Discussion Paper Series – CRC TR 224

Discussion Paper No. 140
Project B 05

Ad Clutter, Time Use, and Media Diversity

Simon P. Anderson [1]
Martin Peitz [2]

July 2020
(First version: December 2019)

[1] Department of Economics, University of Virginia P.O. Box 400182, Charlottesville, VA 22904-4182, USA. sa9w@virginia.edu
[2] Department of Economics and MaCCI, University of Mannheim, 68131 Mannheim, Germany, Martin.Peitz@gmail.com; also affiliated with CEPR, CESifo, and ZEW

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Ad clutter, time use, and media diversity[1]

Simon P. Anderson[2]
University of Virginia

Martin Peitz[3]
University of Mannheim

July 2020

[1]An earlier version of this paper was circulated under the title "Advertising congestion in media markets". We received helpful comments from Emilio Calvano, Michele Polo, Markus Reisinger, Yossi Spiegel, and several seminar audiences. We thank Moonju Cho for spirited research assistance in the extensions, and D. J. Campbell for background simulations. Anderson gratefully acknowledges financial support from the National Science Foundation (1919493), Peitz from Deutsche Forschungsgemeinschaft (DFG) through PE 813/2-2 and CRC TR 224 (project B05).

[2]Department of Economics, University of Virginia P.O. Box 400182, Charlottesville, VA 22904-4182, USA. sa9w@virginia.edu.

[3]Department of Economics and MaCCI, University of Mannheim, 68131 Mannheim, Germany, Martin.Peitz@gmail.com; also affiliated with CEPR, CESifo, and ZEW.

Abstract

We introduce advertising congestion along with a time-use model of consumer choice among media. Both consumers and advertisers multi-home. Higher equilibrium advertising levels ensue on less popular media platforms because platforms treat consumer attention as a common property resource: smaller platforms internalize less the congestion from advertising and so advertise more. Platform entry raises the ad nuisance "price" to consumers and diminishes the quality of the consumption experience on all platforms. With symmetric platforms, entry still leads to higher consumer benefits. However, entry of less attractive platforms can increase ad nuisance levels so much that consumers are worse off.

JEL Classifications: D43, L13

Keywords: media economics, advertising clutter, limited attention, information congestion, two-sided markets

1 Introduction

Commercial media often rely exclusively or predominantly on advertising for revenue. Because platforms compete for viewers and viewers typically dislike advertising, one might think that more competition between media platforms should reduce ad levels. However, the last decades have seen a proliferation of television and radio channels — and now web-sites — and an increase of advertising time.[1]

Critics of mass media decry advertising clutter. Viewers are swamped with ads and ad impressions are wasted. We provide a novel framework that captures both aspects, namely that consumers dislike when content is replaced by advertising *and* they have a limited ability to absorb ads. The model predicts that small, low-quality media platforms feature more advertising minutes than more popular, higher-quality platforms. This result contrasts with the findings in the theoretical literature, and concurs with some casual evidence.[2] We also link advertising choices of media to media diversity. An increase in media diversity (platform entry) leads advertising to replace more content; as ad time reduces net program quality, there is a negative relationship between media diversity and media quality. Advertising becomes more congested making it more difficult for high-quality advertisers to reach consumers. Furthermore, despite a positive gain from variety, consumers can be worse off, as programming carries more advertising.

The standard model of two-sided markets as applied to media economics (Anderson and Coate, 2005; Anderson and Peitz, 2020) builds in a "competitive bottleneck" feature (Armstrong, 2006) which implies there is no direct competition for advertisers. Put briefly, when viewers single-home (meaning they patronize one platform), a platform has a monopoly position over delivering its viewers. The time-use model proposed in this paper

[1] For instance, a Nielsen study documents that a typical U.S. households watches 17.5 channels on a regular basis (see Joe Flint, "TV networks load up on commercials," Los Angeles Times, May 12, 2014). According to a UBS TV report that ad minutes per hour are at an all-time time for almost every network group (see Toni Fitzgerald, "Yes, you are seeing more commercials than ever before," Forbes, December 11, 2018).

[2] Broadcast networks have larger viewerships on average than cable networks. However, according to a Nielsen report with data from 2009 and 2013, the average commercial time on broadcast networks was less than on cable networks (see Joe Flint, "TV networks load up on commercials," Los Angeles Times, May 12, 2014). A UBS TV report authored by John Hodulik finds that network groups with small viewership carry more advertising, explaining: "Network groups with the worst ratings attempt to manage the pressure on advertising revenues with higher ad loads." (quoted in Toni Fitzgerald, "Yes, you are seeing more commercials than ever before," Forbes, December 11, 2018). By contrast, our paper offers an explanation based on profit maximization in market equilibrium.

exhibits the same feature. Even though we model multi-homing consumers who choose how much time to spend on each platform, at any point in time a particular viewer can only be reached through the single channel she is watching at that moment in time. As long as advertising across platforms is coordinated (so as to maximize ad effectiveness), platforms have monopoly power over advertisers. The competitive bottleneck means that competition among platforms is effectively competition for viewers, and so an increase in the number of platforms is predicted to decrease equilibrium ad levels, much like product prices decrease with the number of firms in standard oligopoly models of product competition. This model serves as the starting point of media economics, even though (as discussed by Anderson *et al.* 2012), empirical support for predictions stemming from this model are mixed.

Whereas the time-use model on its own does not change the structure of the media economics interaction, adding the next ingredient changes it quite radically. We enrich the standard media economics model by introducing limited viewer attention for advertising (congestion). This introduces strategic interaction among platforms on the *advertiser* side. Because of multi-homing, no media platform has exclusive access to a viewer's attention – it can be seen as a common property resource to which multiple media platforms have access. Therefore a platform which includes more advertising decreases overall ad effectiveness and thus exerts a negative externality on other platforms.

The upshot is to reverse the standard outcomes quite radically. Suppose that a platform cannot deliver a viewer with certainty to advertisers. Then, through the congestion function, one platform's choice of ad level will affect the willingness to pay for advertising on other platforms when viewers mix their media consumption. Large platforms internalize congestion to a larger extent than small platforms implying that the former have fewer ads and charge more for them.

Entry of a media platform in this setting will lead media platforms to internalize less of the negative congestion effect. Thus, more competition among media platforms will *increase* ad levels (which is in line with some observed market facts, such as the entry of Fox television).[3] This shows a tension between media diversity and media quality. Increasing diversity reduces the fraction of time consumers encounter content on any given platform — i.e. it increases the ad clutter. Our results speak to the connection between ad effectiveness and market structure. A more fragmented industry leads to less-effective advertising. While empirical work in psychology and economics has looked at

[3] See TV Dimensions 2000 (18th Ed), Media Dimensions, Inc.

limited attention for information (for a short overview, see Hefti, 2015), this paper points to the policy-relevant trade-off between ad clutter and media diversity. We have not seen empirical work taking a look at this issue.

