The Impact of the Internet on Advertising Markets for News Media

by

Susan Athey, Emilio Calvano and Joshua S. Gans

First Draft: October, 2009
This Version: 8th October 2014

We provide a model of online advertising display markets where consumer attention may be divided among multiple publishers and, consequently, their advertising attention may be allocated to different platforms. We demonstrate that this gives rise to a mixture of single- and multi-homing advertisers and some consequent matching inefficiency between advertisers and consumers. Thus, as the number of switching consumers expands (associated with, say, the internet’s impact on news publishers), ad prices fall and a number of other competitive effects arise. We demonstrate that increased switching leads advertisers to favor reach over frequency and creates an incentive for contracting and technology improvements that can guarantee impressions to advertisers. Finally, we analyze the strategic choice of ad capacity, showing that, in general, increased switching leads to greater equilibrium ad capacity and lower prices. JEL Classification Numbers: L11, L82

Keywords: advertising, media, newspapers, matching, multi-homing, single-homing, tracking, two-sided markets, platforms

* Stanford (Athey), CSEF (Calvano), University of Toronto & NBER (Gans). This paper was previously circulated with the title “Will the Internet Destroy the News Media?” We thank participants at the ZEW Platforms Conference (Mannheim), Toulouse Network Conference (Seattle), 2010 Media Economics Workshop (CUNY), Annual IO Theory Workshop (Duke), Conference on Internet and Software (Toulouse), 2010 Economics of Information and Communication Technologies (Paris), 2011 International Industrial Organization Conference (Boston), 2011 CRES Conference (St Louis), 2011 NBER Summer Institute (Digitization), 2011 FTC Annual Microeconomics Conference and at seminars at Monash, Dartmouth, Harvard, Yale, Virginia, Toronto, Stanford, Chicago, Columbia, George Mason, London School of Economics, Boston College, Google, Microsoft Research, Harvard Business School, CSEF in Naples, Autonoma, Bologna, Ca Foscari, MIT, Berkeley, Northwestern, Queensland and Melbourne for helpful comments. Special thanks to Michael Riordan, Hanna Halaburda, Francisco Ruiz-Aliseda, Simon Anderson and Glen Weyl for helpful discussions. We thank Jing Xia for outstanding research assistance. Financial support from the Toulouse Network on Information Technology is gratefully acknowledged. Responsibility for all views expressed lies with the authors and not with any affiliated organization. The latest version of this paper is available at research.joshuagans.com.
1 Introduction

Traditional economic analyses of competition in media markets have tended to start from the perspective that consumers have scarce attention and allocate that attention to a single publisher (this is termed “single-homing” in the multi-sided platform literature). Given this, if there are advertisers that want to place ads in front of all consumers, those advertisers will be forced to advertise on all publishers (i.e., “multi-home”). This eliminates competition among publishers for those advertisers (Anderson and Coate, 2005; Armstrong and Wright, 2007). Subsequent empirical work has challenged this prediction with evidence that competition exists among advertisers, larger publishers command a premium in terms of ad revenue per consumer and lifting advertising restrictions on public broadcasters leads to reduced ad revenues for for-profit broadcasters. Indeed, the advertising industry has focused heavily on the concepts of reach (the number of consumers who can be impressed through a given publisher) and frequency (attempting to reach a given consumer a particular number of times), and these concepts, in turn, are central to how publishers position themselves to their advertising customers. For example, a key selling point of Facebook today is that it has an advantage in being able to offer advertisers the ability to reach a large fraction of the population a specified number of times, even as users access the internet through multiple devices.

The response to this empirical challenge has been to relax the assumption of single-homing consumers. Anderson, Foros and Kind (2013) and Ambrus, Calvano and Reisinger (2014) build from the original Anderson and Coate (2005) model to include a share of consumers who consume multiple-media publishers. They demonstrate that, as that share of “overlapping” consumers increases, advertisers’ willingness to pay for ads on individual publishers falls. This creates direct competition between publishers for

---

1 See Brown and Williams (2002); Brown and Alexander (2005).
2 Recently, this has been referred to as the “ITV Premium Puzzle” (Competition Commission, 2003). However, the relationship has been noted previously by Fisher, McGowan and Evans (1980) and Chwe (1988). See also, Crampes, Haritchabalet and Jullien (2009). See Goettler (2012) for recent empirical verification of such advantages.
3 For example, see Filistrucchi, Luini and Mangani (2011) for an empirical analysis of the French advertising ban on prime-time state television.
advertisers and resolves some of the empirical puzzles. While these identify potential competitive forces, here we contend that these models do not closely resemble the features of advertising markets as they pertain to the internet. In particular, the models assume that consumers who visit multiple publishers are able to do so without cost in terms of attention, and in particular, consumers who multi-home across two publishers double their consumption of media and, thus, see twice as many ads as single-homing ones.

Our goal in this paper is to build a model that more accurately captures the impact of the internet on consumer behavior, so that we are able to study how changes in consumer behavior affect advertising markets. A key feature of our model is that we assume that the internet allows consumers to easily spread their attention across multiple publishers, or switch between publishers. Increased consumer switching that many have observed is an essential distinguishing feature of online news consumption (Farhi, 2009; Gentzkow and Shapiro, 2011; Varian, 2010). Web browsers, search engines, aggregators and social network make it easy for consumers to move between publishers and increase consumer switching among publishers (Athey and Mobius, 2012), while free access removes other constraints. However, when attention is spread across publishers, switching consumers actually see fewer ads on a given publisher than their loyal counterparts.

In contrast to the previous literature, we also adopt a number of other features to bring the model closer to reality. First, we assume that publishers face issues in tracking consumer behavior across publishers; specifically, they can monitor consumers while on their publisher but do not know what those consumers saw when their attention was elsewhere (Edelman, 2010). Thus, publishers have imperfect tracking capabilities, which impacts on their value to advertisers. The motivation for this assumption is explored in depth in section 2. Second, and related, we assume that publishers cannot easily tell ex ante whether a given consumer is a switcher or a loyal consumer. Thus, they cannot set different prices to advertisers to access consumers of different types. Third, we assume that advertisers are heterogeneous in the value they receive from impressing consumers, although they value a single impression equally for all consumers.
As we will demonstrate, these assumptions allow a richer understanding of advertiser behavior in response to switching, but they also require a careful accounting for that behavior in equilibrium; specifically, in contrast to previous models, our theory predicts that publishers will serve a mixture of single- and multi-homing advertisers in equilibrium and that this will play an important role in the determination of equilibrium prices. As we show in section 4, advertiser multi-homing behavior is made meaningfully endogenous, so that we can make predictions about the strategy advertisers take towards advertising on multiple, potentially heterogeneous publishers.

The primary contribution of the model, however, is to analyze the consequences of a key source of inefficiency that is created by switching consumers in advertising markets with imperfect tracking technologies. (In other papers in this literature, advertising markets are always efficient in matching consumers to advertisers, outside of considerations of targeting.) With no consumer switching and a single market-clearing price for advertising, advertisers should place ads on all publishers. Consumer switching together with imperfect tracking of consumers across publishers creates inefficiencies in the form duplication; that is, consumers seeing an ad multiple times when an advertiser prefers them to see it just once. Switching by consumers is, thus, a source of diminishing returns to buying ad space on additional publishers (multi-homing).

Consequently, in equilibrium, higher-value advertisers choose to multi-home, as they have a higher opportunity cost of not informing readers, while lower-value advertisers single-home, avoiding wasted impressions. We show that the marginal return from an additional ad is a convex combination of marginal returns on switching consumers and loyal readers. Increased switching decreases marginal returns for multi-homing advertisers. As a consequence, in section 5, we are able to replicate, in a distinct way, the competitive effects that resolve some of the empirical puzzles with traditional

---

5 A countervailing effect outside our model is that with more data about consumers, publishers can sell more targeted advertising. See Iyer, Soberman and Vilas-Boas (2005) who demonstrate advertiser strategy when advertisements can be targeted. See Athey and Gans (2010) for an analysis of the impact of targeting technology on ad prices. See also Bergmann and Bonatti (2011) for an analysis of the interaction between online and offline media competition and targeted advertising.

6 Haigu and Lee’s (2011) analysis of two-sided platforms that connect together content (e.g. games, applications) and users is one of the few studies of multi-sided markets where the focus is on the endogenous choice of multi-homing on one side of the market (content providers). However, the driving force behind multi-homing decisions in their paper is exclusive contracts.
approaches. For example, our model predicts that high-value advertisers tend to buy the premium advertising inventory from the largest publishers and multi-home (e.g., the jeweler Tiffany’s advertises heavily in both the New York Times and the Wall Street Journal), while lower-value advertisers tend to focus their advertising on smaller publishers, where prices are lower, and avoid cross-publisher campaigns among publishers that have overlapping audiences (so they might choose to advertise on a set of locally focused news publishers).

The remainder of the paper is devoted to exploring the strategic implications of our model in a way that could assist empirical researchers looking at the impact of the internet on advertising markets for news and related media. First, in section 6, we examine content strategy. We find that having a larger set of loyal readers allows a publisher to have priority in advertising markets. More critically, however, our approach allows us to consider choices of the amount of content a publisher provides; that is, whether they try to capture a consumer’s full attention or not. We demonstrate that, in the presence of the inefficiencies described above, the relative returns to being a deep rather than a focused content provider fall as the share of switchers increases. This potentially explains the rise of dis-aggregated or unbundled content such as BuzzFeed or Facebook that has emerged as the internet has expanded.

Second, in section 7, we turn to consider how advertising contracts can be enriched to address some of the inefficiencies we identify. Our baseline model takes cues from reality to assume a simple contract space whereby advertisers pay for impressions but are not guaranteed a certain number of impressions. We consider a natural extension whereby publishers could offer a limited number of guaranteed impression contracts. Our model predicts that (a) advertisers will demand a mix of guaranteed and standard contracts; (b) that the prices of guaranteed contracts and of standard contracts will converge as more publishers offer contract menus; and (c) that the existence of those contracts may improve the efficiency of advertising markets and, hence, publisher revenue.

Third, in section 8, we explore publisher’s choice of advertising capacity. This has been a key focus in the literature on advertising as a nuisance (Anderson and Coate, 2005). Here, we consider it as a means of exercising market power in a media duopoly.
We show that consumer switching and inefficient duplication leads to non-monotone best-response functions by publishers in their choice of advertising capacity. This allows a rich set of predictions regarding the association between advertising level and the mix of consumers a publisher can attract.

A final section concludes and offers suggestions for future research and some empirical implications.

2 Display Advertising Markets for Online News Media

Here we describe the key features of online advertising markets for news media that distinguish it from its offline analog and give scope to our analysis. First and foremost, when consumers frequently browse multiple publishers in a given time period, an advertiser paying for multiple ads placed in front of consumers (or “impressions”) with multiple publishers will face the issue of duplication. The concern is that some consumers are impressed too many times and, therefore, some of the advertising budget, together with the consumer attention, is wasted. The flip side of duplication is that advertisers also face the real possibility that, without greater expenditures on impressions, advertisers might miss some consumers altogether.

