmetries in advertiser demand.

Example: Constant Elasticity Demand. Suppose that $p_j\left(q_j\right) = \alpha_j q_j^{-\beta_j}$, where $\alpha_j > 0$, and $\beta_j \in (0, 1)$. Then, $\tilde{\boldsymbol{q}}^{\boldsymbol{U}}$ features no misplacement if and only if $\beta_s = \beta_w$. This follows from the fact that, under this specification, $q_s p_s'(q_s) / q_w p_w'(q_w) = (\alpha_s/\alpha_w) \beta_s/\beta_w$ and $r_s'(q_s) / r_w'(q_w) = (\alpha_s/\alpha_w)(\beta_s - 1) / (\beta_s - 1)$. Note that, unlike in the linear example, here, there is no requirement that the advertising technologies be symmetric. That is, when the two sites' advertiser demand curves have symmetric, constant elasticity, the ad agency has precisely the right incentives, from the industry's point of view, to eliminated misplacement. This translates into the broader point that the more advertiser demand curves exhibit similar elasticities, the more reliable the ad agency will be at eliminating misplacement that arises from asymmetric advertising technologies across sites.

We now turn, in Proposition 9, to the ad agency's impact on double marginalization.

Proposition 9. *For a given set of primitives, consider the equilibrium outcome, $\boldsymbol{q}^*$, the industry's optimum, $\boldsymbol{q}^{\boldsymbol{\Pi}}$, and the ad agency's optimum, $\boldsymbol{q}^{\boldsymbol{U}}$. Recall from Proposition 6*(a) *that equilibrium features double marginalization; that is, $\delta\left(\boldsymbol{q}^{\boldsymbol{\Pi}}\right) < \delta(\boldsymbol{q}^*)$. The level of distraction, $\delta\left(\boldsymbol{q}^{\boldsymbol{U}}\right)$, arising under the ad agency's optimum may be ranked 1st, 2nd or 3rd.*

To appreciate an ad agency's impact on double marginalization, it is easiest to concentrate on a setup in which the sites are fully symmetric and, thus, equilibrium features no misplacement. We do this in the following example.

Example: Variable Tail Advertiser Demand. Suppose $D(\delta) = 1 - \delta$, that $\delta(\boldsymbol{q}) = q_s^2 + q_w^2$, that $c_s = c_w = 0$, and that $m_s = m_w = 2$. For advertisers, we adopt the following demand

$$p\left(q_j\right) = \begin{cases} \psi & 0 \le \frac{q_j}{m_j} \le 0.01 \\ 1 - \frac{q_j}{m_j} & 0.01 < \frac{q_j}{m_j} < 1 \end{cases},$$

where $\psi \ge 1$. Figure 4 illustrates the results.

The crucial feature of this example is the parameter ψ, capturing the willingness to pay, per user, of the "top 1%" of advertisers.[22] Variation in ψ has no impact on the equilibrium outcome, $\boldsymbol{q}^*$, or on the industry optimum, q^{π} (as long as ψ does not grow too large), because these advertisers are

[22]This specification provides a convenient way to mimic the effects of varying the thickness of the right tail of the distribution of advertisers' willingness to pay while holding the rest of this distribution fixed.

always infra-marginal. From the ad agency's perspective, however, ψ has a significant impact. On the one hand, if ψ is close to one, then the top 1% advertisers have relatively similar willingness to pay compared to the rest of the pool. This leads the agency to prioritize low ad prices, leading to the presence of many ads and a high level of distraction. On the other hand, as ψ grows larger, then the interests of the top 1% diverge more from those of the rest of the pool, in favor of attracting many users to the site. Taking this priority into account, the ad level chosen by the agency decreases until, eventually, it falls below, first, the equilibrium level and, then, the industry optimum.

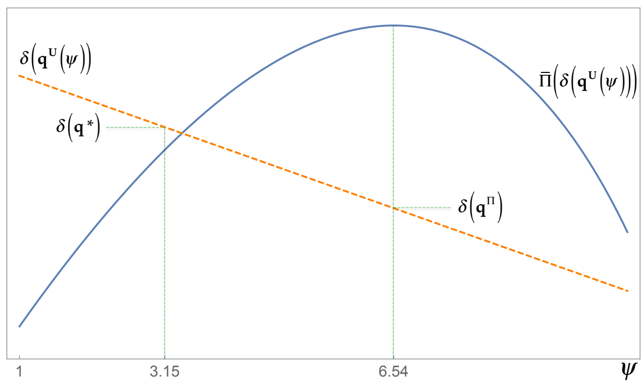


Figure 4: The ad agency's chosen level of distraction and the resulting joint profits for websites, as a function of ψ, the willingness to pay, per user, of top 1% advertisers.

Together, these examples convey the following main point. The ad agency may have an incentive to increase or reduce each of the two distortions. Consequently, if separate websites relinquish their control of ad placement decisions to such an agency, it can (but does not necessarily) increase the profits of the former.

6 Learning Sites' Ad Levels over Time

Throughout the paper, we have used a static model. Also, we have assumed that users learn how much ad-driven distraction will be on each of the complementary websites before making a single decision of whether or not to visit all of them. Thanks to the tractability of this approach, we have been able to identify the misplacement distortion, study its interaction with double marginalization and address a set of questions regarding both collective and individual incentives of the various sites.

A potential worry, however, that may come to mind is that this static, full-information approach simplifies things too much. After all, it seems obvious that most web users sometimes click on links to unfamiliar websites without knowing much about how many ads they will contain. In order to address this concern, this section builds a dynamic model in which (a) users decide sequentially, first, whether to visit the organizer and, then, whether to visit the producer, and (b) they do not learn the level of distraction on any given site until after they have visited it for the first time. We show that this model yields results that our static, full-information model approximates.

6.1 Setup of the Dynamic Model

The game has two phases. In the first phase, the organizer, s, and the producer, w, simultaneously set their ad levels. These choices, which are unobserved by users, can be interpreted as stable decisions regarding the design and layout of their webpages and how much space to devote to ads.

In the second phase there are T periods. In each of the T periods, users decide whether or not to surf. Specifically, in every period, the game proceeds as described here and illustrated in Figure 5.

1. Each user decides whether to *provisionally visit* s or to stay out. If she stays out, her payoff for the period is zero.

2. If she provisionally visits s, she learns its distraction level, δ_s. Upon learning δ_s, the user decides whether to *engage* with s or to immediately *disengage* from surfing. If she engages, she incurs distraction δ_s and obtains access to a link to w. If she disengages, her payoff for the period is zero.