Our results speak to the excessive entry results in oligopoly (von Weizsäcker, 1980; Mankiw and Whinston, 1986; Suzumura and Kiyono, 1987). There is a trade-off between the duplication of fixed costs versus additional variety and more intense use. However, entry in standard oligopoly markets increases consumer surplus because it intensifies competition. This is not necessarily the case here. Thus while consumers enjoy a larger variety of opinion with entry, the associated information is less reliable or nuanced since with higher advertising levels less time is available for information consumption. This latter effect reduces consumer surplus.

In addition, there is a reduction in total surplus because the average match quality between advertisers and consumers tends to deteriorate with entry. Since entry increases overall advertising volume, higher-value ad matches are partly replaced by lower-value matches. Hence, there is a negative total surplus effect with entry due to deteriorated match quality.

In our main model, consumers are still better off from entry when platforms are symmetric. Here, the positive direct effect from more variety dominates the negative indirect effect of lower media quality due to increased ad levels. However, if platforms are asymmetric consumers may actually suffer from platform entry. We consider two extensions under platform symmetry. If advertisers cannot extract the full surplus when selling their products, consumers may suffer from platform entry even under platform symmetry. If platforms charge viewers for subscription, total subscription payments can go up with entry, which constitutes an additional consumer loss. However, the net effect continues to be that viewers are better off with entry under platform symmetry. Depending on the timing of the game, equilibrium subscription prices are set according to "incremental value" or will be set lower to satisfy a "topple-free" condition.

Advertising congestion is related to the classic literature on common property resources and the strand of economics papers on information overload (van Zandt, 2004; Anderson and de Palma, 2009, 2012; Hefti and Liu, 2019). This paper brings information congestion into platform pricing using the approach proposed by Anderson and de Palma (2009). Specifically, it is assumed here that the viewer only has a limited attention span for ads, and is therefore only able to process a fixed number of all the ads to which she is exposed. This analysis renders endogenous the platform ad prices in the presence of congestion, as well as dealing with multiple platforms competing for attention.

Our paper contributes to the literature on media (e.g., Anderson and Coate, 2005; Peitz and Valletti, 2008; Crampes, Haritchabalet and Jullien, 2009) and two-sided platforms (e.g., Rochet and Tirole, 2003; Armstrong, 2006; Tan and Zhou, 2020) more generally by introducing limited viewer attention.

The paper is organized as follows. In Section 2 we spell out the model and in Section 3 we characterize the media equilibrium. In Section 4 we show what happens when media diversity changes. In Section 5 we analyze three extensions. In Section 5.1, we compare our model with limited viewer attention to the one with unlimited viewer attention and show that our results are robust to introducing a small fraction of viewers with unlimited attention. In Section 5.2, we allow for surplus sharing in the advertiser-consumer relationship. This strengthens our result in the sense that entry becomes more problematic from a consumer surplus perspective. This extension also allows us to endogenize ad digestion: in response to entry, viewers pay less attention to ads. In Section 5.3, we introduce subscription prices and establish conditions under which viewers pay more for subscriptions with entry. Section 6 concludes.

2 The Model

We consider a market in which media deliver viewer attention to advertisers. Consumers have a fixed attention span, ϕ. This simple formulation means that a consumer can absorb at most ϕ ads, and we assume that the ads that are retained are chosen randomly from those to which she is exposed (see Anderson and de Palma, 2009). Platform i broadcasts a_i ads (to be determined endogenously). Let λ_i denote the fraction of time a consumer spends on platform i (also to be determined endogenously), which is equivalently the probability she is found on platform i. Therefore the expected number of ads seen on platform i is $\lambda_i a_i$. With n platforms to visit, the expected total number of ads seen by a viewer is $A = \Sigma_{i=1}^{n} \lambda_i a_i$ so that the consumer's probability of *retaining* an ad from platform i is $\min\{1, \frac{\phi}{A}\}$ after being exposed to it.

We focus on situations in which the expected total number of ads A exceeds the viewer attention span ϕ so that there is congestion in equilibrium; our measure of ad congestion is $(A - \phi)/A$. Congestion can only arise in oligopoly because a monopoly media platform would never choose $a > \phi$.[4] Thus a monopolist would always price out congestion by

[4] To see this point, note that with $a > \phi$, a monopolist would only be able to sell an ad at price

delivering impressions with certainty to those with the highest willingness to pay, instead of widening the pool of advertisers.

Advertisers

Advertisers decide whether or not to place an ad on each platform i. We rank advertisers in terms of decreasing willingness to pay, p, to contact viewers and so $p(a_i)$ is the willingness to pay of the marginal (a_ith) advertiser conditional on making contact with the consumer. We assume that the demand for ads is well-behaved, so it is not too convex. We define the advertising demand elasticity as $\varepsilon(a_i) \equiv -\frac{p'(a_i)}{p(a_i)}a_i$.

Assumption 1 $p(a)$ is is twice continuously differentiable. The advertising demand elasticity (in absolute value) is non-decreasing, $\varepsilon'(a_i) \geq 0$, and takes values between 0 and 1.

In other words, we assume that Marshall's Second Law of Demand holds on the advertiser side.

If there are a_i ads on platform i, the ad price per effective viewer is the per-viewer willingness to pay of the marginal advertiser — i.e., $\frac{\phi}{A}p(a_i)$. This willingness to pay is the surplus generated by a advertiser-viewer match and, by assumption, is fully appropriated by the advertiser. The demand price for ads on platform i is then determined as the product of the probability that the viewer is on the platform when the ad is aired, that she retains the ad, and the willingness to pay, in sum $\frac{\lambda_i \phi}{A}p(a_i)$. Here, we implicitly assume that the likelihood of remembering an ad is independent of the particular product that is advertised; so it is independent of the advertiser's willingness to pay.[5]

Even though viewers multi-home, we assume that advertisers do not "waste" impressions. In a setting in which viewers switch between platforms over time, this requires that ad placements are perfectly synchronized.[6] Radio and television markets endogenously lead to this property if media platforms allow each advertiser to choose the time

$\frac{\phi}{a}p(a)$, where $p(a)$ is the advertiser demand price when a ads are broadcast, yielding profit $\phi p(a)$. With a downward-sloping ad demand, this choice is dominated by the choice $a = \phi$ yielding profit $\phi p(\phi)$ because p is decreasing in a.