To address this concern, technologies have been developed to keep track of which ads have been shown to whom; that is, the extent to which single ads are matched to individual consumers. In turn, these so-called tracking technologies and their capabilities determine which kind of products can be sold by media publishers to advertisers. To see this, consider two extreme and opposite examples. Suppose a record of previous ad/consumer matches were available. Then, in principle, it would be possible to match each successive ad from a given advertiser to a different consumer. No consumer would be matched to the same ad twice. Under this technology, which we refer to as perfect internal tracking, the publishers can sell directly to “consumers” (or in industry terms, “unique impressions”). Tracking is labeled “internal” as the publishers do not know to

---

7 Our focus is on display advertising which is distinguished from search or keyword advertising as it usually involves the display of an image-based ad on a website rather than text ads set aside on search results. Typically, display ads are paid on a price per impression basis whereas text ads are paid on a price per click basis. See McAfee and Vassilvitskii (2012) for an overview.
which ads this particular customer has been exposed on the rival’s website. So there is potential for duplicated impressions across publishers for those consumers who visit more than one publisher. At the other end of the spectrum, suppose there is no control whatsoever over the matching process. Then it would be as if each successive ad from a given advertiser were put in front of a consumer chosen at random, possibly one who was already impressed on the same publisher – a “no tracking” technology. Under this technology, the unit being sold is an “impression.” Today, when advertisers buy impressions on advertising exchanges without using cookies to track consumers, their experience is close to the no-tracking extreme.

Example: Suppose a coffee maker is seeking to place ads on the Google Display Network to boost its sales of prime quality coffee. Upon entering the target group of consumers, estimates on how many impressions can potentially be delivered are provided. Figure 1 shows the available inventory of impressions for a hypothetical campaign to individuals who visit websites related to “Gourmet Coffee” that belong to the Google Display Network on a weekly basis. For this case, the figure is between half a million and one billion. The second piece of information in the top-right column denotes the number of consumers who can be reached instead.\(^8\)

Second, unlike other decentralized markets, where supply is a choice at the producer’s discretion, here the inventory, in terms of (a) how many consumers can be reached and (b) how many times a given consumer can be impressed, is a random variable. The realization of this random variable depends on the amount of time consumers spend on a given website, as well as whether they visit the site at all within the relevant timeframe. So, at best, the publishers can sell a promise to impress as many consumers as possible at some price per consumer. This issue is best illustrated by current practice within online display advertising networks (Figure 2).

Example (continued): To perfect the campaign for “Gourmet Coffee” the campaign manager has to supply “a bid,” which is then used together with all other bids from other potential advertisers to allocate the inventory. In practice the advertiser is asked to reveal the maximum willingness to pay for an impression. By entering this bid, the advertiser effectively commits to purchase impressions so long as the price is less than 5 USD per thousand (Figure 2). When Google is also instructed to impress each consumer

\(^8\) The difference between “cookies” and “unique impressions” is due to the fact that, at best, Google can identify the device that displayed a particular ad on its screen. If some devices are shared among different individuals and/or some individuals use multiple devices, then this number differs from that of unique impressions. See Ghosh, Mahdian, McAfee and Vassilvitskii (2012) for an analysis.
at most once, then the actual product being sold is “impress each consumer that visits the Google Display Network once at a maximum price of 5 USD per thousand consumers.”

So, effectively, advertisers are asked to commit to buy as many impressions as possible at or below some given price.\(^9\) In line with this feature, this is precisely the kind of contractual arrangements we focus on in the baseline model. While issue (a) is well understood, common with the offline world and relatively easy to deal with, issue (b), that is, the dispersion of the maximum number of impressions that can be put in front of consumers, creates new managerial challenges as to the optimal way of allocating to advertisers the relatively scarcer attention of some consumers. Indeed, within the Google Display Network, while impressions are offered to advertisers and advertisers bid on consumers they would like to impress, Google offers no guarantee that the contract will be fulfilled. Specifically:

(excerpt from GDN help center) “Display Planner's estimates and statistics are based on the Google Display Network's available network inventory. With other advertisers bidding on this inventory, you won't be able to get the full estimated impressions – even with an unlimited budget, competitive bids, and ad creatives in every possible ad format.”

In a later section, we study how different contractual arrangements can potentially address this by improving the allocation of consumer attention and enhancing advertising revenues at the same time.

Finally, the online publishers (or the advertising networks to which they belong) need a mechanism for eliciting from advertisers their willingness to pay and a mechanism for allocating the impressions to the advertisers. In practice, automated procedures that allocate the impressions in real time via auctions are widely used.\(^10\) As modeling the actual mechanism is not within the scope of this paper, we rely on market clearing prices to allocate the inventory. However, using market clearing prices in this way can also be

---

\(^9\) The option of setting daily or weekly budgets allows limited exposure to the risk that comes with random inventory.

\(^10\) In some cases, these advertisements are procured and generated by news media publishers themselves (Arnosti, Beck and Milgrom, 2014). However, increasingly, such ads have been sold by advertising networks that manage the advertising inventory for a larger number of sites – both news and non-news. Our model applies equally to both structures. The key feature is that the suppliers of advertising inventory are numerous. According to eMarketer (http://www.emarketer.com/newsroom/index.php/google-display-ad-leader/), in 2012 the five largest U.S. display advertisers were Google (15.4%), Facebook (14.4%), Yahoo! (9.4%), Microsoft (4.5%) and AOL (3.6%). Increased mobile usage has fragmented the industry even further since that time.
seen as a way of understanding advertising when publishers and advertisers negotiate directly.11

3 Model Description

Our goal is to build a simple and tractable model that captures the crucial issues of duplication (due to advertisers endogenously purchasing multiple impressions on multiple publishers) and that allocates impressions to consumers in a way that spreads their limited attention across publishers in a heterogeneous manner. Here, we provide a basic model that captures these elements. In later sections, we show that crucial insights are robust to a variety of modeling extensions.

3.1 Consumer attention and advertising inventory

There is a continuum (unit mass) of consumers and two media publishers (indexed by $i$ and $j$). Consumers are endowed with a limited amount of attention – here assumed to be two units – which they allocate across publishers. Thus, given their choices, we can partition the set of consumers into three subsets: (i) those who devote all of their attention to publisher $i$ ($i$’s “loyals”); (ii) those who devote all of their attention to publisher $j$ ($j$’s “loyals”); and (iii) those who devote 1 unit to each (“switchers”). Let $D_i^l$, $D_j^l$ and $D^s$ with $D_i^l + D_j^l + D^s = 1$, denote the measures of these subsets.

The publishers use attention as an input to produce advertising inventory via a basic production function: $a_i$ is the maximum amount of ads (or impressions) that $i$ can put in front of a given consumer per unit of attention. We refer to $a_i$ as the capacity of $i$.12 So a switcher generates $a_i$ and $a_j$ inventory on publisher $i$ and $j$, respectively, while a loyal generates $2a_i$ impressions instead. The total supply of impressions by publisher $i$.

---

11 Our market clearing assumption aggregates outcomes across what is termed in the industry as premium ads (sold primarily through direct negotiation with publishers) and remnant (which uses more arm’s-length auction mechanisms); see Fisher (2012). In section 7, we explore some enrichment of the contract space that takes into account increasingly diverse forms of advertising placement arrangements.

12 As the mass of consumers is normalized to one, throughout the analysis we need also maintain that $a_i \leq \frac{1}{2}$. As will be clear later on, the assumption allows for positive excess demand of $i$’s impressions when their price is close to zero.
is equal to \((2D_i^I + D^r)a_i\). Note that this production function captures the dispersion discussed above in a simple way, as loyal consumers can be matched to twice as many advertisers than switchers on a given publisher.

An important feature of this model is that it allows for different allocations of consumer attention while keeping the total attention per consumer fixed. This allows our results to be grounded in the market fundamentals that do not change as we analyze the impact of the internet.

### 3.2 Advertisers’ preferences

There is a unit mass of advertisers. Advertisers want each consumer to see their ads. They differ as to the value of putting an ad in front of a consumer. This value is denoted \(v\) and is distributed on \([0,1]\) according to a cumulative distribution \(F\).\(^{13}\) For simplicity, this value does not increase if the same consumer sees more than one ad from a given advertiser. Thus, from the perspective of a given advertiser, an “impressed” consumer is a consumer who has been exposed to at least one ad.

### 3.3 Tracking technology

As attention is a primitive of the model, we need to specify how attention is matched to ads – something that is usually not explicitly considered in the economics of advertising literature. Specifically, suppose that an advertiser were allocated an arbitrary number of impressions on each publisher; how many different consumers would be impressed? As discussed, the extent to which ads from the same advertiser can be matched to different consumers so as to avoid duplication relates to a publisher’s “tracking capabilities.” In reality, publishers have some, but not full, capabilities to keep track of which ads have already been served to consumers while browsing their own pages but little or no ability to track which ads their consumers see on other publishers. To capture this in the simplest way, we posit that publishers can perfectly track consumers internally. That is, we assume that if an advertiser is allocated \(n_i\) impressions

\(^{13}\) For the given advertiser, \(v\) is the same for all consumers and independent of the number of distinct consumers receiving an impression. An alternative specification might have advertisers aiming to reach a specific number of consumers (Athey and Gans, 2010) or a specific consumer type (Athey and Gans, 2010; Bergemann and Bonatti, 2010).
of $i$ then the advertiser effectively reaches $n_i$ different consumers of $i$. However, perfect internal tracking does not prevent an advertiser that has purchased impressions on both publishers from reaching a switching consumer twice (and, therefore, wasting an impression). Formally, if an advertiser purchases $n_1$ impressions on publisher 1 and $n_2$ on publisher 2 (where, without loss in generality, each $n_i \leq D_i' + D^*$), then the associated “reach” of the advertiser’s overall campaign is described by a function, $\Phi$, that maps $(n_1, n_2)$ into $[0, 1]$. $\Phi$ is the expected number of unique consumers impressed by at least one ad from the advertiser. So the expected value of an advertising campaign $(n_1, n_2)$ is $v\Phi(n_1, n_2)$ for an advertiser with value $v$.

Let

$$\Phi(n_1, n_2) := n_1 + n_2 - \frac{D^*}{(D_i' + D^*)(D_j' + D^*)} n_1 n_2.$$  

To understand this formula, note that (1) if all consumers were exclusive ($D^* = 0$) then the total reach would be $n_1 + n_2$ and there would be no duplication; (2) if $n_i = 0$ then the total reach would be equal to $n_j$ and there would be no duplication; (3) if all consumers were in common ($D^* = 1$ and $D_i' = D_j' = 0$) then the total reach is equal to 1 minus the probability that a given consumer is not informed, that is, $1 - (1 - n_i)(1 - n_2)$. The formula above, derived in appendix from basic and natural assumptions on the stochastic processes that govern the matching of unique consumers to advertisers, generalizes reach for the case with asymmetric publisher popularity ($D_i' \neq D_j'$) and partial overlap ($0 < D^* < 1$).  