3. If the user engages with s, she follows the link to w to provisionally visit it and learn δ_w. Upon learning δ_w, the user decides whether to engage with w. If she disengages from w, her payoff for the period is $-\delta_s$. If she engages with w, her payoff for the period is $v - \delta_s - \delta_w$.

Allowing users to provisionally visit the sites captures the idea that, when a user opens a webpage that obviously has too many ads to be worthwhile, she can quickly close it without paying attention to its ads or content. This assumption, which seems intuitively reasonable, is theoretically appealing because it prevents a hold up problem analogous to the one identified by Diamond (1971), where lack of information about prices (or ad levels) leads to total market breakdown. Comparable results can be derived without such an assumption if, instead, we truncate the maximum revenue per user that a site can earn. We further assume that, once, in a given period, a user learns the distraction level of a particular site, then she can recall this for all remaining periods. For simplicity, we assume there is no discounting across periods. In this incomplete information setting, our solution concept is Perfect Bayesian Equilibrium.

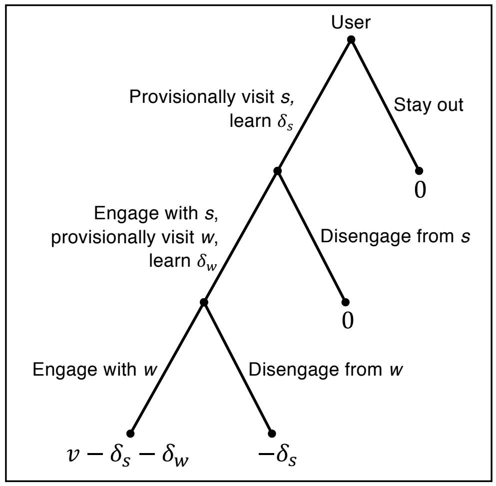


Figure 5: The decisions facing a user surfing for the first time.

6.2 Equilibrium

We now characterize the equilibria of this game. First we consider the case where $T = 1$.

Proposition 10. *In the one-period version of the game, the following statements hold.*

(a) At any equilibrium in which the organizer advertises at all, no users surf (that is, the market unravels).

(b) At any equilibrium in which some users surf, (i) the organizer's ad level is zero, (ii) the producer best-responds, setting $a_w^*(0)$, and (iii) if, off the equilibrium path, users observe any positive amount of advertising on the organizer's site, then they infer that the producer's ad level is prohibitively high, and they all stop surfing.

Proposition 10 strengthens the point made throughout the paper that sites face a tradeoff between solving the misplacement and double marginalization problems. Here, however, such a tradeoff emerges even though there is no competition, either latent or explicit, among sites of a particular type. Consider part (a). The complete market breakdown occurring in the set of equilibria described there represents an extreme version of double marginalization. Since w cannot observably commit to a given ad level before users have incurred distraction from the organizer's ads, δ_s, at the time when users do learn δ_w, organizer-driven distraction δ_s is already a sunk cost for them, and w thus sets its ad level accordingly. Consequently, in such equilibria, no users are willing to incur even a tiny level of distraction from the organizer, and they all choose not to surf.

Now consider part (b). In this type of equilibrium, users' very particular beliefs eliminate an incentive for s to set a positive ad level, thus eliminating double marginalization. Here, however, as highlighted by the following remark, the organizer's advertising technology goes unused.

Remark 3. *An equilibrium of the type described in Proposition 10*(b) *features misplacement so long as* $\delta_s'(0) < \delta_w'(a_w^*(0))$.

We now turn to the case where there are multiple surfing periods. This case represents the likely more realistic scenario where, after sites' basic design decisions have been taken and ad policies have been set, users who find these to be acceptable continue to surf on a regular basis. To discipline the model, we make use of the following, standard definition.[23]

Definition 3. *We say users have* passive beliefs *when their beliefs regarding the producers ad level are unaffected by their observation of the organizer's ad level.*

Proposition 11. *Assume users have passive beliefs. When* $T \geq 2$, *the unique equilibrium features ad levels* $a_s^* = \sqrt{\frac{10(T-1)}{4T-3}}$ *and* $a_w^* = \sqrt{\frac{10T\omega}{4T-3}}$. *As* T *grows to infinity, these approach from below and above, respectively, their counterparts in the static, full-information version of the model,* $\sqrt{\frac{5}{2}}$ *and* $\sqrt{\frac{5\omega}{2}}$.

As T grows large, this equilibrium mimics the one that we analyze in the static, full-information setup used throughout the paper. Thus, taken together, Propositions 10 and 11 show that, although asymmetry of information and dynamics may raise a set of additional subtleties, the simplifications regarding timing and information that we make in the rest of the paper do not, themselves, drive our main findings regarding misplacement and double marginalization.

[23]See Hart and Tirole (1990), which, to the best of our knowledge, is the first paper to make such an assumption in a model of vertically related firms and McAfee and Schwartz (1994), which introduces the term "passive beliefs."

7 Conclusion

Internet users often visit ad-funded *organizers* in order to access content hosted by separate, ad-funded *producers*. In this paper, we show that the platforms involved in such situations face incentives that distort their advertising policies in two ways. The first distortion is the classic problem of double marginalization, reflected by an excessive total amount of advertising. The second distortion is what we call misplacement. It entails the misallocation of nuisance or "demands for attention" across these two different types of websites. This pair of distortions lead to a tradeoff that represents a significant departure from traditional models of this type. In the standard setting, intense competition at one level of a vertical supply chain restores the first-best for the industry as a whole. Here, on the other hand, more competition among producers leads to more misplacement.

Consequently, a moderate level of competition among producers is optimal for the industry as a whole. This tradeoff also has interesting implications for the optimal strategy of an organizer that is deciding which types of content to provide on its own platform and which types of content to link to externally. In two more extreme settings – (a) where content is mainly homogeneous and (b) where it is mainly driven by exclusive "scoops," – content producers face either very intense or very low levels of competition and thus advertise in a way that is highly inefficient. In either of these two cases, organizers have an incentive to acquire content for themselves and take control of advertising that appears alongside it. In contrast, when content is only partially scoop-driven, content producers' advertising is relatively efficient, making it better for the organizer to follow a linking strategy that gives more flexibility to reliably direct users to the freshest content.

Perhaps no current concern facing public policymakers (including antitrust and legislative authorities) is more pressing than the growth of power among a handful of large online platforms. Broadly speaking, in this debate, a particular challenge has been to identify the costs and benefits of centralization, both for the platforms themselves and for society as a whole. By identifying a novel tradeoff influencing platforms' decisions regarding when to (de-)centralize their operations, this paper hopes to contribute to our understanding of this complex and important set of issues.

References

Anderson, Simon P. and Stephen Coate (2005), "Market Provision of Broadcasting: A Welfare Analysis." *Review of Economic Studies*, 72, 947–972.