[5] In our framework this assumption is natural because viewers get zero surplus in an advertiser-viewer interaction and thus are indifferent as to which ads they remember. More generally, one may want to allow for some correlation between product characteristics and the likelihood to recall an ad.

[6] For analyses of duopoly media markets, in which consumers use multiple channels and advertising is non-synchronized see Ambrus, Calvano, and Reisinger (2016) and Anderson, Foros, and Kind (2018). Athey, Calvano, and Gans (2018) develop a model with limited attention in which consumer multi-homing degrades the value of the advertising inventory. They show that the ad price decreases in the share of multi-homing consumers.

slot in which it advertises and thereby to manage advertising decisions across channels. An advertiser which is active on several platforms chooses the same time slot for all ads. It implies that an advertiser's ad can be viewed at most once by any given viewer even though viewers multi-home.[7] This is the most efficient use of an advertiser's advertising budget and thus the optimal choice of an advertiser.

Viewers

We propose a time-use model of media consumption with identical viewers who mix between media.[8] The outside option has index 0 and gives utility v_0^α per unit of time. It may stand for the alternative use of time off media or consumption of an advertising free public broadcaster. Demand follows from maximizing the utility function for media consumption

$$\max_{\lambda_0,\lambda_1,\ldots,\lambda_n} \sum_{i=1}^{n} [s_i(1-a_i)\lambda_i]^\alpha + (\lambda_0 v_0)^\alpha \qquad \text{s. t. } \sum_{i=0}^{n} \lambda_i = 1 \tag{1}$$

with $\alpha \in (0,1)$ so that viewers like to mix between different platforms (and the outside good). Here, λ_i is the fraction of time spent on platform i and s_i stands for the content quality offered by platform i. Only $s_i(1-a_i)$ is actual program content ("net quality"), due to the ads interjecting, so $s_i(1-a_i)\lambda_i$ captures the "quality-time" spent on platform i. The idea here is that the viewer only values the content part of a program and time spent on watching advertising gives a benefit normalized to zero. Viewers ingest $\lambda_i a_i$ ad on platform i. Thus, total ad ingestion is $A = \sum_{j=1}^{n} \lambda_j a_j$. Viewers have an attention span of ϕ for ads. This is the maximal number of ads they digest. If $\phi < A$ a fraction ϕ/A ads is digested. Accordingly, each viewer digests $\phi a_i/A$ ads appearing on platform i.[9]

Define $\tilde{\alpha} = \frac{\alpha}{1-\alpha} > 0$ and denote $\mathbf{a} = (a_1, a_2, ..., a_n)$. The fraction of time spent on

[7] Absent ad congestion, the coordination of advertising across platforms makes the model identical to a model in which heterogenous viewers single-home. Thus, multi-homing by itself will not change the results of the standard model with single-homing viewers; see also Peitz and Valletti (2008).

[8] We are not the first to propose a time-use model. For an alternative utility function, see Gabszewicz, Laussel, and Sonnac (2004).

[9] The consumer's time-allocation problem is the first step of a consumer's decision making. At the second step, she decides which of the advertised products to buy. Under the assumption of full rent extraction by advertisers these consumer choices at the second step do not affect the decision how much time to spend on each platform. One way to think of full rent extraction is that consumers have unit demand for each product and are willing to give up $p(a)$ units of an outside consumption good to consumer one unit of the product of advertiser a. To make sure that consumers can afford to buy all products it is sufficient to assume that income y is sufficiently large. Normalizing the price of the outside good to 1 it is sufficient to assume that $y > \int_0^1 p(a)da$.

platform i is

$$\lambda_i(\mathbf{a}) = \frac{(s_i(1-a_i))^{\tilde{\alpha}}}{v_0^{\tilde{\alpha}} + \Sigma_{j=1}^n (s_j(1-a_j))^{\tilde{\alpha}}}, \qquad i = 1, ...n, \tag{2}$$

while the time spent on the outside option is

$$\lambda_0(\mathbf{a}) = \frac{v_0^{\tilde{\alpha}}}{v_0^{\tilde{\alpha}} + \Sigma_{j=1}^n (s_j(1-a_j))^{\tilde{\alpha}}}. \tag{3}$$

This fractional demand system is in the vein of Luce (1959) and satisfies the "independence of irrelevant alternative" (IIA) property.

Inserting these expressions into (1), consumer surplus is, therefore

$$\begin{aligned} CS &= \left(v_0^{\tilde{\alpha}} + \sum_{i=1}^n [s_i(1-a_i)]^{\tilde{\alpha}} \right)^{1-\alpha} \\ &= \frac{v_0^{\alpha}}{\lambda_0^{1-\alpha}}. \end{aligned} \tag{4}$$

Equation (4) tells us that, for given parameters α and v_0 with $v_0 > 0$, consumer surplus depends only on the market share of the outside option. We immediately obtain the following lemma.

Lemma 1 *An increase in consumer surplus is equivalent to a decrease in the time spent on the outside option, λ_0.*

Therefore the direction of a change in equilibrium welfare (for example, of a platform program quality increase or platform merger) can be determined by time spent on platforms altogether. Hence time spent online is a simple observable measure that indicates that welfare has gone up when more time is spent.

Under symmetry and full coverage ($\lambda_i = 1/n$, $i \neq 0$), the consumer surplus is $n^{1-\alpha}(s(1-a))^{\alpha}$. Ignoring strategic effects (i.e., for given a) this surplus is increasing with entry since viewers are variety-loving.

Platforms

We analyze the platform balance problem of delivering reluctant viewers to advertisers. The profit function takes the form

$$\Pi_i = \begin{cases} \frac{\lambda_i \phi a_i p(a_i)}{A} & \text{for } \phi < A \\ \lambda_i a_i p(a_i) & \text{for } \phi \geq A \end{cases} \tag{5}$$

It depends on whether there is congestion. Clearly, a more-attractive outside option (i.e. v_0 and thereby λ_0 goes up) eats into viewing time on channels $i = 1, ..., n$. According to the IIA property, relative viewing time on the n media platforms, λ_i/λ_j for $i, j \in \{1, ..., n\}$, is unaffected and thus a more-attractive outside option lowers λ_i. This implies that a media platform's profit decreases in λ_0 when there is no advertising congestion.

By contrast, with congestion (i.e. $\phi < A$) profits are independent of λ_0. To see this, we rewrite profits as

$$\phi \frac{\lambda_i a_i}{\sum_{j=1}^{n} \lambda_j a_j} p(a_i) = \phi \frac{a_i}{\sum_{j=1}^{n} (\lambda_j/\lambda_i) a_j} p(a_i).$$

Since the IIA property says that λ_j/λ_i is independent of λ_0, profits do not depend on v_0. A better outside option implies that viewers spend less time on ad-financed media, but the probability that advertisers reach viewer increases. The two effects cancel each other out. In other words, while viewers get to see fewer ads when they stay less time tuned (i.e. when λ_0 is up), they continue to be attentive to a total of ϕ ads as long as there is still ad congestion. Profits depend on this number ϕ and the fraction of ads watched on the platform, $\lambda_i a_i$, over the total intake A.