3.4 Advertising products

As discussed in section 2, in reality, due to the random nature of the inventory, advertisers are asked to commit to buy as many impressions as possible in a given time

---

14 The familiar “Butters’ technology,” first employed in Butters (1977) and then extensively in the literature, approximates instead the outcome under no tracking with $\Phi$ taking an exponential form: $1 - e^{-n_1-n_2}$.
period at or below a given price. In our basic framework, with publishers facing two consumer types (loyal and switchers) but no ability to differentiate the two at a given moment in time, we capture this by assuming that publisher $i$ offers to impress all consumers in a single-attention period. Over two periods, this formally corresponds to the following contract: "Impress all loyal consumers and a random half of the switchers at price $p_i$ per impression," being the expected shares of each consumer type an advertiser will impress. All advertisers can, in principle, accept this contract at neither, one, or both outlets. In line with the literature, in what follows, we refer to these actions as “not purchasing,” “single-homing on $i$,” and “multi-homing,” respectively. The convention adopted means that each publisher can fulfill exactly $2a_i$ distinct contracts with advertisers. Finally, note that accepting contract $i$ entitles a quantity of $D'_i + \frac{1}{2}D^r$ impressions. For future reference, let $\Pi_i(v, p_1, p_2)$ be an indicator equal to one if type $v$ accepts $i$’s contract and 0 otherwise, and let $\tilde{n}_i(\Pi_i) := \left(D'_i + \frac{1}{2}D^r\right)\Pi_i$ denote the number of impressions consumed on publisher $i$ given $D'_i$, $D^r$, and choice $\Pi_i$.

In the baseline case, all publishers are price takers with $p_1$, $p_2$ determined in equilibrium via a market clearing condition that equates the number of contracts demanded to $2a_i$. In section 7, we discuss alternative contractual terms (i.e., advertising products). In section 8, we allow the publishers to affect market-clearing prices by endogenizing the supply of impressions.

### 3.5 Equilibrium definition

To close the model, we assume a basic-price mechanism that equates the individual publishers’ demand and supply of advertising contracts and therefore of impressions. A market equilibrium is a tuple $\left\{ (\Pi'_1(v, \hat{p}_1, \hat{p}_2), \Pi'_2(\hat{v}, \hat{p}_1, \hat{p}_2)), \hat{p}_1, \hat{p}_2 \right\}$ where:

1. For each advertiser $v$, $(\Pi'_1, \Pi'_2)$ maximizes $v \Phi(\tilde{n}_1(\Pi'_1), \tilde{n}_2(\Pi'_2)) - \hat{n}_1(\Pi'_1)\hat{p}_1 - \hat{n}_2(\Pi'_2)\hat{p}_2$,

---

15 As will be explored in Section 7, increasingly advertisers are able to negotiate for a certain number of guaranteed impressions (see Fisher, 2014). However, because of imperfect tracking, there is no guarantee of a certain number of impressions on unique visitors to a publisher and, thus, our approach here initially has some broad coverage of practices to date.
and (ii) For each publisher $i$, $(\hat{p}_1, \hat{p}_2)$ is such that $\int_0^{I^*_i} (v, \hat{p}_1, \hat{p}_2) dF(v) = 2a_i$.

The first condition formalizes the idea that individual demand is determined by advertisers maximizing profits taking prices as given, while the second condition says that the market for impressions (or equivalently the market for advertising contracts) of each publisher clears. To understand (ii), note that every advertiser that accepts the contract of publisher $i$ is effectively allocated $D^i_l$ impressions on loyalys (one for each loyal customer) and, $\frac{1}{2}D^s$ impressions on switchers whose aggregate supplies are respectively $D^i_l 2a_i$ and $D^s 2a_i$. So (ii) is (trivially) equivalent to clearing, simultaneously, the market for impressions on loyalys and switchers:

(ii’) For each publisher $i$, $(\hat{p}_1, \hat{p}_2)$ is such that $D^i_l \int_0^{I^*_i} (v, \hat{p}_1, \hat{p}_2) dF(v) = D^i_l 2a_i$, and

(ii’’) For each publisher $i$, $(\hat{p}_1, \hat{p}_2)$ is such that $D^s \int_0^{I^*_i} (v, \hat{p}_1, \hat{p}_2) dF(v) = D^s 2a_i$.

Accordingly, in what follows, we refer to $(\hat{p}_1, \hat{p}_2)$ as the prices that clear the market for impressions.

3.6 Benchmark: Perfect tracking

As a final step in the model set-up it is useful to describe a benchmark in this environment. The first-best allocation of consumers’ attention to advertisers is such that the highest value advertisers are allocated with priority to scarce advertising inventory and there is no duplication. Let $v_i$ denote the marginal advertiser allocated to consumers loyal to publisher $i$, and let $v_s$ denote the marginal advertiser allocated to consumers who switch between publishers $j$ and $i$. An efficient allocation of advertisers to consumers involves allocating all advertisers with $v \geq v_i$ to publisher $i$’s loyal consumers and those with $v \geq v_s$ to those who switch between $i$ and $j$. Thus, the marginal advertisers are defined as the unique solution to $2a_i = 1 - F(v_i)$ and $a_i + a_j = 1 - F(v_s)$.

To see how this first best might be implemented in practice, consider a scenario where there exists a public record that keeps track of all consumer/ad matches. More realistically, suppose that both publishers outsource their advertising to a third party,
labeled “ad-platform.” The platform acquires the publishers’ entire advertising inventory and can keep track, say by planting “cookies” on the consumers’ web browsers, of all previous consumer/ad matches.\(^\text{16}\) Moreover suppose that the platform can price discriminate based on consumer type (loyal or switcher).

In this scenario there would still be two markets: one market for impressions on loyals and another one for impressions on switchers. The prices that equate demand and supply are equal to \(\hat{p}_i = v_i\) and \(\hat{p}_s = v_s\), respectively. Thus, advertisers will choose to advertise so long as their value exceeds the impression price. Prices for loyals and switchers will be the same if publishers are symmetric (if \(a_i = a_j\), then \(\hat{p}_i = \hat{p}_j = \hat{p}_s\)), but if publishers are asymmetric (\(a_i > a_j\)), then \(\hat{p}_i < \hat{p}_s < \hat{p}_j\), since outlet \(j\) as a lower supply of ad space and thus market-clearing prices for consumers who must be reached there are higher. In equilibrium, the publishers’ inventory is worth \(\pi_i = \hat{p}_i 2a_i D_i^l + \hat{p}_s a_s D_s^s\).

### 4 Equilibrium Analysis: Existence and Uniqueness

Before assessing the role of switching in shaping market outcomes, we show that the equilibrium model presented is well behaved, verifying the existence and uniqueness of a market equilibrium. The additional challenge with respect to previous works is that advertisers are heterogeneous. So it is necessary to identify the different advertising strategies associated with different valuations in order to construct the aggregate demand for impressions for a publisher.

Given consumers’ allocation of attention and prices, we proceed to characterize the advertisers’ demand as a function of their idiosyncratic valuation. By definition, in equilibrium, advertisers maximize the expected value of their reach less the expenditures associated with achieving that reach. Specifically, the payoffs associated to the different choices are:

\(^{16}\) An alternative (but probably less realistic) assumption would be that the ad platform shares information with the publisher about the consumer type, so that the publisher can set different capacities for different types. This additional flexibility would lead to a scenario with essentially distinct markets, so that firms compete for switchers but have a monopoly over access to loyal users. It is a bit more complicated to think how this would work in practice, since consumer types would only be fully determined in the second period, after the consumer had already experienced a first-period ad capacity. We omit the formal analysis of this case.
i. Single-home on 1 \[(v - p_1)(D_1^l + \frac{1}{2} D^s)\]

ii. Single-home on 2 \[(v - p_2)(D_2^l + \frac{1}{2} D^s)\]

iii. Multi-home \[v(D_1^l + D_2^l + \frac{3}{4} D^s) - p_1 D_1^l - p_2 D_2^l - (p_1 + p_2)\frac{1}{2} D^s\]

Note that, conditional on patronizing one publisher, the restriction to buy as many impressions as possible on that publisher is innocuous. The reason is that the marginal value of an impression does not diminish with the number of impressions already made, as there is no duplication within publishers by assumption.

A key observation is that the combination of switching and no tracking across publishers is a source of diminishing returns from multi-homing; that is, from purchasing additional impressions on a different publisher. To build intuition, suppose that publishers are symmetric \(D' \equiv D_1^l = D_2^l\) and \(a \equiv a_1 = a_2\), and assume equal prices \(p \equiv p_1 = p_2\) (indeed, this will be the case in equilibrium). Recall that, by definition, \(2D' + D^s = 1\). By single-homing on either publisher, the expected return is equal to the value of informing half the population \(\frac{1}{2}v\) while the expected expenditure is \(\frac{1}{2}p\). If the advertiser were to multi-home instead, it would purchase an additional \(\frac{1}{2}\) impressions at a cost of \(\frac{1}{2}p\). However, the return to that would be less than the value of informing the other half of the population \((D' + \frac{1}{4} D^s)v \leq \frac{1}{2}v\) due to duplicated impressions. Returns are constant only if \(D^s = 0\).

The property of diminishing returns immediately implies that advertisers sort, in equilibrium, with relatively higher types patronizing more publishers overall. Figure 3 illustrates the sorting that arises in the symmetric case. Type \(v_m\) is indifferent between single-homing on either publisher or multi-homing; that is, \(v_m = p\frac{D' + D^s}{D + \frac{1}{4} D^s}\). Type \(v_s\) is, instead, indifferent between single-homing and nothing; that is, \(v_s = p\). In what follows, we refer to advertisers in \(v_m \leq v\) and in \(v_s \leq v \leq v_m\) as “high” and “low” types, respectively.

To solve for the market equilibrium, the publishers’ respective aggregate demands have to equal their supply. Aggregate demand is the sum of individual demands across
advertisers’ types. Note that with equal prices, all low types are simultaneously indifferent between single-homing on either publisher. In the symmetric publisher case, per impression prices must equalize across publishers. If that were not the case, then the aggregate demand for the cheaper publisher would be relatively higher and the markets for impressions would not be clear. For the same reason, equal prices require single-homing advertisers to split equally across publishers. Note that this is a trivial source of multiplicity, as, due to indifference, advertisers can split in multiple ways keeping the aggregate demands fixed. As both thresholds $v_m$ and $v_s$ are monotone decreasing in $p$, a market equilibrium necessarily exists. The following proposition generalizes this argument to allow for asymmetric publishers. All proofs are in the appendix.

**Proposition 1.** For all $a_1,a_2 > 0$ and $D_1', D_2', D' \geq 0$, a market equilibrium exists and is unique up to indifferences.

In the rest of the paper, we first assess the impact of switching for the simplest case of symmetric publishers. We then proceed to the description of the equilibrium allocation and prices with asymmetric publishers. Section 6 tackles the case with asymmetric readership shares ($D_1' \neq D_2'$), section 7 considers alternative advertiser contracts and section 8 tackles the case with asymmetric capacities ($a_1 \neq a_2$).

## 5 Switching and Market Prices

How does consumer switching shape equilibrium advertiser behavior and publisher pricing? With fixed supply, the market equilibrium properties are basically inherited from the properties of the aggregate demand for ads. In particular, a reduction in the aggregate demand of one publisher implies lower prices and profits for that publisher.

In equilibrium, the aggregate demand for impressions of publisher $i$ equals:

$$(D'_i + \frac{1}{2} D') \left(1 - F(v_m) + \frac{1}{2} \left( F(v_m) - F(v_s) \right) \right) = \frac{1}{2} \left( 1 - \frac{1}{2} \left( F(p \frac{1}{1+p}) + F(p) \right) \right)$$

where $D'_i + \frac{1}{2} D' = \frac{1}{2}$ follows by symmetry. Thus, aggregate demand is falling in $D'$. The arrow in Figure 3 depicts the effect of a marginal increase in $D'$. As switching increases, the high types scale back on advertising and become low types due to the increased
duplication on switchers. The following result directly follows from the observation that demand is monotone decreasing in $D^*$. 