Anderson, Simon P and Andre De Palma (2009), "Information Congestion." *The RAND Journal of Economics*, 40, 688–709.

Anderson, Simon P. and André de Palma (2012), "Competition for Attention in the Information (Overload) Age." *RAND Journal of Economics*, 43, 1–25.

Anderson, Simon P, Øystein Foros, and Hans Jarle Kind (2017), "Competition for Advertisers and for Viewers in Media Markets." *Economic Journal*, 128, 34–54.

Athey, Susan, Emilio Calvano, and Joshua S Gans (2018), "The Impact of Consumer Multi-Homing on Advertising Markets and Media Competition." *Management Science*, 64, 1574–1590.

Berger, Paul D. (1972), "Vertical Cooperative Advertising Ventures." *Journal of Marketing Research*, 9, 309–312.

Burguet, Roberto, Ramon Caminal, and Matthew Ellman (2015), "In Google we Trust?" *International Journal of Industrial Organization*, 39, 56–70.

Cao, Xinyu and T. Tony Ke (Forthcoming), "Cooperative Search Advertising." *Marketing Science*.

Casadesus-Masanell, Ramon, Barry J. Nalebuff, and David Yoffie (2007), "Competing Complements." *Mimeo.*, URL `https://ssrn.com/abstract=1032461`.

Cheng, Leonard K. and Jae Nahm (2007), "Product Boundary, Vertical Competition, and the Double Mark-Up Problem." *RAND Journal of Economics*, 38, 447–466.

Chiou, Lesley and Catherine Tucker (2017), "Content Aggregation by Platforms: The Case of the News Media." *Journal of Economics & Management Strategy*, 26, 782–805.

Choi, Hanna, Carl F. Mela, Santiago Balseiro, and Adam Leary (2017), "Online Display Advertising Markets: A Literature Review and Future Directions." *Mimeo*, URL `https://ssrn.com/abstract=3070706`.

Choi, Jay Pil (2006), "Broadcast Competition and Advertising with Free Entry: Subscription vs. Free-to-Air." *Information Economics and Policy*, 18, 181 – 196.

Cournot, Augustin (1838), *Researches into the Mathematical Principles of the Theory of Wealth* (N.T. Bacon trans.), 1971 edition. Augustus M. Kelley, New York.

Crampes, Claude, Carole Haritchabalet, and Bruno Jullien (2009), "Advertising, Competition and Entry in Media Industries." *Journal of Industrial Economics*, LVII, 7–31.

Daughety, Andrew F (2008), "Cournot Competition." *New Palgrave Dictionary of Economics: Volume 1–8*, 1197–1203.

de Cornière, Alexandre (2016), "Search Advertising." *American Economic Journal: Microeconomics*, 8, 156–188.

de Cornière, Alexandre and Greg Taylor (2014), "Integration and search engine bias." *RAND Journal of Economics*, 45, 576–597.

Dellarocas, Chrysanthos (2012), "Double Marginalization in Performance-Based Advertising: Implications and Solutions." *Management Science*, 58, 1178–1195.

Diamond, Peter A. (1971), "A Model of Price Adjustment." *Journal of Economic Theory*, 3, 156–168.

Eliaz, Kfir and Ran Spiegler (2011), "A Simple Model of Search Engine Pricing." *Economic Journal*, 121, 329–339.

Eppler, Martin J. and Jeanne Mengis (2004), "The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines." *The Information Society*, 20, 325–344.

Evans, David S. (2008), "The Economics of the Online Advertising Industry." *Review of Network Economics*, 7, 359–391.

Evans, David S. (2009), "The Online Advertising Industry: Economics, Evolution, and Privacy." *Journal of Economic Perspectives*, 23, 37–60.

Gabszewicz, Jean J., Didier Laussel, and Nathalie Sonnac (2004), "Programming and Advertising Competition in the Broadcasting Industry." *Journal of Economics & Management Strategy*, 13, 657–669.

Gentzkow, Matthew (2007), "Valuing New Goods in a Model with Complementarity: Online Newspapers." *American Economic Review*, 97, 713–744.

Goldfarb, Avi and Catherine Tucker (2011), "Online Display Advertising: Targeting and Obtrusiveness." *Marketing Science*, 30, 389–404.

Gomes, Renato (2014), "Optimal Auction Design in Two-Sided Markets." *RAND Journal of Economics*, 45, 248–272.

Hagiu, Andrei and Bruno Jullien (2011), "Why Do Intermediaries Divert Search?" *RAND Journal of Economics*, 42, 337–362.

Hart, Oliver and Jean Tirole (1990), "Vertical Integration and Market Foreclosure." *Brookings Papers on Economic Activity: Microeconomics*, 1990, 205–286.

Jeon, Doh-Shin and Nikrooz Nasr (2016), "News Aggregators and Competition among Newspapers on the Internet." *American Economic Journal: Microeconomics*, 8, 91–114.

Jørgensen, Steffen and Georges Zaccour (2014), "A Slurvey of Game-Theoretic Models of Cooperative Advertising." *European Journal of Operational Research*, 237, 1–14.

Katona, Zsolt and Miklos Sarvary (2008), "Network Formation and the Structure of the Commercial World Wide Web." *Marketing Science*, 27, 764–778.

Lerner, Josh and Jean Tirole (2004), "Efficient Patent Pools." *American Economic Review*, 94, 691–711.

Lerner, Josh and Jean Tirole (2015), "Standard-Essential Patents." *Journal of Political Economy*, 123, 547–586.

Liu, Yu-Hsin (2018), "The Impact of Consumer Multi-homing Behavior on Ad Prices: Evidence from an Online Marketplace." *Mimeo.*, URL https://sites.google.com/umail.iu.edu/yuhsinliu.

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green (1995), *Microeconomic Theory*. Oxford University Press, New York.

McAfee, R Preston and Marius Schwartz (1994), "Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity, and Uniformity." *American Economic Review*, 210–230.

Peitz, Martin and Tommaso M. Valletti (2008), "Content and Advertising in the Media: Pay-TV versus Free-to-Air." *International Journal of Industrial Organization*, 28, 949–965.

Pilkington, Ed (2018), "How the Drudge Report Ushered in the Age of Trump." *The Guardian*, URL https://www.theguardian.com/us-news/2018/jan/24/how-the-drudge-report-ushered-in-the-age-of-trump.

Rey, Patrick and Jean Tirole (1986), "The Logic of Vertical Restraints." *American Economic Review*, 76, 921–939.

Schwartz, Marius (1989), "Investments in Oligopoly: Welfare Effects and Tests for Predation." *Oxford Economic Papers*, 41, 698–719.