We restrict attention to the case with ad congestion.[10] This means that ad digestion ϕ is less than total ad ingestion A. The degree of ad congestion can then be defined as $(A - \phi)/A$. Profits of platform i are the product of the viewing time spent on the platform, λ_i; the number of ads on the platform, a_i; and the per-viewer ad price $\frac{\phi}{A} p(a_i)$. These profits depend on competitors' decisions through two channels. Interdependence on the viewer side comes from the assumption that consumers decide how to allocate their viewing times λ_i. Interdependence on the advertising side comes from the joint assumption that the A ads are seen across multiple channels (because viewers are mixing between platforms) and that there is advertising congestion. Here, competitors' choices affect the total intake A and thereby the probability that an ad on platform i is digested by viewers which in turn is reflected in the per-viewer ad price $\frac{\phi}{A} p(a_i)$.

[10] For $\phi \geq A$, platform i's profit is $\Pi_i = a_i \lambda_i(\mathbf{a}) p(a_i) = R(a_i)\lambda_i(\mathbf{a})$ where Assumption 1 implies that the revenue per viewer, $R(a) = ap(a)$ is also well-behaved. The case $\phi > A$ resembles standard models of media platforms (e.g., Anderson and Coate, 2005). Aggregative game tools can be used as in Anderson and Peitz (2020). In the extension section, we look at the symmetric case.

The profit function has a kink at $a_i = \widehat{a}_i \equiv (\phi - A)/\lambda_i$ with the property that marginal profits jump downward; i.e.,

$$\left.\frac{\partial \Pi_i}{\partial a_i}\right|_{a_i \uparrow \widehat{a}_i} > \left.\frac{\partial \Pi_i}{\partial a_i}\right|_{a_i \downarrow \widehat{a}_i}.$$

Thus, there is a corner equilibrium with $\phi = A$ under some parameter constellations.

3 Analysis

The structure of the model enables us to cast the oligopoly interaction as an aggregative game. This construct was introduced by Selten (1971), and further developed by Acemoglu and Jensen (2013) and Anderson, Erkal, and Piccinin (2020) *inter alia*. For our purpose, an aggregative game is one in which players' strategic actions can be recast in such a manner as to render each player's payoffs as a function solely of its own action and the sum of all players' actions. The latter sum is termed the aggregate. The aggregative game construct enables considerable simplification by uncovering the basic structure so as to write the oligopoly problem as a two-dimensional problem (instead of the n dimensions one would generally have with n players). Equilibrium is then simply described as a fixed point, at which aggregate equals the sum of each player's action as a function of the aggregate. It is important to recognize that this does not just apply to symmetric situations. Indeed, payoff functions are allowed to be idiosyncratic: one of the main useful properties of the approach is that it leads to a tight characterization of individual actions as a function of players' differing fundamental characteristics (program quality in the model below). And, as we shall see, the analysis of free entry equilibrium is also readily enabled, even when infra-marginal players are asymmetric (this analysis draws on Anderson, Erkal, and Piccinin, 2020: an important complication of the current situation is that consumer surplus *cannot* be written as a function of the aggregate despite the IIA property of viewer demand).

3.1 Equilibria in media markets with ad congestion

Pursuant to the discussion above, we want to write platform i's profit $\Pi_i(\psi_i, \Psi)$ as a function of its own action ψ_i and the corresponding aggregate $\Psi = \sum_j \psi_j$. We will then proceed by determining the function $\psi_i(\Psi)$, which is the *inclusive best reply* that maps the aggregate into own action. Notice that a player's own action is part of the aggregate, contrasting this approach to the standard way to think about best replies as functions solely of the actions of others.

The primitive action variable for a platform is its ad level, a_i, so that we seek a monotonic transform of this variable to use as the action variable (in order to preserve the strategic equivalence of the game in actions and the game in ad levels).

For $\phi < \sum_{j=1}^{n} \lambda_j a_j$, finding an action variable to yield an aggregator function is some-what challenging.[11] The action variable $\psi_i = a_i[s_i(1-a_i)]^{\widetilde{\alpha}}$ is net-quality adjusted ad level (or ad quality time), which is defined on $[0, \overline{\psi}_i]$, where $\overline{\psi}_i \equiv \overline{a}[s_i(1-\overline{a})]^{\widetilde{\alpha}}$ and $\overline{a} = \arg\max_a a(1-a)^{\widetilde{\alpha}} = 1/(1+\widetilde{\alpha}) \in (0,1)$. Recall too that $\Psi = \sum_{j=1}^n \psi_j$. The profit of channel i is then:

$$\begin{aligned} \Pi_i &= \frac{a_i[s_i(1-a_i)]^{\widetilde{\alpha}}}{\sum_{j=1}^n a_j[s_j(1-a_j)]^{\widetilde{\alpha}}} \phi p(a_i) \\ &= \frac{\psi_i}{\Psi} \phi p(a_i(\psi_i)). \end{aligned} \tag{6}$$

where the ratio term in the first expression is i's ad share $\lambda_i a_i / \sum_{j=1}^n \lambda_j a_j$: notice the key property that the denominators from (2) cancel out. This implies that profits are independent of the attractiveness of the outside option as long as $\phi < A$. This means that we can allow viewers with heterogeneous valuations of their outside option, v_0, as long as the $\phi < A$ for all viewers including the one with the highest v_0 (who spends the least time watching advertising and thus has the lowest A).

Notice that ad quality time $a_i[s_i(1-a_i)]^{\widetilde{\alpha}}$ (from which we have drawn the aggregate) is hump-shaped. Nonetheless, the formulation still yields a viable aggregative game because p is decreasing, and so we can restrict attention to the increasing part of $\psi_i(a_i)$ along the inclusive best reply. That is, a platform will never choose a_i beyond the monopoly level

$$\overline{a} = \arg\max_{a_i} a_i[s_i(1-a_i)]^{\widetilde{\alpha}}$$

because to do so would mean ad minute exposure would be already decreasing. Thus, $\psi_i(a_i)$ can be inverted in the relevant range. We have

$$\frac{da_i}{d\psi_i} = \frac{a_i(1-a_i)}{\psi_i(1-(1+\widetilde{\alpha})a_i)} > 0, \tag{7}$$

where a_i is a function of ψ_i. We define the inverse action elasticity as $\eta(a_i) \equiv \frac{\psi_i}{a_i}\frac{da_i}{d\psi_i}$.