**Proposition 2.** Under symmetry, equilibrium prices and profits are decreasing in $D^*$. 

What can be said about the equilibrium allocation? Duplication causes demand and, therefore, prices to drop below the level that would arise without switching. This means that some low-value advertisers that would not have had access to consumer attention in a perfect tracking benchmark now would. So two sorts of “mismatches” occur in equilibrium. The combination of inefficient depletion (due to duplication) and inefficient use of attention (due to ad/consumer mismatches) is necessarily suboptimal from the perspective of the total advertiser and publisher welfare (and would also be suboptimal from a consumer perspective, had we spelled out a richer model where consumers receive surplus from matching with advertisers, or are annoyed by repetitious ads).

In this model, when there is no switching (i.e., $D^* = 0$), equilibrium publisher profits become $D_i^*F^{-1}(1 - 2a_i)2a_i$; the same profits that would arise under perfect tracking.

The mechanism that leads to lower equilibrium prices in our model here is quite distinct from that identified in alternative models (e.g., Anderson et al., 2013; Ambrus et al., 2014). In this context, switching has real effects in that it degrades the value of the inventory. One way to understand this property of our model is to consider the advertisers’ dual problem of minimizing the expenditure needed to guarantee a certain reach

$$\min_{n_1, n_2} (n_1 + n_2)p \quad s.t. \quad \Phi(n_1, n_2) \geq \phi$$

where $\phi > \frac{1}{2}$. Switching degrades the inventory in the sense that a higher $D^*$ raises the minimum number of impressions $n_1 + n_2$ required to get at least a reach of $\phi$. This reduces the market price of ad impressions. By contrast, in other models (Anderson et al., 2013; Ambrus et al., 2014), prices decrease due to publishers competing down the price charged to advertisers for access to multi-homers. In those models, switching causes a reallocation of the advertising surplus while leaving the total surplus unchanged.
Moreover, our model delivers at least two distinct predictions. First, *as the share of consumers who are switchers increases, the share of advertisers that single-home in equilibrium increases*. Second, it predicts that those advertisers that place the lowest value on reaching consumers will be the single-homers. Advertisers for which the value of reaching the marginal consumer is high will spread their advertising across many publishers.\(^{17}\)

What does this result imply for the impact of mergers and public (i.e., non-advertising) publishers? With regard to media publisher mergers, there will only be gains from trade in our baseline model if (1) the merger permits perfect tracking (internal to the merged firm) and, hence, leads to a more efficient allocation in the advertising markets, or (2) if perfect tracking is not possible, the publishers can offer a distinct advertising product for multi-homing advertising as opposed to single-homing advertising and use this as a basis for price discrimination among different advertiser types.\(^{18}\) In either case, we would expect to see media mergers associated with higher prices for advertising; however, these may be associated with better matching and so may lead to a more efficient outcome.\(^{19}\)

With regard to competition from non-advertising (e.g., public) media, studies have documented that for-profit media publishers object to public publishers being able to sell advertising.\(^{20}\) This stance is seen as a puzzle in traditional media economics as advertising usually annoys consumers and, hence, public advertising would allow competing publishers to attract users away from public publishers. Suppose that, in our model here, one publisher were to switch from advertising to non-advertising but otherwise would remain the same. This has two effects. First, while that publisher still captures (scarce) consumer attention, it decreases the effective supply of advertising capacity in the market. Second, unlike switchers between mainstream publishers,

\(^{17}\) The possibility that advertisers will purchase multiple impressions at a rate that likely leads to waste is borne out by the ComScore data. For instance, they estimate that in the first quarter of 2011, almost 1.1 trillion display ads were delivered in the U.S. Of these, 19.5 billion were purchased by AT&T, 16.6 billion by Experian Interactive and 11.2 billion by Scottrade. If the entire American population surfed the net daily during that time, they would see one AT&T ad per day.

\(^{18}\) Our earlier working paper (Athey, Calvano and Gans, 2013a) demonstrates this.

\(^{19}\) In section 8 below, we explore the impact of switching on the total available advertising space that may be a better reflection of efficiency considerations.

\(^{20}\) For example, see Filistrucchi et al. (2011) for an empirical analysis of the French advertising ban on prime-time state television.
switchers between non-advertising and advertising publishers do not contribute to the wasted impressions problem. Consequently, the switch to being non-advertising reduces duplication. This increases the demand for advertisements. These two effects – a decrease in supply and an increase in demand – combine to raise equilibrium impression prices and profits for the advertising publisher.

6 Switching and Content Strategy

We now turn to consider the impact of consumer switching on other aspects of the news media. In this section, we examine the impact on content, looking first at “vertical” differentiation between publishers in terms of quality and then turning to the amount of content a publisher might choose to provide.

6.1 Vertical content differentiation: “Bigger is better”

To shed light on how switching affects the returns to acquiring consumers and, therefore, the incentive to invest in “quality,” suppose that one publisher is able to generate a higher readership share than the other; that is, it has higher quality. Formally, assume that \( D_1 < D_2 \) but the publishers are otherwise symmetric.

To build intuition, we work here through an indirect argument. Specifically, we suppose that the market clearing unit prices are equal across publishers and conclude that these prices cannot be part of a market equilibrium. To this end, consider the advertisers’ demand. Recall that, in the baseline case, relatively low-valued advertisers were indifferent between single-homing on either publisher. Figure 4 illustrates the sorting that would arise instead with equal prices, assuming that one publisher is marginally more attractive than the other. When publisher 2 is larger, advertisers will sort on to that publisher first. The reason is that revenues from single-homing on publisher 2 exceed that from single-homing on publisher 1 for all active advertisers:

---

21 A recent paper by de Corniere and Taylor (2013) develops this point. They show that if Google were to “divert” attention to websites with little or no ads it could achieve higher ad-prices (due to scarcer supply) trading off search traffic (due to lower attractiveness). So they show that the supply-side effect described here can be a basis for “search engine bias.”

22 We use “quality” in the economic sense that it drives more consumer demand as opposed to some other criteria that might be applied to news content. It is well-established that changes in content strategy do impact on the shares of newspaper readers (George and Waldfogel, 2006).
\((D'_2 + \frac{1}{2}D')(v - p) > (D'_1 + \frac{1}{2}D')(v - p)\). Everything else held constant, this creates upward pressure on the relative price of publisher 2’s ads to rebalance supply and demand on both markets. So an (arbitrarily small) competitive edge in quality gives a positional advantage to publisher 2 in the sense that its impressions command a price premium. Note that this argument relies on there being single-homing advertising and, hence, as discussed, on \(D^s > 0\). The following proposition takes stock (proof omitted).

**Proposition 3.** Let \(D'_1 < D'_2\) and \(a_1 = a_2\). Then \(\hat{p}_1 < \hat{p}_2\) if and only if \(D^s > 0\).

Figure 5 shows the equilibrium allocation of advertisers.

The fact that “larger” publishers, in terms of readership share, command a premium for their ad space is a known puzzle in traditional media economics (Goettler, 2012). The reason being that, in a canonical model, consumers are equally valuable regardless of the publisher they are on, yet in practice advertising rates are typically higher on larger publishers. Here, because ads are tracked more effectively internally, placing ads on the larger publisher only involves less expected waste than when you place ads on the smaller publisher or spread them across publishers. So, the larger publisher can command a (tracking-related) premium and this fact can, arguably, contribute to account for the observed wedge in impression prices. It is useful to note that the price premium here is commanded over all advertising impressions. Again, the mechanism here is quite distinct than that in Anderson et al. (2013) and Ambrus et al. (2014). In this models, as platforms cannot appropriate the rent associated to switching consumers, larger publishers command a per viewer premium as they have more exclusive viewers. Here, on the other hand, there are single-homing advertisers that are choosing between both publishers. To them, a bigger publisher (i.e., with more consumers) is valuable “per se,” regardless of the relative composition; that is, regardless of the ratio of switchers to exclusive viewers.\(^{23}\)

While we are agnostic here about the sources of differences in readership shares, the notion of positional advantage allows us to shed light on some of the incentives as well as the unintended consequences of various strategies that may affect publisher size.

\(^{23}\) Furthermore, with three publishers, the premium can potentially disappear in these other models, as “bigger” does not mechanically imply to have the most exclusive. In contrast, our model is still there.
For example, when publishers compete for readers, they will be competing on dimensions that raise the likelihood that consumers visit them in any given period. Not only does that increase their total scale (and hence, the base upon which they can earn advertising revenue), it also increases their priority in the advertising market. Thus, imperfect tracking makes competitive pressures for consumers somewhat more intense than traditional media economics would imply.\textsuperscript{24}

For instance, there has been much discussion regarding paywalls and how these interact with advertising businesses. Common wisdom is that paywalls are a substitute for media publishers than advertising that might otherwise cause consumer annoyance (see Anderson and Coate, 2005; Casadesus-Masanell and Zhu, 2010). Here, however, a paywall would have an indirect effect through the advertising market. Regardless of how a paywall is actually structured,\textsuperscript{25} we speculate that the effect would be to reduce the relative size of the publisher establishing it. Thus, in addition to losing consumers (and their associated advertising revenue), the paywall also reduces or harms the publisher’s priority in the advertising market. This suggests yet another competitive constraint on establishing paywalls specific to our theory.

\subsection*{6.2 Investing for reach: “Focused is better”}

The previous subsection suggests that investing for readership might be a desirable strategy for publishers seeking a positional advantage in online advertising markets. But it is also the case that a publisher’s content strategy may be to provide less but higher quality content overall. While the analysis thus far has followed the existing literature and assumed that publishers have sufficient content to potentially absorb both periods of attention, here we endogenize the quantity of content offered and investigate the circumstances under which a publisher prefers to supply less content. Specifically,

\textsuperscript{24}While Anderson et.al. (2013) and Ambrus et.al. (2014) do not consider investments in quality, it is straightforward to see that, in their model where switching consumers are both marginal to publishers and also do not attract advertising revenue, that switching mutes the incentives to invest in quality and a larger readership compared with the model presented here. Thus, the impact of switching on investments in attracting readers is another point of distinction between our model and theirs.

\textsuperscript{25}In the working paper (Athey, Calvano and Gans, 2013a), we demonstrated this for micropayments (which reduced the overall number of consumers visiting a publisher), subscriptions (which reduced the incentive for would-be switchers to allocate partial attention to a publisher and so become exclusive to the other publisher) and a limited paywall (that only charged if a consumer wanted to allocate more than one unit of attention to a publisher and, therefore, induced them to become a switcher).
when is “reach” a better investment than “depth of engagement”? Or, to put it differently, how do websites that capture the attention of a vast, albeit a shared, audience such as Facebook compare to websites that engage deeply loyal readers?