Shleifer, Andrei and Robert W Vishny (1993), "Corruption." *Quarterly Journal of Economics*, 108, 599–617.

Spengler, Joseph J (1950), "Vertical Integration and Antitrust Policy." *Journal of Political Economy*, 58, 347–352.

Van Zandt, Timothy (2004), "Information Overload in a Network of Targeted Communication." *RAND Journal of Economics*, 542–560.

White, Alexander (2013), "Search Engines: Left Side Quality versus Right Side Profits." *International Journal of Industrial Organization*, 31, 690–701.

Appendices

A Proofs and Second-Order Conditions

Proof of Proposition 1 . (a) We first show that $\delta(\boldsymbol{a}^*) > \delta\left(\boldsymbol{a}^{\boldsymbol{\Pi}}\right)$. Let $a_j\left(\delta_j\right)$ denote the inverse function of $\delta_j\left(a_j\right)$, and rewrite equation (3) as

$$\overline{a}^* - \overline{c} = h(\delta^*) \sum_{j=1}^{n} a_j'\left(\delta_j^*\right). \tag{11}$$

Note that, in equation (11), the LHS measures the (net) marginal gain, summed across all sites, of serving one more user, and the RHS measures the infra-marginal loss, summed across sites, from the same perturbation. Now consider the reallocation of individual sites' ad levels that maximizes ad revenue, while holding fixed δ^* and thus total demand. We show that, following such a reallocation, for the industry as a whole, the marginal gain of adding another user strictly exceeds the infra-marginal loss.

Following the reallocation of ad levels, the industry's marginal gain from adding another user becomes $k + \overline{a}^* - \overline{c}$, for some $k \geq 0$, and its infra-marginal loss is $h(\delta^*)\,\hat{a}'$, for some $\hat{a}' > 0$, capturing the rate, common across all sites, at which distraction is transformed into revenue, following the reallocation. Moreover, $\hat{a}' \leq \max_j \left\{a_j'\left(\delta_j^*\right)\right\} < \sum_{j=1}^{n} a_j'\left(\delta_j^*\right)$. The former inequality holds because $a_j'' < 0$, $\forall j$, and, following the reallocation, any sites in the set $\arg\max_j \left\{a_j'\left(\delta_j^*\right)\right\}$ must have (weakly) increased their ad levels. The latter inequality holds because, at any Nash equilibrium, $\delta_j^* > 0$, $\forall j$. Thus, the analog to equation (1), evaluated at the vector of post-reallocation ad levels, $\hat{\boldsymbol{a}}$, is $\overline{\hat{a}} - \overline{c} > h(\delta(\hat{\boldsymbol{a}}))/\delta_j'(\hat{\boldsymbol{a}})$, $\forall j$, and our claim is proved.

(b) The claim about misplacement is straightforward.

□

Proof of Proposition 2. (a) The Nash Equilibrium advertising level chosen by w satisfies $\tilde{a}_w - c_w = -\frac{D}{D'\delta_w'}$. Plugging this into the right-hand side of equation (5) gives

$$\overline{\Pi}'(\tilde{a}_w) = -D \cdot \left(\delta_s' a_s^{*\prime} + \frac{\delta_w'}{\delta_s'}\right). \tag{12}$$

Therefore, an exogenous decrease in a_w leads to an increase in $\overline{\Pi}$ if and only if the expression in (12) is negative, which requires that the sum of the terms in brackets be positive. Moreover, $a_s^{*\prime} = h'/(\delta_s'(1-h') + h\delta_s''/\delta_s')$. Plugging in this expression for $a_s^{*\prime}$ and rearranging gives the stated result.

(b) We have

$$\overline{\Pi}'(c_w) = D \cdot \left(1 - \frac{\delta_w'(c_w)}{\delta_s'(a_s^*(\delta_w(c_w)))}\right),$$

and thus $\overline{\Pi}'(c_w) > 0 \Leftrightarrow \delta_w'(c_w) < \delta_s'(a_s^*(\delta_w(c_w)))$.

□

Proof of Proposition 3. Since we focus on symmetric equilibrium under linking mode, and acquisition mode is also symmetric for the l producers, we use w to denote each of the l producers. When $\rho = 0$, under acquisition mode, the organizer sets a_s, a_w to maximize

$$\Pi^A(a_s, a_w) = (a_s + a_w) D\left(\frac{l}{m}(a_s^2 + a_w^2/\omega)\right) = (a_s + a_w)\left(1 - \frac{(a_s^2 + a_w^2/\omega)l}{10m}\right).$$

Thus the organizer sets

$$a_s^* = \sqrt{\frac{10m}{3l(1+\omega)}}, \;\; a_w^* = \omega\sqrt{\frac{10m}{3l(1+\omega)}},$$

which brings in profits

$$\Pi^{A*} = \Pi^A(a_s^*, a_w^*) = \frac{2}{3}\sqrt{\frac{10m}{3l}(1+\omega)}.$$

Under linking mode, the profits of s and w, are, respectively,

$$\Pi_s^L(a_s, a_w) = a_s D(a_s^2 + a_w^2/\omega) = a_s\left(1 - (a_s^2 + a_w^2/\omega)/10\right),$$
$$\Pi_w^L(a_s, a_w) = \frac{1}{l} a_w D(a_s^2 + a_w^2/\omega) = \frac{1}{l} a_w\left(1 - (a_s^2 + a_w^2/\omega)/10\right).$$

Their incentives to choose ad level are the same as in the "bilateral monopoly" case in Section 3. The equilibrium outcome is

$$a_s^* = \sqrt{\frac{5}{2}}, \;\; a_w^* = \sqrt{\frac{5\omega}{2}},$$
$$\Pi_s^{L*} = \Pi_s^L(a_s^*, a_w^*) = \frac{1}{2}\sqrt{\frac{5}{2}},$$
$$\Pi_w^{L*} = \Pi_w^L(a_s^*, a_w^*) = \frac{1}{2l}\sqrt{\frac{5\omega}{2}}.$$

The condition that linking mode is preferable compared to acquisition mode is $\Pi_s^{L*} \geq \Pi^{A*}$, which can be reduced to $\frac{m}{l} \leq \frac{27}{64(1+\omega)}$. □

Proof of Propositions 4 and 5. To establish the stated results, we first outline the logic behind Proposition 4. Then we lay out the formal proof and calculations for Propositions 4 and 5 and Fig. 3.