Lemma 2 *The inverse action elasticity $\eta(a_i)$ takes positive values and is increasing in a_i.*

Proof. The inverse action function elasticity simplifies to

$$\eta(a_i) = \frac{1-a_i}{1-(1+\widetilde{\alpha})a_i} > 0. \tag{8}$$

[11] Absent congestion, we could write profit as the function $\Pi_i\left(\tilde{\psi}_i, \tilde{\Psi}\right) = a_i(\tilde{\psi}_i)\tilde{\psi}_i p(a_i(\tilde{\psi}_i))/(v_0^{\widetilde{\alpha}} + \tilde{\Psi})$ with $\tilde{\psi}_i = [s_i(1-a_i)]^{\widetilde{\alpha}}$ and $\tilde{\Psi} = \sum_i \tilde{\psi}_i$ giving rise to a different aggregative game structure.

The derivative is $\eta'(a_i) = \widetilde{\alpha}/[1-(1+\widetilde{\alpha})a_i]^2 > 0$. ■

From (6), the first-order condition defining the inclusive best reply is (recalling that ψ_i enters Ψ)

$$p\left(\frac{1}{\Psi} - \frac{\psi_i}{\Psi^2}\right) + \frac{\psi_i}{\Psi} p' \frac{da_i}{d\psi_i} = 0, \tag{9}$$

where $\frac{da_i}{d\psi_i}$ is given as the reciprocal of (7). We can rewrite this expression as

$$1 - \frac{\psi_i}{\Psi} = -\psi_i \frac{p'}{p} \frac{da_i}{d\psi_i}$$

Recalling that $\varepsilon(a_i) = -p'(a_i)a_i/p_i$, the first-order condition in elasticity form can be written as

$$\frac{\psi_i}{\Psi} = 1 - \varepsilon(a_i)\eta(a_i), \tag{10}$$

and thence equilibrium profit is

$$\phi(1 - \varepsilon(a_i)\eta(a_i))p(a_i). \tag{11}$$

We can now show the following result.

Lemma 3 *For $p' < 0$, inclusive best replies $r_i(\Psi)$ satisfy $0 < r_i'(\Psi) < \psi_i/\Psi$ and thus there exists a unique equilibrium.*

Proof. We observe from (8) that

$$\begin{aligned} \frac{d\eta(a_i)}{d\psi_i} &= \frac{\widetilde{\alpha}}{[1-(1+\widetilde{\alpha})a_i]^2} \frac{da_i}{d\psi_i} \\ &= \frac{\widetilde{\alpha} a_i (1-a_i)}{\psi_i [1-(1+\widetilde{\alpha})a_i]^3} > 0. \end{aligned}$$

Furthermore, Assumption 1 states that $\varepsilon(a_i)$ is non-decreasing and so is $\frac{d\varepsilon(a_i)}{d\psi_i}$.

To determine the slope of the inclusive best reply, we differentiate $\Psi = \psi_i/[1 - \varepsilon(a_i)\eta(a_i)]$ with respect to ψ_i.

$$\frac{d\Psi}{d\psi_i} = \frac{1 - \varepsilon(a_i)\eta(a_i) + \psi_i[\varepsilon(a_i)\eta(a_i)]' da_i/d\psi_i}{[1-\varepsilon(a_i)\eta(a_i)]^2}. \tag{12}$$

Since $da_i/d\psi_i > 0$, the inclusive best reply is upward sloping if $[\varepsilon(a_i)\eta(a_i)]' > 0$, which holds since $\varepsilon(a_i) > 0$ and $\varepsilon'(a_i) \geq 0$ by Assumption 1 and $\eta(a_i) > 0$ and $\eta'(a_i) > 0$ by Lemma 2.

Using (10) we can rewrite equation (12) as

$$\frac{d\psi_i}{d\Psi} = \frac{\psi_i}{\Psi} \frac{1 - \varepsilon(a_i)\eta(a_i)}{1 - \varepsilon(a_i)\eta(a_i) + \psi_i[\varepsilon(a_i)\eta(a_i)]' da_i/d\psi_i}.$$

As shown in the previous paragraph, $\psi_i[\varepsilon(a_i)\eta(a_i)]' da_i/d\psi_i > 0$. Hence, $\frac{d\psi_i}{d\Psi} < \frac{\psi_i}{\Psi}$. ■

As we have just shown, inclusive best replies are upward sloping, so that actions are inclusive strategic complements with the aggregate. This also implies that ad levels are strategic complements. The economics of this property can be understood from the economics of a common property resource, in which participants have different interest shares (for instance, think of a common property fishery in which participants have different valuations of its continued health). Here the common property resource is consumer attention. The more that others (over-)fish it (i.e., advertise), the more it is degraded and the bigger an individual's incentive to do likewise in the dwindling value.

The second key property in Lemma 3 is that average action shares exceed marginal ones. This property implies that the sum of the actions has slope below 1 and so the equilibrium is unique. For later use, we also report the properties when inverse advertiser demand $p(a)$ is flat.

Remark 1 *The limit case $p' = 0$ has the feature that all platforms set their actions at the boundary of their action spaces; that is, where $\overline{\psi}_i \equiv \overline{a}[s_i(1-\overline{a})]^{\widetilde{\alpha}}$ and $\overline{a} = \arg\max_a a(1-a)^{\widetilde{\alpha}} = 1/(1+\widetilde{\alpha}) = 1 - \alpha \in (0,1)$. In this case, inclusive best replies are flat.*

In our analysis above we have assumed that advertisers post at most one ad per platform. If there were no congestion, and the market were "fully covered" (meaning that the outside option of non-purchase is not exercised) they would have no advantage from a second ad because they already get the consumer's attention with probability one by placing a (synchronized) ad on each platform. Otherwise though, there is a benefit from a second ad, or more.[12] Clearly, if the highest value advertiser does not want a second ad, then none do: denote by v $(= p(0))$ the uncongested demand price per viewer of this advertiser (the inverse demand function intercept). Consider the case when

[12] Several papers have addressed the effects on competition in the ad market when consumers multi-home and some advertisers place multiple ads on platforms, including Ambrus and Reisinger (2006), Ambrus, Calvano, and Reisinger (2016), Anderson, Foros, and Kind (2018), and Athey, Calvano, and Gans (2018). However, to consider entry, Ambrus and Reisinger (2006), only compare monopoly and duopoly in a Hotelling model (in which model the transition from one to two firms generically involves rather different aspects); Anderson, Foros, and Kind (2018) assume a fixed number of advertisers with the same willingness to pay for impressions and only obliquely allow for advertising nuisance to consumers.