We tackle this question by studying the strategic game in which the publishers choose simultaneously whether to provide “deep” or “focused” content. The baseline model is extended as follows. Suppose that, prior to consumers and advertisers making any choices, the publishers simultaneously choose whether to have focused \((f)\) or deep \((d)\) content. Producing one unit of content costs \(c \geq 0\). So choosing \(f\) over \(d\) entails an additional payoff of \(c\) interpreted as cost savings. Our goal is to characterize the Pure Strategy Nash Equilibria of the game and study how these are affected by the cost \(c\). Let \(D_i^f(d,d)\) and \(D_i^d(d,d)\) denote the amount of loyal consumers of \(i\) when both publishers choose deep content, the first argument denoting \(1\)’s choice. For simplicity, assume symmetry \(D_1^f(d,d) = D_2^f(d,d)\) and \(a_1 = a_2\). A “focused content publisher” is defined as a publisher that can be allocated at most one unit of attention and therefore attracts only switchers. Formally, we assume throughout this section that if one publisher, for example publisher 1, chooses \((f)\), then \(D_1^f(f,d) = 0\) regardless of publisher 2’s choice. Note that this assumption implies that if publisher 2 mimics 1 then \(D_2^f(f,f) = 1\). If it chooses \(d\) instead, then \(D_2^f(f,d) = D_2^f(d,d)\). In addition, as we keep the aggregate supply of attention fixed, if 1 chooses \(f\) then 2 acquires all of one’s potentially loyal customers: \(D^d(f,f) = D_1^f(d,d) + D^d(d,d)\). Note that the two extra cases we need to consider here, corresponding to the subgames following choices \((f,d)\), \((d,f)\) and \((f,f)\), are all special instances of the model with asymmetric readerships considered in the previous subsection. Therefore, the associated payoffs can be easily recovered and this completes the description of the game. Lastly, to simplify the analysis, assume \(F(v) = v\).

Note that, excluding the degenerate case in which \(D^d(d,d) = 1\), choosing \(d\) over \(f\) always increases revenues (i.e., profits net of costs) as it allows them to reduce the extent of switching. However, to compute the return to provide a “second unit” of content, we need to compare these additional revenues to the cost saving \(c\). Most
importantly, we need to establish how these returns change, depending on whether one’s rival adopts a deep versus a focused strategy.

As we shall see a key driver of the equilibrium outcome is a fundamental property of the game, which is (strong) strategic substitutability. That is:

\[ \pi_1(d, f) - \pi_1(f, f) > \pi_1(d, d) - \pi_1(f, d). \]

Intuitively, the first publisher to supply deep content taps into the rival customer base persuading all “would-be loyal” consumers. The second (or follower) publisher investing merely retakes some of those consumers back. This observation is at the heart of the following proposition. The proposition establishes the conditions on costs such that an asymmetric equilibrium emerges from otherwise symmetric publishers. It also establishes uniqueness up to permutation of the indexes.

**Proposition 4.** Suppose that \( F(v) = v \). For all \( D' < 1 \) there exists \( c < \bar{c} \) such that in equilibrium (i) for \( c < \xi \) both publishers choose to provide deep content; for \( c \in (\xi, \bar{c}) \) one publisher chooses deep content while the other chooses focused content and for \( c > \bar{c} \) neither publisher chooses deep content. Finally \( \xi \rightarrow \bar{c} \rightarrow 0 \) as \( D' \rightarrow 1 \).

The second and, in light of our motivating question, most important result is that increasing the share of switchers decreases the returns for one and eventually both publishers to provide deep content. Intuitively, deep content is most valuable when providing it builds up a loyal customer base. So with switching, the deep content publisher’s premium evaporates. Proposition 4 provides an interesting perspective on the way news and other content might be organized on the internet. Because of the duplication issues associated with imperfect tracking, the relative returns to deep versus focused content are skewed towards focused content.\(^{26}\)

\(^{26}\) Indeed, while we do not demonstrate it here, given our assumption that simple contracts are offered to advertisers, as we address it in the next section, it is possible to consider more elaborate contracts that may allow advertisers to single-home on one publisher, when all consumers are switching, and still impress all consumers. In this situation, there will be pressures towards unbundling of news content across publishers with consumers choosing their own aggregate bundles and seeing distinct ads on each one. Thus one view is that the pressures on publishers are not to reduce switching, but rather to ensure that consumers switch more and include their publisher as part of the bundle: precisely the notion of investing for reach.
7 Switching and Advertiser Contracts

We now focus on the benefits to publishers of offering a “guarantee” to impress all customers on a given publisher. Such guarantees are sometimes offered directly, but a similar outcome arises when inventory is sold through second price auctions (and analogous mechanisms) that allocate consumer attention sequentially on a highest bid / first served basis. As we shall see, our framework allows us to consider these in a simple way. Specifically, we show that the duplication concerns that drive the previous findings, together with the dispersion of consumer attention described in section 2, create demand for richer advertising products. Surprisingly, we show that, in the context of our specific model, such guarantees generate a first best allocation of attention. So these contractual arrangements are at least, in theory, capable of making up for the lack of a perfect-tracking technology by allowing advertisers to construct a portfolio of impressions across publishers that, in equilibrium, avoids duplication. Finally, we characterize the equilibrium that arises when both publishers are allowed to sell guarantees.

We modify the baseline model as follows. Each publisher sells two contracts:

(G)uaranteed: impress all consumers who visit across both time periods at price \( P_i \) per impression.

(S)tandard: impress as many consumers as possible at price \( p_i \) per impression.

---

27 McAfee, Papineni and Vassilvitskii (2013) examine auctions for guaranteed delivery in advertising exchanges. They examine a situation where there are many advertising options and products available and run into the same management issues highlighted here. They analyze optimal auction design to manage this issue from the perspective of a single advertising exchange. By contrast, we look at how this impacts on the allocation of advertisers across exchanges.

28 In Athey, Calvano and Gans (2013a), we explored another expansion of the product space whereby advertisers could purchase multiple impressions on each publisher but internal tracking was also imperfect to some extent. In this situation, we demonstrated that all of the above results continued to hold, however, when the number of switchers became high, some high-value advertisers would purchase more impressions on one publisher and “crowd out” low-value advertisers; potentially, raising publisher impression price and profits. We found, however, that this outcome was not robust to changes in capacity; especially when that capacity was endogenous.

29 While at present there have been some moves towards contracts that guarantee a certain number of impressions for advertisers (see Fisher, 2014), here our notion of guarantee is far stronger as it is a guarantee to impress every unique visitor to a publisher. While our specific model assumes that the tracking capabilities exist internally to fulfill such a guarantee, it is far from clear that those exist in reality, as many readers cannot even be tracked internally.
Let \( I_i^G(v, p_1, p_2) \) be an indicator equal to one if type \( v \) accepts \( i \)'s guaranteed and standard contract, respectively, and 0 otherwise. Note that since we have assumed that there are only two consumer types, \( n_i^G(I_i^G) := (D_i^I + D_i^S)^G \) and \( n_i^S(I_i^S) := D_i^I D_i^S \).

\( P_i \) and \( p_i \) are determined in equilibrium via the following market clearing conditions

\[(ii'') \text{ For each publisher } i, \, (\hat{P}_i, \hat{P}_2) \text{ is such that } \int_0^1 I_i^G(v, p) dF(v) = a_i \]

\[(iii'') \text{ For each publisher } i, \, (\hat{p}_i, \hat{p}_j) \text{ is such that } \int_0^1 I_i^S(v, p) dF(v) = a_i \]

where \( p \) denotes the entire vector of prices. To understand (ii''), note that every advertiser that accepts the contract (G) of publisher \( i \) is effectively allocated \( D_i^I + D_i^S \) impressions. So (G) can only be sold to at most \( a_i \) advertisers.

Consider, first, the symmetric case and so assume \( P_1 = P_2 = P \) and \( p_1 = p_2 = p \).

Note first, so long as \( P \geq p \), that for an advertiser that chooses to multi-home, it prefers to buy a guaranteed product on one publisher only. Namely,

\[(v - P)(D_i^I + D_i^S) + (v - p)D_i^I \geq (v - P)2(D_i^I + D_i^S) \quad (1)\]

By combining product types, an advertiser can impress every consumer without the additional expense and wasted impressions associated with guaranteed impressions on each publisher.

This immediately suggests a natural candidate equilibrium when guaranteed contracts are offered. Suppose that a set of high-value advertisers, \([v_m, 1]\), each demand a guaranteed contract from one publisher and a standard contract from the other. Also, assume that the advertisers are evenly split across publishers, therefore inducing an aggregate demand of (G) and (S) contracts equal to \( \frac{1}{2}(1 - F(v_m)) \). Then, the market will clear at a price that guarantees that \((D_i^I + D_i^S)a + D_i^I a = \frac{1}{2}(1 - F(v_m)) \) where \( v_m \) is such that \((v_m - P)(D_i^I + D_i^S) + (v_m - p)D_i^I = 0 \) as the marginal advertiser would also bid for that bundle across publishers. Thus, \( v_m = P(D_i^I + D_i^S) + pD_i^I \). This gives us one equilibrium equation with two unknowns. However, as the marginal advertiser could also bid for any of the four available products separately, it is clear that \( v_m = P \) and \( v_m = p \) so that, in
equilibrium, both the guaranteed and normal products have the same price. It should, therefore, be readily apparent that the allocation of advertisers to publishers is efficient and that each publisher earns profits of $F^{-1}(1-2a)a$; the same outcome as the perfect tracking benchmark.

Notably, this outcome results in allocative efficiency. Quality differences between publishers do not change this outcome. For instance, if $D'_1 > D'_2$, then high-value advertisers bid more for a bundle of guaranteed and standard impressions. However, as each component of the bundle is set by different publishers, the ability to substitute between them causes prices to be bid to equality. Hence, no sorting occurs.

This also suggests that if publishers can choose to have focused content, then a publisher that does so can offer all advertisers guarantees of impressing all switchers. So, if say, publisher 1 chose focused content (with $D'_1 = 0$), then it could sell a guaranteed product to all switchers. Publisher 2’s products would remain the same and, because a guaranteed product on publisher 2 would be equivalent to bundle of a guaranteed product on publisher 1 and a standard product on Publisher 2, impression prices on each publisher would be the same. Thus, publisher 1’s revenue would be $D'F^{-1}(1-2a)a$ while publisher 2’s revenue would be $(2D'_2 + D')F^{-1}(1-2a)a$. Therefore, the difference between the two is $2D'_1F^{-1}(1-2a)a$. Also, note that the thresholds in Proposition 4 would be $\zeta = (1-D')F^{-1}(1-2a)a < 2(1-D')F^{-1}(1-2a)a = \overline{\zeta}$. $\overline{\zeta}$ is the same as the baseline case considered in Proposition 4, while $\zeta$ is less than the same threshold in Proposition 4. Thus, the range of parameters where focused content provision becomes the equilibrium strategy for at least one of the publishers is expanded. As $D' \to 1$, all advertisers become single-homers when a guaranteed product is offered.\(^{30}\)

In summary, our baseline model reflects the current reality for online advertising markets as there are few outlets that offer guarantees of the type specified here. In addition, it is not readily apparent whether the result we obtain here is an artifact of our

---

\(^{30}\) What will change the outcome is if there are differences in ad capacities between publishers. In this case, the bundle across publishers could only be offered up to the minimum ad capacity. Beyond that point, additional capacity could not be sold as a part of the bundle and so the higher capacity publisher would sell the excess consumers to single-homing advertisers. This outcome is still allocatively efficient, however, but sorting means that the profits differ from the benchmark outcome.
simple environment and that future research examining more complex auction environments would yield a similar result. Nonetheless, it does demonstrate that a criteria that would lead us to conclude that guaranteed contracts are leading to efficient outcomes is that per impression prices of guaranteed and standard contracts will be the same; something also that does not exist in reality.