Under linking mode, the strategies of producers affect the profits of the organizer only by entering the demand function of the organizer. When both producers get the scoop, the organizer picks the one with fewer ads. When only one gets the scoop, the organizer picks that one. As the demand we specify is linear in total distraction δ, what matters for the organizer is just the expected distraction of the producer it picks, which is $\delta_0 \equiv q\mathbb{E}\left[\min_{\{w1,w2\}}\left\{\delta_j\right\}\right] + (1-q)\,\mathbb{E}\left[\delta_j\right]$. The latter part of Proposition 4(b) comes as a result of Proposition 4(a). The organizer plays a pure strategy $a_s^*(\delta_0)$ as best response, and $a_s^*(\cdot)$ is a strictly decreasing function.

From the perspective of producers, taking the organizer's ad level a_s^* as given, the profits of each producer is the sum of the profits it will earn in two regimes: the *exclusive regime* and the *competitive regime*, where, by the former we mean the case in which it gets the exclusive scoop, and by the latter we mean the the case in which both producers get the scoop. This is because the producer earns zero profits when only its rival gets the scoop, as the organizer will pick its rival. In the exclusive regime, each producer would optimally set its ad level in the manner of bilateral monopoly with the organizer. The competitive regime resembles Bertrand competition between two producers, as the two producers are incentivized to undercut each other. The producer's strategy takes into account the incentives it faces under both of these regimes.

Suppose by contradiction the producer's symmetric equilibrium strategy is a pure strategy or a mixed strategy with some mass point in the support. In this case, one producer can improve its payoff by marginally lowering this mass point, so as to undercut its rival in the competitive regime while keeping its profits in the exclusive regime almost unchanged. Thus we conclude that each producer must play a mixed strategy with an atomless distribution on the support in the symmetric equilibrium, which is the first part of Proposition 4(a).

To see that $0 < \underline{a} < \overline{a} < a_w^*$, first we suppose by contradiction $\underline{a} = 0$. When the realization of the ad level, a_w, is 0, the producer earns zero profits. But it can earn positive profits under the exclusive regime by setting a positive ad level, which contradicts $\underline{a} = 0$. To see why $\overline{a} < a_w^*$, we start from the bilateral monopoly equilibrium, where the producers set their ad levels at a_w^* and the organizer best responds by choosing $a_s^*((a_w^*)^2/\omega)$. Keeping the organizer's $a_s^*((a_w^*)^2/\omega)$ unchanged, suppose by contradiction that, in this mixed strategy equilibrium, we have $\overline{a} > a_w^*$. Then the producer could improve its payoff by playing a_w^* instead of $\overline{a}$ against its rival's randomization over $[\underline{a}, \overline{a}]$. This is because it earns strictly higher profits under the exclusive regime by choosing a_w^*, and it earns nonnegative profits under the competitive regime as opposed to zero profits under competition regime if it chooses $\overline{a}$. Hence $\overline{a} \leq a_w^*$. Randomizing over $[\underline{a}, \overline{a}]$ thus yields lower expected distraction compared to the case in which both producers play a_w^* with probability one, which exerts upward pressure on the organizer's choice of a_s. A larger a_s in turn makes the best response of each producer in the exclusive regime lower, which means $\overline{a}$ is lower than a_w^*, completing Proposition 4(a). Note that this adjusting-to-equilibrium argument holds only when there is a unique symmetric equilibrium, which we will show in the calculations below.

Below we present the formal proof and calculations.

With $\delta_0^* = q\mathbb{E}\left[\min_{\{w1,w2\}}\left\{\delta_j\right\}\right] + (1-q)\,\mathbb{E}\left[\delta_j\right]$, the organizer sets $a_s^* = \sqrt{\frac{10-\delta_0^*}{3}}$. For the producer to best respond to a_s^* in the exclusive regime, the producer sets $a_w = \sqrt{\frac{\omega}{3}(10-(a_s^*)^2)}$. Now we argue $\overline{a} = \sqrt{\frac{\omega}{3}(10-(a_s^*)^2)}$. Suppose by contradiction $\overline{a} \neq \sqrt{\frac{\omega}{3}(10-(a_s^*)^2)}$, then when playing against a random strategy over $[\underline{a}, \overline{a}]$, the producer earns higher profits by playing $\sqrt{\frac{\omega}{3}(10-(a_s^*)^2)}$ instead of $\overline{a}$. This is because $\sqrt{\frac{\omega}{3}(10-(a_s^*)^2)}$ brings in strictly higher profits in the exclusive regime than does $\overline{a}$, and it brings in nonnegative profits in the competitive regime, in which $\overline{a}$ gives rise to zero profits.

We denote the profits earned by each producer at equilibrium as $\widetilde{\Pi}$, which should be equal to the profits it earns from any realized ad level in the support, $[\underline{a}, \overline{a}]$, of its strategy, including the case in which the realized ad level is exactly $\overline{a} = \sqrt{\frac{\omega}{3}(10-(a_s^*)^2)}$. By choosing a, the probability that this producer is picked by the organizer is $q(1-F_w^*(a)) + \frac{1-q}{2} = \frac{1+q}{2} - qF_w^*(a)$. Thus,

$$\widetilde{\Pi} = \left(\frac{1+q}{2} - qF_w^*(a)\right) \cdot a\left(1 - ((a_s^*)^2 + a^2/\omega)/10\right), \forall a \in [\underline{a}, \overline{a}].$$

As we have shown that $\overline{a} = \sqrt{\frac{\omega}{3}(10-(a_s^*)^2)}$, the indifference condition above gives

$$F_w^*(a) = \frac{1}{2q}\left((1+q) - \frac{(1-q)\cdot\frac{10-(a_s^*)^2}{15}\cdot\sqrt{\frac{\omega}{3}(10-(a_s^*)^2)}}{a\,(1-((a_s^*)^2 + a^2/\omega)/10)}\right).$$

To find $\underline{a}$, we set $F_w^*(\underline{a}) = 0$, which implies

$$\underline{a}\left(1 - ((a_s^*)^2 + \underline{a}^2/\omega)/10\right) = \frac{1-q}{1+q}\cdot\frac{10-(a_s^*)^2}{15}\cdot\sqrt{\frac{\omega}{3}(10-(a_s^*)^2)}.$$

Define $\eta \equiv \underline{a}/\overline{a}$ and thus $\underline{a} = \eta \cdot \sqrt{\frac{\omega}{3}(10-(a_s^*)^2)}$. Plugging this into the equation above yields

$$\eta(3-\eta^2) = \frac{2(1-q)}{1+q}. \tag{13}$$

Note that $\eta(3-\eta^2)$ is increasing on $[0,1]$ and ranges from 0 to 2. Hence, for any given q, there exists a unique η that solves the equation.