an advertiser places a second ad on platform i in the same time bracket as its first ad (literally, an ad at the same time, a synchronized ad). The first ad is a "hit" with probability $\frac{\phi}{A}\lambda_i$. The second ad raises the chance of a hit by giving an extra chance of breaking into a consumer's perception. Conditional on the consumer being on platform i at the time (which happens with probability λ_i), with two ads the advertiser gets at least one ad through with probability $1-\left(1-\frac{\phi}{A}\right)^2$ (one minus the chance of neither ad getting through). So the conditional incremental probability is $\frac{\phi}{A}\left(1-\frac{\phi}{A}\right)$. This value is maximal for $\frac{\phi}{A}=1/2$, so we can use $\frac{\phi}{A}=1/2$ to find a (very loose) upper bound. Then we have that a second impression is not wanted if

$$\frac{v}{2}<p\left(a_i\right),$$

which can be interpreted as the requirement that the advertiser demand not be too heterogeneous over the relevant range: the marginal advertiser's uncongested willingness to pay should be at least half the willingness to pay of the advertiser with the highest willingness to pay. A similar (but modified) logic applies for asynchronous ads.[13] Thus, the advertiser demand must not be too heterogeneous over the relevant range. In this case, our analysis applies even when advertisers can place multiple ads on a platform.

[13] First note that an ad in a different time bracket will be a contender for reaching a previously unreached consumer with probability $\lambda_i\Sigma_{j\in N}\lambda_j$ (given that viewing times are random and independent) where N is the set of platforms on which the advertiser places ads (so N is all of them for the advertiser with value v). Then the chance of potentially hitting a consumer is $1-(1-\lambda_i)(1-\Sigma_{j\in N}\lambda_j)$, and so the extra chance is $1-(1-\lambda_i)(1-\Sigma_{j\in N}\lambda_j)-\Sigma_{j\in N}\lambda_j=(1-\Sigma_{j\in N}\lambda_j)\lambda_i$. Absent ad congestion, $v\lambda_i\lambda_0$ would therefore be the top advertiser's willingness to pay for an asynchronous second ad, as opposed to a willingness to pay of $v\lambda_i$ for the first ad. So second ads would not be aired if λ_0 were close to 1 because a first ad on each platform would almost surely do the job.

Now introduce ad congestion. An ad in the first time slot along with the other ads is a contender and hits while the second lone one is not a contender with probability $\frac{\phi}{A}(1-\lambda_i)\Sigma_{j\in N}\lambda_j$. The synchronous first one is not a contender while the second one hits with probability $\frac{\phi}{A}\lambda_i(1-\Sigma_{j\in N}\lambda_j)$. Both are contenders with probability $\lambda_i\Sigma_{j\in N}\lambda_j$, and, conditional on this, at least one hits with probability $1-\left(1-\frac{\phi}{A}\right)^2$. Adding up these terms, and subtracting the chance of a hit $\frac{\phi}{A}\Sigma_{j\in N}\lambda_j$ when just using the synchronized ads gives the incremental chance of a hit as $\frac{\phi}{A}(1-\lambda_i)\Sigma_{j\in N}\lambda_j+\frac{\phi}{A}\lambda_i(1-\Sigma_{j\in N}\lambda_j)+\frac{\phi}{A}\left(2-\frac{\phi}{A}\right)\lambda_i\Sigma_{j\in N}\lambda_j-\frac{\phi}{A}\Sigma_{j\in N}\lambda_j=\frac{\phi}{A}\lambda_i\left(1-\frac{\phi}{A}\Sigma_{j\in N}\lambda_j\right)$. Given that the price of an ad on platform i is $\frac{\phi}{A}\lambda_i p(a_i)$, the highest willingness to pay advertiser therefore does not want a second ad if $v\geq\frac{p(a_i)}{\left(1-\frac{\phi}{A}\Sigma_{j\in N}\lambda_j\right)}$.

3.2 Equilibrium characterization

In this subsection we elaborate upon the cross-section properties of the equilibrium. First, we analyze the link between equilibrium action ψ_i and the associated advertising level a_i.

Lemma 4 *In a congestion equilibrium, a larger quality s_i implies a larger action ψ_i and a lower advertising level a_i.*

Proof. From the first-order condition $\frac{\psi_i}{\Psi} = 1 - \varepsilon(a_i)\eta(a_i)$, we know that left-hand side is increasing in ψ_i (for given Ψ). Since ε is non-decreasing and positive by Assumption 1 and η increasing and positive by Lemma 2, we must have that the product $\varepsilon(a_i)\eta(a_i)$ is increasing in a_i. Thus, the right-hand side is decreasing in a_i. This establishes that if for any two platforms, in equilibrium, $\psi_i > \psi_j$ we must have $a_i < a_j$ and vice versa. Furthermore, the inclusive best reply is larger for higher s_i. Therefore, the equilibrium value of ψ_i is increasing in s_i and a lower equilibrium value a_i is observed for a higher-quality platform. ■

This is illustrated in Figure 1. In the figure, we consider a media market with two platforms in which platform 2 has higher quality and thus is the larger platform in equilibrium. The equilibrium value of the aggregate, Ψ^*, is determined by the intersection of $\psi_1 + \psi_2$ with the diagonal. Equilibrium values of ψ_1 and ψ_2 are then obtained by the value of ψ_i evaluated at the inclusive best replies evaluated at Ψ^*. We can then represent equilibrium values $\frac{\psi_i^*}{\Psi^*} = 1 - \varepsilon(a_i^*)\eta(a_i^*)$. The function $\varepsilon\eta$ is increasing and hence we can infer equilibrium values a_1^*, a_2^* with $a_1^* > a_2^*$.