8 Switching and Advertising Levels

We now turn to consider the impact of consumer switching on the amount of impressions supplied by publishers and, in turn, prices. We study the strategic game in which the publishers choose simultaneously their capacity levels. As we shall see, the key driver of the publishers’ equilibrium strategies is that switching is a source of payoff externalities between them. Intuitively, different publishers provide advertisers substitutable means to reach switching customers. Imperfect substitutability creates, in equilibrium, upward pressure on advertising quantities much as it does in ordinary oligopoly settings. Indeed, the equilibrium we characterize is locally akin to a Cournot capacity game. The main hurdle is that with heterogeneous advertisers and consequent sorting over types, issues of non-monotonicity of best response functions arises; something that we will discuss in depth below. Nonetheless, we find that switching, generally, leads to higher advertising levels in equilibrium and, in turn, lower equilibrium prices.

This finding complements a large literature on the determinants of ad levels in media markets. As we shall see, a necessary ingredient of our theory is advertisers’ heterogeneity that, as discussed, implies in equilibrium advertisers sorting into different publishers (possibly none). The strategic considerations we focus on are therefore distinct from the “duplication” and “business sharing” considerations which drive the results in Ambrus et al. (2014) where advertisers are all alike and all served in equilibrium. Our

31 For some aspects arising in advertising exchanges and the design of auction for guaranteed slots, see Ghosh, McAfee, Papineni and Vassilvitskii (2009).
32 This has been a particular focus of models of media economics. Anderson and Coate (2005) modeled viewership levels as endogenous to the level of advertising on a publisher as those ads created a nuisance for consumers. They demonstrate that when the number of publishers increase (i.e., there was more publisher competition), advertisers would, in fact, reduce annoying advertising levels as a result of strengthened competition for consumers.
argument is also different from the “business stealing” argument of Anderson and Coate (2005), which, on the contrary, drives advertising levels downward. Business stealing relies on consumers being driven away by an increase in the extent of (annoying) advertising. All the above forces are suppressed here as advertising does not affect consumer demand.

Formally, we study how the fixed point of the best reply correspondence

$$a_i^*(a_j) := \arg\max a_i \quad 2a_ip_i(a_i,a_j)(D_i + \frac{1}{2}D')$$

changes when $D_s$ is allowed to exceed zero. $p_i$ denotes the market clearing price as a function of the (endogenous) capacity choices.33 (For simplicity, we assume in this section that $F(v) = v$.) In addition, to improve the exposition, we relegate the formal characterization of the equilibrium in the appendix. Specifically we derive there the closed-form expression of the unique Pure Strategy Nash Equilibrium. We focus here, instead, on illustrating the marginal incentives over capacity by studying the best reply of publisher $i$.

Consider first the benchmark case $D^s = 0$. As all consumers are exclusive, the problem reduces to a basic textbook monopoly problem. In the absence of switching, there are no externalities across publishers. Advertisers purchase $i$’s impressions if and only if $v \geq p_i$. One can easily derive that $p_i(a_i,a_j) = 1 - 2a_i$ which leads to $a_i^* = \frac{1}{4}$ at the monopoly price of 1/2.

Suppose now that $D^s > 0$ and for simplicity that the publishers are otherwise symmetric: $D^s_1 = D^s_2$. Figure 7 depicts the qualitative shape of the best reply correspondence of publisher 1. The four regions labeled A to D correspond to the different competitive regimes that arise as a function of publisher 2’s capacity. The figure reveals that capacities are at first strategic complements and then strategic substitutes. Figure 6 aids intuition showing how advertisers sort along the best reply locus.

The two extreme regions A and D correspond to the cases in which there is no competitive pressure. The two publishers corner different markets. To see which ones,

33 In the appendix, we show that the model is well behaved in that to each $(a_i,a_j)$ pair corresponds a unique pair of market clearing prices.
consider first the case where \( a_2 \) is close to zero. Then publisher 1 can sell all of its contracts at the monopoly price as if publisher 2 were not active. On the contrary, suppose that \( a_2 \) is arbitrarily high. Then publisher 1 can only hope to extract rents on those advertisers that multi-home. In both cases, the optimal quantity equals the monopoly level \( a_i^* = \frac{1}{4} \). In regions B and C, publisher 2, by increasing \( a_2 \), exerts a price externality on publisher 1. In region B, publisher 1 reacts by supplying the lowest capacity that keeps the rival out of the market for single-homers. It basically behaves as a constrained monopolist. Eventually (region C), publisher 1 accommodates entry and both firms compete head-on for single-homers. The local slope of the best reply reflects the basic Cournot intuition. The larger \( a_2 \), the lower the optimal \( a_1 \), as publisher 1 seeks to absorb some of the price decrease. Accordingly, we show that the unique feasible equilibrium outcome is a purely strategic one with \( a_1^* = a_2^* = \frac{1}{3} \) whenever \( D_1^l = D_2^l \).

As a final exercise, it is instructive to consider how viewership asymmetries impact on chosen advertising levels. Specifically, does the publisher whose share of switchers is relatively high choose more or less advertising than its rival?

We find that the publisher with the higher share of switchers chooses the higher advertising level. Recall that while there is a tendency for publishers to expand their capacity as there are more switchers, at some point, they can, by exercising market power, reduce the number of single-homing advertisers on their publisher to zero by restricting advertising levels. For a publisher with a smaller share of switchers, this is “easier” than one with a higher share of switchers. Consequently, in equilibrium, the publisher with a smaller share of switchers will choose a smaller advertising level than a publisher with a higher share. This type of outcome is depicted in Figure 8(b). The implication is that publishers with focused content are likely to have higher advertising levels per consumer than publishers with full content.

9 Conclusions and Directions for Future Research

This paper has analyzed the impact of the internet (in terms of increased consumers switching and imperfect tracking of these switches) on advertising market equilibrium and strategies for news media. The model predicts that the internet will be
associated with increased inefficient matching in advertising markets and a corresponding large fall in advertising revenues for display ads. This is, indeed, something that we have seen with regard to news media. A recent report of the Federal Communication Commission found that U.S. newspaper advertising revenues dropped 47% from 2005 to 2009.\(^{34}\) The ad revenue decline is pronounced even when controlling for obvious explanatory factors such as circulation, decline in revenues from classified ads and the business cycle.\(^{35}\)

The decline in advertising revenue has been almost unanimously attributed to the rise of the internet. However, the adverse impact of the web represents an economic puzzle because, in many respects, the forces influencing supply and demand appear to be as favorable for the industry, if not more so, than before. Online consumption of news media has created new and improved advertising products and services that should be, in principle, more valuable to advertisers (e.g., enhanced ads, targeting capabilities and improved measurement).\(^{36}\) Moreover, the internet dramatically increased the accessibility of many publishers for a wider audience.

A variety of theories have been proposed to explain the drop in advertising revenue. A common theme is that there is a glut in the supply of advertising space (Rice, 2010). However, this argument fails to account for the fact that, while there may be space for every advertiser on the web, those ads must still be viewed by actual consumers: human attention is naturally scarce, which limits the amount of advertising that can be supplied. Another theme is that online or digital ads are far less effective than ads that are on paper. However, the evidence is not consistent with that hypothesis (see Dreze and Hussersherr, 2003; Lewis and Reiley, 2009; Goldfarb and Tucker, 2011a, 2011b).

---


\(^{35}\) According to the Newspaper Association of America (www.naa.org), since 2000, total advertising revenue earned by its member U.S. newspapers declined by 57% in real terms to be around $27 billion in 2009. Much of this decline was in revenue from classifieds, but total display advertising revenue fell around 40%. In contrast, circulation over the same period declined by 18%. Ad revenue as a share of GDP also declined by 60%. According to ComScore, total U.S. display advertising revenue online was around $10 billion in 2010, which includes all sites and not just newspapers.

\(^{36}\) The internet has also created new types of opportunities such as “search ads.” However, many observers and regulators have noted that these new forms of advertising are complements rather than substitutes for the kind of advertising typically used by the news media; see Evans (2008, 2009). Chandra and Kaiser (2011) demonstrate that magazines which are better able to tailor content to specific consumer groups can continue to command a premium in ad rates and that this premium is associated with a consumer base with higher internet use.
Our paper is novel in that it focuses attention on the operation of advertising markets. Moreover, it generates a number of predictions regarding the impact of consumer switching, many of which have not heretofore been presented in the literature. The predictions include:

1. Consumer switching will result in an increased share of advertisers single-homing on individual publishers.
2. Consumer switching will be associated with a fall in publisher advertising prices and profits.
3. Publishers with higher readership shares will attract higher per-consumer advertising revenues.
4. Consumer switching increases a publisher’s incentives to invest in quality content that attracts a greater share of consumers.
5. Consumer switching reduces a publisher’s incentives to provide full content that can serve each consumer’s full attention and instead offer focused content; that is, reach becomes relatively more valuable than readership.
6. Consumer switching increases a publisher’s incentives to offer more complex contracts that guarantee advertisers’ impressions.
7. Consumer switching increases the advertising levels of publishers.
8. Publishers offering focused content (or content with lower loyal readership shares) will offer higher advertising levels than other publishers.

While empirical predictions 1 and 2 have been offered by other papers examining multi-homing consumers,\(^{37}\) predictions 1, 4 and 7 are distinct points of differentiation, while predictions 3, 6 and 8 are made possible by our model that holds consumer attention as fixed.

Nonetheless, the approach here is a first step. The model can be potentially enriched in a number of directions to offer insights into other aspects of online advertising markets. For example, throughout this paper, we have assumed that advertisements were equally effective on both publishers. However, in some situations, it may be that the expected value from impressing consumers on one publisher is higher than that from impressing consumers on another. For instance, consider (as in Athey and

\(^{37}\) See the reviews in Reisinger (2011) and Anderson, Foros, Kind and Peitz (2012).
Gans, 2010) a situation in which all advertisers are in a given local area. One local publisher publishes in that given area only, while the general publisher publishes across local areas.\footnote{Location is only one aspect upon which consumers and advertisers might sort according to common interests. Any specialized media content can perform this function and give a publisher a matching advantage over more general publishers.} Without the ability to identify consumers based on their location, a consumer impressed on the local publisher will still generate an expected value of $v$ to advertiser $v$, whereas one impressed on the general publisher will only generate an expected value of $\theta v$ with $\theta < 1$. In this situation, even if there are no switching consumers, advertisers on the general publisher will be paying for wasted impressions.

While this situation may be expected to generate outcomes similar to when readership shares are asymmetric, the effects can be subtle. A general publisher may have fewer consumers who are of value to advertisers, but it may also have a larger readership.\footnote{Levin and Milgrom (2010) argue that targeting may be limited because it conflicts with goals of achieving market thickness (see also Athey and Gans, 2010).} Also, when consumers switch between publishers, the switching behavior is information on those hidden characteristics. Thus, switching behavior may actually increase match efficiency. Consequently, the effects of tailored content, self-selection and incentives to adopt targeting technologies that overcome these are not clear and are likely to be an area where future developments can be fruitful.
10 Appendix

10.1 Supplementary analysis

For convenience, define $D_i(p_i, p_{-i}) := \int_0^1 F(v) p_i(v) dF(v)$ as the aggregate demand of publisher $i = 1, 2$ where we adopt the convention that $-i$ is equal to 2 if $i = 1$ and is equal to 1 if $i = 2$. Also, note that there must be an arbitrary real number, denoted $\overline{p}$, such that $D_i = D_{-i} = 0$ for all $p_i, p_{-i} \geq \overline{p}$.