Now we solve for the equilibrium.

$$\begin{aligned}
\delta_0^* &= q\mathbb{E}\left[\min_{\{w1,w2\}}\left\{\delta_j\right\}\right] + (1-q)\,\mathbb{E}\left[\delta_j\right] \\
&= \frac{1}{\omega}\left((1+q)\int_{a_{w1}<a_{w2}} a_{w1}^2 dF_w^*(a_{w1})dF_w^*(a_{w2}) + (1-q)\int_{a_{w1}>a_{w2}} a_{w1}^2 dF_w^*(a_{w1})dF_w^*(a_{w2})\right) \\
&= (1+q)\int_{a_{w1}<a_{w2}} a_{w1}^2 dF_w^*(\sqrt{\omega}a_{w1})dF_w^*(\sqrt{\omega}a_{w2}) + (1-q)\int_{a_{w1}>a_{w2}} a_{w1}^2 dF_w^*(\sqrt{\omega}a_{w1})dF_w^*(\sqrt{\omega}a_{w2}) \\
&= \int_{\eta\cdot\bar{a}/\sqrt{\omega}}^{\bar{a}/\sqrt{\omega}} \Big((1+q)(1-F_w^*(\sqrt{\omega}a_{w1})) + (1-q)F_w^*(\sqrt{\omega}a_{w1})\Big)\cdot a_{w1}^2 dF_w^*(\sqrt{\omega}a_{w1}) \\
&= \frac{(1-q)^2}{27q}\cdot\left(\left(\frac{3\eta^2}{(3-\eta^2)^2} + \ln\frac{3-\eta^2}{\eta^2}\right) - \left(\frac{3}{4}+\ln 2\right)\right)\cdot(10-(a_s^*)^2).
\end{aligned}$$

Let $\zeta = \frac{(1-q)^2}{27q}\cdot\left(\left(\frac{3\eta^2}{(3-\eta^2)^2} + \ln\frac{3-\eta^2}{\eta^2}\right) - (\frac{3}{4}+\ln 2)\right)$, which depends only on q. We have

$$\delta_0^* = \zeta\cdot(10-(a_s^*)^2).$$

Solving the condition above together with $a_s^* = \sqrt{\frac{10-\delta_0^*}{3}}$ yields

$$a_s^* = \sqrt{\frac{10(1-\zeta)}{3-\zeta}}.$$

From the definition of δ_0^*, it is apparent that $\delta_0^* \geq 0$ and thus $\zeta \geq 0$. In addition, if we can show that $\zeta \leq 1$, then according to equation above, there exists a unique equilibrium. In fact, we can prove a stronger statement, $\zeta \leq \frac{1}{3}$. We do this by first showing $\lim_{q\to 0}\zeta = \frac{1}{3}$ and then showing $\frac{\partial\zeta}{\partial q} < 0$.

For the first part, according to equation (13) and the implicit function theorem, $\frac{\partial\eta}{\partial q} = \frac{-4}{3(1+q)^2(1-\eta^2)}$ and η is continuously differentiable in q. Since $\eta = 1$ when $q = 0$, we have $\lim_{q\to 0}\eta = 1$. Using L'Hôpital's rule, $\lim_{q\to 0}\zeta = \lim_{q\to 0}\frac{(1-q)^2}{27}\cdot\frac{\left((\frac{3\eta^2}{(3-\eta^2)^2}+\ln\frac{3-\eta^2}{\eta^2}) - (\frac{3}{4}+\ln 2)\right)}{q} = \frac{1}{27}\cdot\lim_{q\to 0}\frac{\partial(\frac{3\eta^2}{(3-\eta^2)^2}+\ln\frac{3-\eta^2}{\eta^2})}{\partial q} = \lim_{q\to 0}\frac{8}{3(1+q)^2\eta(3-\eta^2)^3} = \frac{1}{3}$.

For the latter part, rearranging equation (13) yields $q = \frac{2-3\eta+\eta^3}{2+3\eta-\eta^3}$. Plugging this into the expression of ζ leads to $\zeta(\eta) = \frac{(1-\frac{2-3\eta+\eta^3}{2+3\eta-\eta^3})^2}{27\cdot\frac{2-3\eta+\eta^3}{2+3\eta-\eta^3}}\cdot\left((\frac{3\eta^2}{(3-\eta^2)^2}+\ln\frac{3-\eta^2}{\eta^2}) - (\frac{3}{4}+\ln 2)\right)$. Thus we have $\frac{\partial\zeta}{\partial q} = \zeta'(\eta)\cdot\frac{\partial\eta}{\partial q}$. Recall that $\frac{\partial\eta}{\partial q} = \frac{-4}{3(1+q)^2(1-\eta^2)} < 0$. Hence, it suffices to show that $\zeta'(\eta) \geq 0$ when $\eta \in [0,1]$.

We have $\zeta'(\eta) = \dfrac{8\eta\left((30+4\ln 2)\eta^2 - 9\eta^4 + 4(3-\eta^2)\ln\frac{3-\eta^2}{\eta^2} - (21+12\ln 2)\right)}{9(4-\eta^2)^2(1-\eta^2)^3}$. To simplify the equation, now define $\iota(\eta) = (30+4\ln 2)\eta^2 - 9\eta^4 + 4(3-\eta^2)\ln\frac{3-\eta^2}{\eta^2} - (21+12\ln 2)$ and then $\zeta'(\eta)$ is simplified into $\dfrac{8\eta\cdot\iota(\eta)}{9(4-\eta^2)^2(1-\eta^2)^3}$. Thus it is sufficient to show that $\iota(\eta) \geq 0$. We do this by checking the derivatives of ι in $[0,1]$ up to third order. We have $\iota'''(\eta) = -\dfrac{216(\eta^4(3-\eta^2)^2 + 2(1-\eta^2))}{\eta^3(3-\eta^2)^2} \leq 0$. Then $\iota''(\eta) \geq \iota''(1)$ for $\eta \in [0,1]$. Note that $\iota''(1) = 0$ and thus $\iota''(\eta) \geq 0$, so $\iota'(\eta) \leq \iota'(1)$ for $\eta \in [0,1]$. Note, additionally, that $\iota'(1) = 0$, so $\iota'(\eta) \leq 0$ and $\iota(\eta) \geq \iota(1) = 0$ when $\eta \in [0,1]$. This completes the proof that there exists exactly one symmetric equilibrium.