As the previous Lemma has established, ψ_i and a_i are negatively related in equilibrium. Since $\psi_i = a_i\lambda_i[v_0^{\tilde{\alpha}} + \Sigma(s_j\,(1-a_j))^{\tilde{\alpha}}]$ (where the term in square brackets is the same for all platforms), a platform with larger action ψ_i must have a larger market share, and moreover, the larger actions emanate from platforms with larger s_i. Hence we have:

Proposition 1 *Consider any two platforms i and j. In congestion equilibrium, (i) $s_i > s_j$ implies that $\lambda_i > \lambda_j$ and $a_i < a_j$ and (ii) $s_i = s_j$ implies that $\lambda_i = \lambda_j$ and $a_i = a_j$.*

From the proposition, a larger λ_i entails a smaller a_i, and, hence, a larger price per ad per viewer, $\frac{\phi}{A}p\,(a_i)$.[14] This results is in line with some empirical findings. Fisher et al. (1980) find that the per-viewer fee of an advertisement on programmes with more viewers

[14] Even though a larger platform has fewer ads, it is more profitable than a smaller one. To see this, recall that $\Pi_i = \phi\frac{\psi_i}{\Psi}p\,(a_i)$. A larger platform entails both a larger ψ_i and a larger $p(a_i)$, so its profits must be larger. This result can also be derived from the maximized value function, which

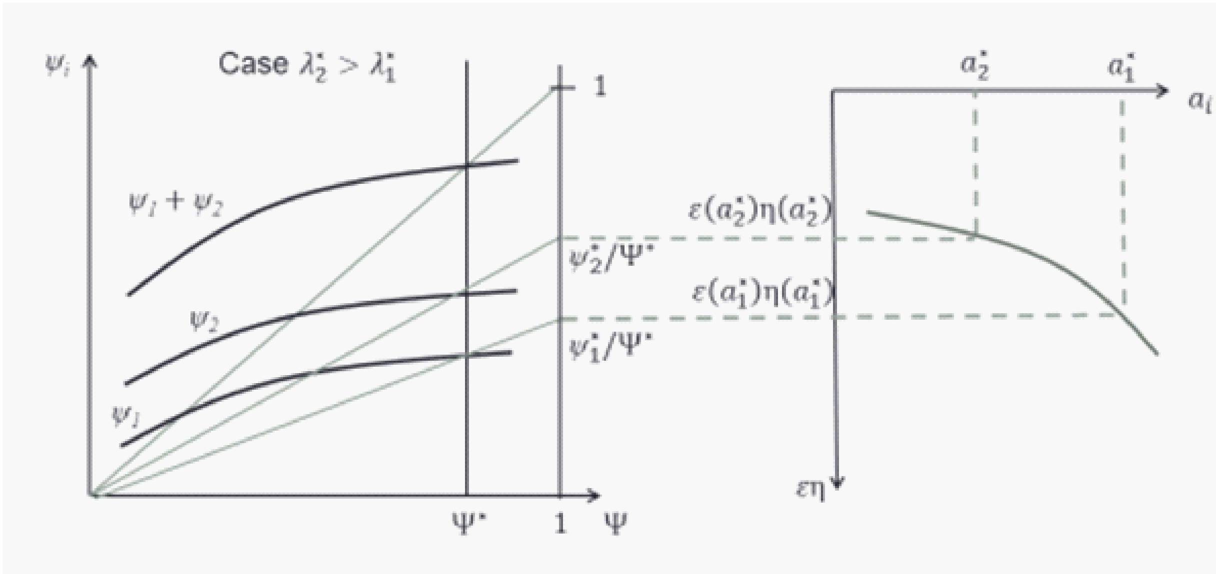

Figure 1: Relationship between λ_i, ψ_i, and a_i

is larger. It is also consistent with the "ITV premium" noted by other authors (see e.g. the discussion in Anderson et al., 2012). It is also a form of cross-sectional "see-saw" effect: interpreting a_i as the "price" paid by viewers, then this price is high when the price per ad per viewer (on the other side of the market) is low. Indeed, as argued in Anderson and Peitz (2020), the viewer single-homing model (and, by extension, the version of the current time-use model without congestion) exhibits a see-saw, but induced by quality differences in the opposite direction. That is, while a quality advantage induces a higher market share in both cases, the platform in the no-congestion case has more ads and a consequently lower price per ad per viewer.

Proposition 1 says that a platform uses a quality advantage to take a higher equilibrium market share. This effect is reinforced because it also wishes to carry a *lower* ad level. Market shares are therefore more dispersed than the quality levels that drive them (the ratio of high to low shares exceeds the ratio of their qualities). Put another way, the distribution of market shares has greater variance than the quality distribution. This is a type of "superstar phenomenon". In standard one-sided oligopoly models (e.g. the logit model of differentiated products in Anderson and de Palma, 2001), higher qualities

writes $\Pi_i = \phi \frac{\psi_i(\Psi)}{\Psi} p\left(a_i \psi_i\left(\Psi\right)\right)$. This function is decreasing in Ψ so that larger values of the aggregate constitute greater competition, and hurt profit: see Anderson, Erkal, and Piccinin (2020) for more on the competitiveness property. Also, higher values of s_i entail higher profit.

are parlayed into *both* higher qualities and higher mark-ups, which mutes market share variance as compared to quality variance. The same is true for standard models of media competition (see Anderson and Peitz, 2020) in which "better" programs want to broadcast more ads than inferior rivals. The result here is due to the congestion effect.

The congestion effect works by giving higher quality platforms a greater stake in not bloating overall congestion. As mentioned earlier, consumer attention is treated as common property, so that a platform with a higher quality catering to a larger market base has a bigger incentive than smaller rivals to internalize the extra congestion from its ads. It therefore wants to broadcast fewer ads. Both effects combine to give a higher price per ad: price per ad per viewer is higher, and also the viewer base is bigger.

While this price per ad effect is empirically well supported (e.g., it is implied by the ITV premium discussed above), casual evidence on the ad/viewership relation seems quite mixed. There are clearly high quality publications with few ads, and many late-night TV programs seem to carry many ads. Our analysis suggests that such results should be seen in markets where congestion effects are strong enough.

4 Media diversity

Here we look at the effects of adding more varieties, i.e., platform entry, on ad levels and ad congestion. We consider market environments in which there is congestion before and after entry. The consumer surplus analysis is more intricate, so we defer it to the subsequent subsection. We first look at the effects on incumbents.

4.1 Entry and incumbent platforms

To evaluate the effect of changes in the market, we have to understand how the aggregate changes. Under entry an additional platform will contribute by adding a new term to the aggregate. Hence the equilibrium value of the aggregate after entry must be larger than before. By Lemma 3, inclusive best replies slope up, and so individual actions of incumbent platforms rise. Ad levels rise too because they vary directly with actions (see (7)). Hence we have:

Proposition 2 *Platform entry raises advertising levels a_i on all channels.*

This unambiguous result holds even though platforms compete for viewers and a larger a puts them at a disadvantage. Thus the externality effect through congestion dominates the competition effect.[15] Since more advertising reduces the quality time on consumers experience on platforms, consumers may reduce their time spent with those media or reshuffle their demand towards platforms with less advertising. Thus a priori it is unclear whether consumer ingest more ads.