10.1.1 Inverse demand under the assumption of uniformly distributed valuations

Let $F(v) = v$. Recall that without loss of generality we choose the index 1 to denote the relatively smaller publisher. That is we assume $D_1 \leq D_2$. Given a pair of prices, advertisers choose to either multi-home, single-home or do nothing. Let $v_{k,l}$ denote the advertiser indifferent between action $k$ and action $l$ where $k, l \in \{1, 2, 0\}$. For future reference define the following thresholds:

$$ a_{1L}(a_2) := \left\{ a_i \geq 0 : v_{1,2} = v_{2,0} \right\} \text{ and } a_{1H}(a_2) := \left\{ a_i \geq 0 : v_{1,2,2} = v_{1,2} \right\}. $$

(2)

Suppose that $a_1 < a_{1L}$ . Then it is possible to show that the only possible threshold configuration is $v_{1,2} \leq v_{2,0} \leq \max \{v_{1,2,1}, v_{1,0}\} \leq v_{1,2,2}$ . Intuitively if $a_1$ is very small relative to $a_2$ then high-type (aka multi-homing) advertisers will absorb the entire inventory. With this threshold configuration, the aggregate demand for 1 and 2, respectively, under the uniform distribution assumption, is $D_1(p_1, p_2) = 1 - v_{1,2}$ and $D_2(p_1, p_2) = 1 - v_{2,0}$ . Imposing $D_1 = 2a_1$ and $D_2 = 2a_2$ allows the recovery of prices:

$$ \hat{p}_1 = (1 - 2a_1) \frac{D_1'}{D_1' + D_2'} \text{ and } \hat{p}_2 = 1 - 2a_2. $$

(3)

Finally, plugging these prices into the thresholds $v_{k,l}$ allows recovering

$$ a_{1L} := a_2 - \frac{D_1' (1 - 2a_1)}{8D_1' + 2D_2'}. $$

(4)

and verifies that indeed, for all $a_1 < a_{1L}$ , the threshold configuration holds, so advertisers’ choices are indeed optimal and these prices clear the market for impressions. Given the uniqueness result above, then, this must be the only pair of market clearing prices for all $(a_1, a_2)$ such that $a_1 < a_{1L}$ . However, given the uniqueness result above, we need only verify that for all $a_1 < a_{1L}$ there is a pair of market clearing prices. So we get

$$ \hat{p}_1(a_1, a_2) = \frac{D_1'}{D_1' + D_2'} (1 - 2a_1) \text{ and } \hat{p}_2(a_1, a_2) = 1 - 2a_2 \text{ if } a_1 < a_{1L}. $$

(5)

Similarly, in order to recover the market clearing prices for $a_{1L} \leq a_1 \leq a_{1H}$ , it suffices to redo the exercise for the following threshold configuration: $v_{0,1} \leq v_{1,2} < v_{1,2,2} \leq v_{0,2}$ and $v_{0,1} < v_{0,2}$ . This allows recovering prices:
\[ \hat{p}_i(a_i,a_s) = \frac{2(1 - a_i - a_s)(4D'_i + D^*)}{8D'_i + 3D^*} \]  
(6)

\[ \hat{p}_i(a_i,a_s) = \frac{16D'_i(a_i(D'_1 - 2D'_1) + D'_3) + 6(D'_1 + D'_2 - 2a_iD'_1)D^* - 2(1 + a_i)(D')^2 - 2a_i(2D'_1 + D')^2(4D'_1 + D^*)}{(2D'_1 + D')(8D'_1 + 3D^*)} \]

and the threshold:

\[ a_{iH} = a_2 + \frac{D'(1 - 2a_i)}{8D'_1 + 4D^*}. \]  
(7)

Finally, suppose the thresholds satisfy \( v_{0,1} < v_{12,1}, \) \( v_{12,2} \leq v_{12,1}, \) \( v_{12,2} > v_{0,2} \) and \( v_{0,2} > v_{0,1}. \) This is the case in which publisher 2 serves only multi-homers and publisher 1 serves all active advertisers. Then, by the same token, the only candidate prices must satisfy is:

\[ \hat{p}_i(a_i,a_s) = 1 - 2a_i \] \ and \ \[ \hat{p}_s(a_i,a_s) = (1 - 2a_i) \frac{D'_1 + D'}{D'_1 + D^*} \]  
for \( a_i > a_{iH}. \)  
(8)

10.1.2 Endogenous capacity

**Proposition A2.** In the game where both publishers simultaneously set their capacities and \( F(v) = v, \) if a Pure Strategy Nash Equilibria exists then it is unique with

\[ a^*_1 = \frac{16(D'_1)^2 + 3D'_1(D' - 8D'_1) - D'(9D'_1 + D^*)}{40(D'_1)^2 + D'_1(6D' - 64D'_1) - 3D'(8D'_1 + D^*)} \]

and

\[ a^*_2 = \frac{8(D'_1)^2 - 16D'_1D'_1 - D'(6D'_1 + D^*)}{40(D'_1)^2 + D'_1(6D' - 64D'_1) - 3D'(8D'_1 + D^*)}. \]

For the special case in which \( D'_1 = D'_2, \) then the expressions simplify to \( a^*_1 = a^*_2 = \frac{1}{2}. \) Recall that \( (a^*_1,a^*_2) \) denotes a PSNE of the game. So firm \( i \) solves:

\[ a^*_i = \arg \max_{a_i} \hat{p}_i(a_i,a_s)2a_i. \]

Given the expression for the inverse demand derived above, the profit function is a piecewise function defined by three different subfunctions, each subfunction applying to three different subdomains, depending on whether \( a^*_i \) (given \( a^*_j \) ) is such that \( a^*_i < a_{iL}(a^*_j) \), such that \( a_{iL}(a^*_j) \leq a^*_i \leq a_{iH}(a^*_j), \) or such that \( a^*_i > a_{iH}(a^*_j). \) In what follows, we will refer to the different inverse demand functions that apply to the three sub-mentioned subdomains as \( \hat{p}_{iL}, \hat{p}_{iM} \) and \( \hat{p}_{iH}, \) respectively.

It can be verified by checking the second order derivatives that each subfunction is itself a strictly concave function of \( a_i \) for all admissible values of \( D'_1, D'_2 \) and \( D^*. \) Note that the objective function is piecewise continuous for all admissible values of \( D'_1, D'_2 \) and \( D^*. \)

In the first part of the proof we show that \( a^*_i \in (a_{iL}(a^*_j), a_{iH}(a^*_j)). \) This immediately implies that the global maximum of the profit function must be a solution to the usual system of first order conditions found by deriving the subfunction corresponding to the \( a^*_i \in (a_{iL}(a^*_j), a_{iH}(a^*_j)) \) case.
To show that $a^*_i \in (a_i^*(a_2^*), a_{i\theta}^*(a_2^*))$ is necessary, we proceed by contradiction. Suppose that $a^*_i > a_{i\theta}^*(a_2^*)$. If $D' > 0$, by definition (7) we get $a^*_i > a^*_i$. However, the first order condition of the subfunction that applies to publishers 1 and 2 when $a^*_i > a_{i\theta}^*(a_2^*)$ imply, respectively, $a_i^* = \frac{1}{4}$ and $a_2^* = \frac{1}{4}$, a contradiction. The same reasoning applies when $a^*_i < a_{i\theta}^*(a_2^*)$. Suppose now that $a_i^* = a_{i\theta}^*(a_2^*)$. To see that this can’t be optimal for publisher 1, we use lemma 1 (whose proof is relegated at the end for expositional purposes).

**Lemma 1.** Let $f: A \rightarrow \mathbb{R}$ denote a real valued piece-wise continuous function characterized by continuous and differentiable subfunctions $g(a)$ and $h(a)$ in the subdomains $a \leq a \leq a$, respectively. Suppose $h$ and $g$ are strictly quasi-concave and differentiable over the entire domain of $f$. Let $x^f$, $x^g$ and $x^h$ denote the unique maximizers of $f$, $h$ and $g$, respectively, over $A$ and suppose they all belong to the interior of $A$. If $x^g > x^h$ then $x^f \neq \tilde{x}$.

**Proof of Lemma 1.** Suppose that $x^f = \tilde{x} = x^h$. Then the left derivative of $f$ at $\tilde{x}$ must be equal to zero and since $x^g > x^h$ by assumption then the right derivative must be strictly positive, a contradiction. Suppose now that $x^f = \tilde{x} = x^h$, then the right derivative of $f$ at $\tilde{x}$ must be equal to zero while the left derivative must be positive. Finally, suppose $x^f = \tilde{x}$ and $x^g, x^h \neq x^f$. Then the left and right derivative of $f$ around $\tilde{x}$ must be strictly negative. This contradicts $x^g > x^h$. To see this, note that a negative left (right) derivative implies $x^g > \tilde{x}$ ($x^h < \tilde{x}$).

To verify the premises of lemma 1, we shall show that the maximum over $(0, \frac{1}{2})$ of the two different subfunctions that apply when $a_i < a_{i\theta}^*(a_2^*)$ and $a_{i\theta}^*(a_2^*) \leq a_i \leq a_{i\theta}^*(a_2^*)$ are ordered. By solving the two first order conditions that obtain when differentiating the respective subfunctions, one finds that the first maximum equal to $\frac{1}{4}$ is always lower than the second one whose lower bound obtains when $D' = 0$ and equals $\frac{1-a_i^*}{2}$. To see this, note that $a_2^* > \frac{1}{2}$ and the cross partial derivative of $p_{1\alpha}^*(a_i^*, a_2^*)2a_i$ w.r.t. $a_i$ and $D'$ is lower than zero, implying that the solution is increasing in $D'$. More precisely, one can verify that the unique stationary point of $p_{1\alpha}^*(a_i^*, a_2^*)2a_i$ is minimized over the set $\{(D_1^*, D_2^*, D') \geq 0 : D_1^* + D_2^* + D' = 1\}$ at $D' = 0$. Lemma 1, therefore, leads to a contradiction, and we conclude that $a_i^* \neq a_{i\theta}^*$. Finally, suppose that $a_i^* = a_{i\theta}^*(a_2^*)$. By definition $a_2^*$ must maximize:

$$\begin{align*}
\hat{p}_{2\alpha}(a_i^*, a_2)2a_2 & \quad a_i^* \leq \frac{D'(1-a_i^*)}{8D_1^* + 2D'} \\
\hat{p}_{2\beta}(a_i^*, a_2)2a_2 & \quad a_2 \leq \frac{D'(1-a_2^*)}{8D_2^* + 2D'} \\
\hat{p}_{2\gamma}(a_i^*, a_2)2a_2 & \quad a_i^* \geq a_2 + \frac{D'(1-a_a^*)}{8D_1^* + 4D'}
\end{align*}$$

Plugging $a_i^* = a_{i\theta}^*(a_2^*)$ reveals that the candidate equilibrium must be such that:
\[ a^*_i = a^*_2 + \frac{D^i(1-2a^*_2)}{8D^i_i + 4D^i}. \]

So one can apply Lemma 1 to firm’s 2 decision problem to establish an analogous contradiction. We conclude that if a SPNE exists then \( a^*_i \in (a^*_1, a^*_2, a^*_2) \) and

\[
(a^*_1, a^*_2) = \begin{cases} 
\frac{\partial p_{1i}(a^*_1, a^*_2)}{\partial a_1} 2a_1 = 0 \\
\frac{\partial p_{2i}(a^*_1, a^*_2)}{\partial a_2} 2a_2 = 0 
\end{cases}
\]

(10)

Solving the above system delivers the closed form expressions presented.