Under linking mode, the profits for the organizer are thus

$$\Pi^{L*} = a_s^* \cdot \left(1 - ((a_s^*)^2) + \delta_0^*)/10\right).$$

Under acquisition mode, the profits for the organizer are

$$\Pi^{A*} = \max_{a_s, a_w} (a_s + a_w) D\left(\left(\frac{1+q}{2}\right)^{-1}(a_s^2 + a_w^2/\omega)\right) = \frac{2}{3}\sqrt{\frac{5}{3}(1+q)(1+\omega)}.$$

We compute Π^{L*} and Π^{A*} numerically, and compare them in Figure 3, from which we establish Proposition 5. □

Proof of Proposition 6. (a) We first show that $\delta(\boldsymbol{q}^*) > \delta\left(\boldsymbol{q}^{\Pi}\right)$. Consider the RHS of equations (7) and (9), characterizing the industry optimum and equilibrium, respectively. For $j = s, w$ and $k \neq j$, let $I_j\left(q_j\right) \equiv r_j'\left(q_j\right) \big/ \delta_j'\left(q_j\right)$ denote a generic "infra-marginal" term. We now show, using an analogous argument to the one given in the proof of Proposition 1, that, following a shift from $\boldsymbol{q}^*$ to $\hat{\boldsymbol{q}} \equiv \arg\max_{\boldsymbol{q}} \{r_s(q_s) + r_w(q_w) : \delta(\boldsymbol{q}) = \delta(\boldsymbol{q}^*)\}$, the total marginal gain to the industry from serving one more user strictly exceeds the infra-marginal loss.

Regarding the total marginal gain from adding another user, by the definition of $\hat{\boldsymbol{q}}$, it holds that $r_s(\hat{q}_s) + r_w(\hat{q}_w) - \bar{c} \geq r_s(q_s^*) + r_w(q_w^*) - \bar{c} = h(\delta(\boldsymbol{q}^*))(I_s(q_s^*) + I_w(q_w^*))$.

Regarding the infra-marginal loss, first note that $I_s(\hat{q}_s) = I_w(\hat{q}_w)$. It remains to show that $I_s(\hat{q}_s) \leq \max\{I_s(q_s^*), I_w(q_w^*)\} < I_s(q_s^*) + I_w(q_w^*)$. The first inequality holds because these are both decreasing functions, and, in the shift from $\boldsymbol{q}^*$ to $\hat{\boldsymbol{q}}$, either q_s or q_w must have weakly increased. The second inequality holds because $I_j\left(q_j^*\right) > 0$, for $j = s, w$.

(b) The claim about misplacement is straightforward.

□

Proof of Proposition 7. Let $\boldsymbol{q}^*$ denote a vector of equilibrium numbers of ad slots. There is no misplacement if and only if $\boldsymbol{q}^* \in \arg\max_{\boldsymbol{q}} \{r_s(\boldsymbol{q}) + r_w(\boldsymbol{q}) : \delta(\boldsymbol{q}) = \delta(\boldsymbol{q}^*)\}$. The Lagrangian for this problem is given by $r_s(\boldsymbol{q}) + r_w(\boldsymbol{q}) - \lambda\left(\delta\left(\boldsymbol{q}\right) - \delta(\boldsymbol{q}^*)\right)$, where λ is the multiplier. Differentiating the Lagrangian with respect to $\boldsymbol{q}$ gives first order conditions

$$\frac{\frac{\partial r_j(\tilde{\boldsymbol{q}}^U)}{\partial q_j} + \frac{\partial r_k(\tilde{\boldsymbol{q}}^U)}{\partial q_j}}{\delta_j'(\tilde{q}_j^U)} = \lambda, \quad j = s, w,\ k \neq j,$$

which establishes our claim.

□

Proof of Proposition 8*.* Let $\delta\left(\tilde{\boldsymbol{q}}\right)$ denote the level of total distraction to which the ad agency is constrained. Solving $\max_{\boldsymbol{q}}\ \{U\left(\boldsymbol{q}\right)$ subject to $\delta\left(\boldsymbol{q}\right) = \delta\left(\tilde{\boldsymbol{q}}\right)\}$ gives rises to first-order conditions $-q_j p_j'\left(q_j\right) \Big/ \delta_j'\left(q_j\right) = \lambda$, for $j = s, w$, $k \neq j$, where λ is the Lagrange multiplier.

□

Proof of Proposition 9*.* The example following the Proposition establishes the claim. □

Proof of Proposition 10*.* First we show that in any equilibrium, s must set $a_s^* = 0$. Then we establish for all the equilibria with $a_s^* = 0$, it must be the case that the user's off-path beliefs when they observe $a_s \neq 0$ that none of them engage in s.

Define residual demand function $D_\chi(\delta) = \min\{D(\delta), D(\chi)\}$, which is the demand w will face if only the users with valuation higher than cutoff χ, i.e. $v \geq \chi$, will observe δ_w and then decide to engage in w or not.

For any equilibrium with $a_s^* > 0$, on the equilibrium path, the users' beliefs about the ad level on w must be correct, which is a_w^*. Users with $v \geq \delta_s^* + \delta_w^*$ will engage in s and then observes $a_w = a_w^*$. From the perspective of w, w sets a_w to maximize its profits with residual demand $D_{\delta_s^*+\delta_w^*}(\delta)$. However, unless $D_{\delta_s^*+\delta_w^*}(\delta) = 0$, the unique optimal $\delta_w(a_w)$ would be no less than $\delta_s^* + \delta_w^*$, which is strictly larger than δ_w^*. This completes the proof for the first part of the proposition.

For any equilibrium with $a_s^* = 0$, suppose by contradiction that the off-path beliefs of the users when they observe $a_s^{off} > 0$ are such that $D(\delta_s^{off} + \mathbb{E}\delta_w)$ is nonzero. In that case, deviating to $a_s^{off} > 0$ brings in positive profits for s, which contradicts $a_s^* = 0$. At equilibrium, w best responds to $a_s^* = 0$.

□

Proof of Proposition 11*.* We assume the users hold so-called "passive beliefs". That is, if users observe an ad level set by the organizer that differs from its equilibrium value, they continue to believe that the producer's ad level is at the equilibrium level. For the organizer, the mass of users that will engage is always $D(\delta_s + \delta_w^*)$, when evaluating its deviation from a_s^*. For the producer, the mass of users that will engage in the first period is $D_{\delta_s^*+\delta_w^*}(\delta_w)$. However, in all subsequent periods, those users with $v \geq \delta_s^* + \delta_w^*$ would have learned the true value of δ_w, and thus the demand it would face is $D_{\delta_s^*+\max\{\delta_w^*,\delta_w\}}(\delta_w)$. Hence, the equilibrium conditions are

$$\begin{aligned} a_s^* &= \arg\max_{a_s} \quad T \cdot a_s \cdot D(\delta_s(a_s) + \delta_w(a_w^*)) \\ &= \arg\max_{a_s} \quad T a_s \left(1 - \frac{a_s^2 + (a_w^*)^2/\omega}{10}\right), \end{aligned} \tag{14}$$

and

$$\begin{aligned} a_w^* &= \arg\max_{a_w} \quad a_w \cdot D_{\delta_s^*+\delta_w^*}(\delta_w(a_w)) + (T-1) \cdot a_w \cdot D_{\delta_s^*+\max\{\delta_w^*,\delta_w(a_w)\}}(\delta_w(a_w)) \\ &= \arg\max_{a_w} \quad a_w \left[D(\max\{\delta_w(a_w), \delta_s^* + \delta_w^*\}) + (T-1)D(\delta_s^* + \max\{\delta_w^*, \delta_w(a_w)\})\right]. \end{aligned}$$