### 10.1.3 Deriving the “Reach function”

Suppose \( n_i \leq D^i_i + D^i \) denotes a random selection among all consumers who devote at least one unit of attention to publisher \( i \). It follows that the probability that a given consumer of \( i \) is informed on publisher \( i \) is equal to \( \frac{n_i}{D^i_i + D^i} \) and is constant across consumers. Suppose that the selections \( n_1 \) and \( n_2 \) are independent. It follows that the probability that a given switching consumer is included in both selections is \( \frac{n_1 n_2}{(D^1_i + D^1)(D^2_i + D^2)} \). The expected number of consumers informed is, therefore, equal to \( n_1 + n_2 \) minus the expected number of switching consumers included in both selections. That is, \( D^i \) times \( \frac{n_1 n_2}{(D^1_i + D^1)(D^2_i + D^2)} \).

### 10.2 Proofs

#### 10.2.1 Proof of Proposition 1

The advertiser set of actions (or choice set) is finite. The objective function is well defined for all \( p_1, p_2 \geq 0 \) and \( v \geq 0 \). It follows that the advertiser program always admits at least a solution.

Define \( D_i(p_1, p_{-i}) := \int_0^1 \mathbb{I}_i(v, p_1, p_2) dF(v) \) as the aggregate demand of publisher \( i = 1, 2 \). Adopt the convention that \(-i\) is equal to 2 if \( i = 1 \) and is equal to 1 if \( i = 2 \). Also, note that there always exists an arbitrary real number, denoted \( \mathcal{P} \), such that \( D_i = D_{-i} = 0 \) for all \( p_1, p_{-i} \geq \mathcal{P} \). Given this object, a market equilibrium exists if and only if there is a pair of prices such that \( D_i = 2a_i \) for \( i = 1, 2 \). We say that the equilibrium is “unique up to indifferences” if, given a pair of equilibrium prices, there is more than one family of solutions to the advertisers’ problem \( (\mathbb{I}_1(v, \hat{p}_1, \hat{p}_2), \mathbb{I}_2(v, \hat{p}_1, \hat{p}_2)) \), that gives rise to the same aggregate demand. The proof therefore reduces to showing the following.

**Proposition A1.** To each pair of positive capacities \( (a_1, a_2) \) corresponds a unique pair of prices \( (\hat{p}_1, \hat{p}_2) \) such that \( D_i(\hat{p}_1, \hat{p}_{-i}) = 2a_i \) for \( i = 1, 2 \).

The following properties, left unproved, follow immediately from the properties of the solution of the advertisers’ optimization problem.
D-1. $D_1$ and $D_2$ are well defined and upper hemi-continuous correspondences for all positive $p_1, p_2$.

D-2. (Strict monotonicity in own price.) If $p'_i > p_i$ and $D_i(p_i, p_{-i}) > 0$ then $D_i(p'_i, p_{-i}) - D_i(p_i, p_{-i}) < 0$ for $i = 1, 2$.

D-3. (Weak monotonicity in rival’s price.) If $p'_{-i} > p_{-i}$ then $D_i(p_i, p'_{-i}) - D_i(p_i, p_{-i}) \leq 0$.

D-4. (Strict monotonicity in both prices.) If $p'_i > p_i$ and $p'_{-i} > p_{-i}$ then the aggregate demand for at least one publisher should be strictly lower.

Define the following auxiliary mappings $p^*_i : [0, \bar{p}] \times [0, \bar{a} = \frac{1}{2}] \rightarrow [0, \bar{p}]$:

$$p^*_i(p_{-i}, a_i) = \left\{ p_i \geq 0 : D_i(p_i, p_{-i}) = 2a_i \right\} \text{ for } i = 1, 2.$$  

Note that this mapping is single-valued by D-2. That is, given a price $p_{-i}$, to each $a_i$ corresponds a unique $p_i$ such that the market for $i$’s impressions clears. Furthermore, the function is weakly increasing D-3. By definition, a pair of equilibrium prices is equivalent to any point in the square $[0, \bar{p}] \times [0, \bar{p}]$ at which the two auxiliary functions intersect. Existence follows, applying Tarski’s theorem to the auxiliary mapping $(p^*_i, p^*_j)$ from the compact set $[0, \bar{p}] \times [0, \bar{p}]$ into itself. Uniqueness follows from combining D-3 and D-4. That is, if two market clearing price vectors were to exist for given $(a_i, a_j)$, then weak monotonicity implies that one needs to be strictly larger than the other (component-wise), and this contradicts D-4.

10.2.2 Proof of Proposition 4

To simplify the notation we adopt the following conventions. As the game is symmetric, all publisher indexes are dropped. $D'$ and $D''$ denote $D'(d, d)$ and $D''(d, d)$ . It follows that $D'(f, d) = D'(d, f) = D' + D'' . \hat{p}_{f,d}$ denotes the market clearing price of the publisher’s inventory when it plays $f$ and the rival plays $d$ . $\hat{p}_{d,d}, \hat{p}_{d,f}$ and $\hat{p}_{f,f}$ are defined analogously. Let $\pi(\cdot, \cdot)$ denote advertising revenues net of cost as a function of the actions. Finally, let $\underline{c} := \{ c \geq 0 : \pi(d, d) - \pi(f, d) = c \}$ and $\bar{c} := \{ c \geq 0 : \pi(d, f) - \pi(f, f) = c \}$ . Suppose the following holds (strong strategic substitutability):

$$\pi(d, f) - \pi(f, f) > \pi(d, d) - \pi(f, d).$$

Then $\underline{c} < \bar{c}$. If $c > \bar{c}$, it is a dominant strategy not to invest and the unique Nash equilibrium is $(f, f)$. If $c < \underline{c}$, then it is a dominant strategy to invest and $(d, d)$ is the unique equilibrium outcome. If $\underline{c} < c < \bar{c}$, the only a pair of asymmetric pure strategy equilibria exists. That is, an equilibrium exists and is unique up to permutations of the indexes. When $c$ is equal to one of the two thresholds, then, the symmetric and asymmetric equilibria coexist. We are left to show that strong strategic substitutability (SSS) holds whenever $D' < 1$. Recall from the baseline model that $\pi(d, d) := 2aD' \hat{p}_{d,d} + aD'' \hat{p}_{d,d} = a\hat{p}_{d,d}$ . Next we have $\pi(f, d) := a(D' + D'') \hat{p}_{f,d}$; $\pi(f, f) := a\hat{p}_{f,f}$ and $\pi(d, f) := 2aD' \hat{p}_{d,f} + a(D' + D'') \hat{p}_{d,f} = a(1 + D') \hat{p}_{d,f}$. 
So (SSS) is equivalent to:

\[ a(1 + D') \hat{p}_{d,f} - a \hat{p}_{f,f} > a \hat{p}_{d,d} - a(D' + D') \hat{p}_{f,d} \iff \]

\[ a(\hat{p}_{d,f} - \hat{p}_{f,f}) > a(\hat{p}_{d,d} - \hat{p}_{f,d}) + aD'(\hat{p}_{f,d} - \hat{p}_{d,f}) \]

Recall also that the notion of positional advantage implies that \( \hat{p}_{d,f} > \hat{p}_{f,f} \) and \( \hat{p}_{f,d} > \hat{p}_{d,d} \). So the inequality holds if \( \hat{p}_{f,d} - \hat{p}_{d,f} < 0 \). That is, the case can be verified using (6), which under symmetry reduces to (using \( 2D' + D' = 1 \)):

\[ \hat{p}_{f,d} = \frac{2(1-2a)(4D'_1(f,d) + D'(f,d))}{8D'_1(f,d) + 3D'(f,d)} = \frac{2(1-2a)}{3}, \]

and

\[ \hat{p}_{d,f} = \frac{2(1-2a)(4D'_1(d,f) + D'(d,f))}{8D'_1(d,f) + 3D'(d,f)} = \frac{2(1-2a)(1+3D')}{3+5D'} \]

Note that \( \hat{p}_{f,d} - \hat{p}_{d,f} < 0 \) if and only if \( D' > 0 \) (or equivalently \( D' < 1 \)) and equal otherwise.

The last statement of the proposition follows from observing that with \( D' = 1 \) providing focused content has no impact on prices. That is \( \hat{p}_{d,f} = \hat{p}_{f,f} = \hat{p}_{d,d} = \hat{p}_{f,d} = \hat{p}_{f,d} = \hat{p}_{d,d} = \frac{2}{3}(1-2a) \) and hence \( \pi(d,f) - \pi(f,f) = \pi(d,d) - \pi(f,d) = 0 \), which implies \( \zeta = \xi = 0 \).
Figure 1: Estimates by the “Google Display Planner” on target audience for “gourmet coffee”

Figure 2: Bidding options for a hypothetical campaign on the GDN
Figure 3: Sorting with Symmetric Publishers

Case $D^s > 0$

\[ v_s = p \]

\[ v_m = \frac{D_1^l + \frac{1}{2} D^s}{D_1^l + \frac{1}{4} D^s p} \]

\[ v = 1 \]

Case $D^s = 0$

\[ v_s = p \]

\[ v = 1 \]

Figure 4: Sorting with Asymmetric Readership and Single Price

Case $D_1^l < D_2^l$, $D^s > 0$

\[ \hat{p} = \frac{D_1^l + \frac{1}{2} D^s}{D_1^l + \frac{1}{4} D^s \hat{p}} \]

\[ v = 1 \]

Figure 5: Sorting with Asymmetric Readership and Differential Prices

Case $D_1^l < D_2^l$, $D^s > 0$

\[ \hat{p}_1 \]

\[ \frac{\hat{p}_2 - \hat{p}_1 \delta}{1 - \delta} \]

\[ \frac{D_1^l + \frac{1}{2} D^s}{D_1^l + \frac{1}{4} D^s \hat{p}_1} \]

\[ v = 1 \]

\[ \delta := \frac{D_1^l + \frac{1}{2} D^s}{D_2^l + \frac{1}{2} D^s} \]; measure of relative size
Figure 6: Sorting with Endogenous Capacity

$v_{i,j}^{k,l}$ denotes the advertiser indifferent between action $k$ and action $l$ where $k, l \in \{12, 1, 2, 0\}$.

- **Region A**
  - $0 \rightarrow v_{1,0} \rightarrow sh_1 \rightarrow mh \rightarrow v_{12,1}$

- **Region B**
  - $v = 0 \rightarrow v_{1,0} \rightarrow sh_1 \rightarrow mh \rightarrow v_{12,1}$

- **Region C**
  - (case $a_j > a_i$)
  - $v_{j,0} \rightarrow sh_j \rightarrow v_{12,i} \rightarrow mh \rightarrow v_{12,1}$

- **Region D**
  - $v_{2,0} \rightarrow sh_2 \rightarrow mh \rightarrow v_{12,2}$

Figure 7: Best Reply of Publisher 1
Figure 8: Endogenous Capacity Equilibria

(a) $D'_1 = D'_2$ and $D' = 0.8$

(b) $D'_1 = 0$, $D'_2 = 0.2$ and $D' = 0.8$
11 References