Note that by raising a_w to a level of at least a_w^*, the producer increases its revenue per user without lowering the quantity. Thus, the optimal a_w must be larger than or equal to a_w^*. Hence,

$$a_w^* = \arg\max_{a_w} \quad a_w \left[D(\max\{\delta_w(a_w), \delta_s^* + \delta_w^*\}) + (T-1)D(\delta_s^* + \delta_w(a_w))\right]. \tag{15}$$

A necessary condition for the equilibrium to hold is that the optimal a_w satisfies $\delta_w(a_w) \leq \delta_s^* + \delta_w^*$. We assume this is true to characterize the equilibrium ad levels, and then go back to verify that it is indeed optimal for w to choose $a_w = a_w^*$ instead of any a_w such that $\delta_w(a_w) > \delta_s^* + \delta_w^*$. Thus,

$$\begin{aligned} a_w^* &= \arg\max_{a_w} \quad a_w \left[D(\delta_s^* + \delta_w^*) + (T-1)D(\delta_s^* + \delta_w(a_w))\right] \\ &= \arg\max_{a_w} \quad a_w \left[1 - \frac{(a_s^*)^2 + (a_w^*)^2/\omega}{10} + (T-1)\left(1 - \frac{(a_s^*)^2 + a_w^2/\omega}{10}\right)\right] \end{aligned} \tag{16}$$

Combining conditions (14) and (16) yields

$$a_s^* = \sqrt{\frac{10(T-1)}{4T-3}}, \quad a_w^* = \sqrt{\frac{10T\omega}{4T-3}}.$$

To verify that these are in fact equilibrium ad levels, we must show w does not prefer some ad level satisfying $\delta_w(a_w) \geq \delta_s^* + \delta_w^*$ instead. Note that

$$\begin{aligned} &\frac{\mathrm{d}a_w \left[D(\delta_w(a_w)) + (T-1)D(\delta_s^* + \delta_w(a_w))\right]}{\mathrm{d}a_w} \\ =& a_w \left[1 - \frac{3a_w^2/\omega}{10} + (T-1)\left(1 - \frac{(a_s^*)^2 + 3a_w^2/\omega}{10}\right)\right] \\ =& a_w \frac{(3T-2)(a_w^*)^2/\omega + (a_s^*)^2 - 3Ta_w^2/\omega}{10}, \end{aligned}$$

which is negative for any a_w such that $a_w^2/\omega \geq (a_s^*)^2 + (a_w^*)^2/\omega$.

Therefore, under passive beliefs, the unique equilibrium features $a_s^* = \sqrt{\frac{10(T-1)}{4T-3}}$ and $a_w^* = \sqrt{\frac{10T\omega}{4T-3}}$, which, respectively, increases and decreases to their counterparts in the static case, $\sqrt{\frac{5}{2}}$ and $\sqrt{\frac{5\omega}{2}}$, as T goes from 2 to ∞.

$\square$

Second-Order Conditions (Section 2). We now state the second-order conditions associated with the total industry profit maximization problem considered in Subsection 2.2. The first-order derivative of industry profits, with respect to ad level a_j, is $\overline{\Pi}_j'(\boldsymbol{a}) = (\bar{a}-\bar{c})D'(\delta)\delta_j'(a_j) + D(\delta)$. Therefore, for any j and k, the own and cross second-order derivatives are given by

$$\overline{\Pi}_{jj}''(\boldsymbol{a}) = (\bar{a}-\bar{c})\left(D''(\delta(\boldsymbol{a}))\left(\delta_j'(a_j)\right)^2 + D'(\delta(\boldsymbol{a}))\delta_{jj}''(a_j)\right) + 2D'(\delta(\boldsymbol{a}))\delta_j'(a_j),$$
$$\overline{\Pi}_{jk}''(\boldsymbol{a}) = (\bar{a}-\bar{c})D''(\delta(\boldsymbol{a}))\delta_j'(a_j)\delta_k'(a_k) + D'(\delta(\boldsymbol{a}))\left(\delta_j'(a_j) + \delta_k'(a_k)\right).$$

In order to guarantee that the stated first-order conditions are sufficient for industry profit maximization, we assume that, for all $\boldsymbol{a}$ such that $\bar{a}-\bar{c} \geq 0$, the Hessian matrix with generic element $\overline{\Pi}_{jk}''(\boldsymbol{a})$ is negative definite.

Second-Order Conditions (Section 5). We now state the second-order conditions associated with the total industry profit maximization problem considered in Subsection 5.1. The first-order derivative of industry profits, with respect to ad level q_j, is $\overline{\Pi}_j'(\boldsymbol{q}) = (r_j(q_j) + r_k(q_k) - \bar{c})D'(\delta)\delta_j'(q_j) + r_j'(q_j)D(\delta)$. Therefore, for $j = s, w$ and $k \neq j$, the own and cross second-order derivatives are given by

$$\overline{\Pi}_{jj}''(\boldsymbol{q}) = (r_j(q_j) + r_k(q_k) - \bar{c})\left(D''(\delta(\boldsymbol{q}))\left(\delta_j'(q_j)\right)^2 + D'(\delta(\boldsymbol{q}))\delta_{jj}''(q_j)\right) + 2r_j'(q_j)D'(\delta(\boldsymbol{q}))\delta_j'(q_j) + r_j''(q_j)D(\delta(\boldsymbol{q})),$$
$$\overline{\Pi}_{jk}''(\boldsymbol{q}) = (r_j(q_j) + r_k(q_k) - \bar{c})D''(\delta(\boldsymbol{q}))\delta_j'(q_j)\delta_k'(q_k) + D'(\delta(\boldsymbol{q}))\left(r_k'(q_k)\delta_j'(q_j) + r_j'(q_j)\delta_k'(q_k)\right).$$

In order to guarantee that the stated first-order conditions are sufficient for industry profit maximization, we assume that, for all $\boldsymbol{q}$ such that $r_j(q_j) + r_k(q_k) - \bar{c} \geq 0$, the Hessian matrix with generic element $\overline{\Pi}_{jk}''(\boldsymbol{q})$ is negative definite.