Proposition 3 *Platform entry increases advertising ingestion A.*

Proof. Advertising ingestion is $A = \sum_{i=1}^{n} \lambda_i a_i = \sum_{i=1}^{n} \psi_i / [v_0^{\tilde{\alpha}} + \Sigma(s_j(1-a_j))^{\tilde{\alpha}}] = \Psi / [v_0^{\tilde{\alpha}} + \Sigma(s_j(1-a_j))^{\tilde{\alpha}}]$. The numerator goes up with entry since Ψ increases. The denominator goes down with entry since the a_i increase with entry by Proposition 2. ■

Since advertising ingestion increases, advertising *congestion* $(A-\phi)/A$ goes up as well. Stronger competition among platforms leads to wasteful advertising, as a constant number of ads enter the attention span of viewers and thus an increasing fraction is purely wasteful from a welfare perspective (as it replaces valuable content). In addition, it decreases the match quality for advertisers as some high-value advertisers are replaced by lower value advertisers in a consumer's consciousness.

The economics here are once more best represented by reference to the common property problem. When more agents claim the common property resource through entry, each exploits it more because it internalizes the effect of its actions over a smaller base.

The effects on ad prices are quite interesting. First, because the ad level goes up on each platform i, then the "uncongested" price per ad per minute, $p(a_i)$, goes down. Because A rises, then each incumbent's full price per ad per minute, $\frac{\phi}{A} p(a_i)$, goes down by a further percentage. Finally, because market shares are lost to the new rival, the price per ad, which is $P_i \equiv \lambda_i \frac{\phi}{A} p(a_i)$ goes down even more still. Therefore all prices on the ad market tumble. One might indeed expect that prices should fall with entry, but, as noted in the Introduction, standard media economics models predict ad prices *per viewer* to rise (ad prices are ambiguous because of the share effect). Because of the "competitive bottleneck" problem (Armstrong, 2006), competition for viewers is dominant and this leads entry to reduce the "price" paid by viewers to fall – that price is the number of ads suffered. It is because ad levels fall that price per ad per viewer rises as we go back up the demand for ads relation.

With congestion, matters are much different – the see-saw effect with respect to prices works in the opposite direction completely. Indeed, with congestion effects, we have seen

[15] Softer forms of congestion in which attention increases with ad ingestion could yield ambiguous effects.

in Proposition 2 above that ad levels rise too, meaning that the implicit price paid by consumers on each channel goes up. In turn, consumers ingest more ads and new media entry has a harmful effect on consumers. The upshot is thus that new entry leads to the consumption experience deteriorating on each channel as the amount of non-advertising minutes in the program goes up. More competition causes worse advertising clutter on incumbent channels and more overall advertising overload. Whether this pricing effect can overturn the per se benefits of new options is the topic of the next sub-section.

While mergers are not our prime concern in this paper, it is instructive to track how they change ad prices. First, a merged entity has a larger stake in the common property and so it reduces its ad levels (it reduces its actions). This raises the uncongested price per viewer, with a further fillip from the reduced overall ad level. Ad levels on rival programs fall too, by strategic complementarity and as rivals now get a bigger stake in the total ad level. Insofar as the market share of the combined entity rises, taking customers from both the non-viewing option and the rivals, then *ad prices go up*.[16] The higher prices from merger play out on the advertiser side of the market (as opposed to on the viewer side, which is the case in the competitive bottleneck setting). The see-saw now works in favor of consumers who face less advertising clutter across the board.

4.2 Consumer Surplus

The equilibrium advertising per channel increases with more platforms, as we argued above. This ad level effect reduces consumer surplus. A utility-maximizing consumer adjusts its consumption pattern to changes in advertising level across all platforms. However, since all platforms offer lower quality time after entry, this must reduce consumer surplus. Hence, if entry simply means that one extra platform includes advertising together with its programming, the effect of entry on consumer surplus is unambiguously negative. Instances of such entry is that a zero advertising restriction on a public broadcaster is removed. However, entry of an additional platform changes the diversity of content and thus has an additional effect on consumer surplus.

Including the variety effect, the effect of entry on consumer surplus is not obvious. As in standard differentiated products oligopoly, entry increases product variety, which is something consumers like. In the standard oligopoly context, entry also leads to lower prices, which is also something consumers like. In a media context the corresponding

[16] However, advertisers with higher willingness to pay may be better off because of the reduced congestion.

result would be that consumers suffer from less nuisance after the entry of an additional media platform. While this property holds in the Anderson and Coate (2005) framework (see also Anderson and Peitz, 2020), this is not the case in our current setting with advertising congestion, as has been shown above.

Thus, we have to evaluate the overall effect of entry on consumer surplus. This, as we noted earlier, is not a simple function of the aggregate (in contrast to the central CES/Logit examples in Anderson, Erkal, and Piccinin, 2020). We start by considering a symmetric setting. Under symmetry consumer surplus is $[ns^{\tilde{\alpha}}(1-a)^{\tilde{\alpha}} + v_0^{\tilde{\alpha}}]^{1-\alpha}$. The conflict is this. Consumer surplus moves the same way as $ns^{\tilde{\alpha}}(1-a)^{\tilde{\alpha}} = \frac{\Psi}{a}$. However, Ψ rises with entry, while a rises too, so it is ambiguous a priori. The next result determines the net effect given Assumption 1.

Proposition 4 *In a symmetric market with at least two platforms, under Assumption 1, the entry of an additional platform always increases consumer surplus.*

Proof. Because consumer surplus tracks $ns^{\tilde{\alpha}}(1-a)^{\tilde{\alpha}}$, the effect of entry on consumer surplus, dCS/dn, is positive if and only if

$$s^{\tilde{\alpha}}(1-a)^{\tilde{\alpha}} - n\tilde{\alpha}s^{\tilde{\alpha}}(1-a)^{\tilde{\alpha}-1}\frac{da}{dn} > 0,$$

which is equivalent to

$$(1-a) - n\tilde{\alpha}\frac{da}{dn} > 0. \tag{13}$$

Using (10), (8), and the fact that, under symmetry, $\psi_i/\Psi = 1/n$, the equilibrium advertising level as a function of firms n is

$$a = \frac{n-1-n\varepsilon(a)}{(\tilde{\alpha}+1)(n-1)-\varepsilon n}. \tag{14}$$

Hence (after simplifying)

$$\frac{da}{dn} = \frac{\varepsilon(a)/(n-1)^2}{\frac{n}{n-1}\varepsilon'(a) + \frac{\tilde{\alpha}}{a^2}} > 0,$$

and, using this expression in (13), because $\varepsilon'(a) \geq 0$, it suffices to show that

$$(n-1)[(\tilde{\alpha}+1)(n-1)-\varepsilon n] - \tilde{\alpha}n\varepsilon > 0. \tag{15}$$

For $a > 0$ to hold in (14), we must have $n\varepsilon(a) < (n-1)$. Given this restriction, (15) holds if (the inequality is implied by):

$$(n-1)[(\tilde{\alpha}+1)(n-1)-(n-1)] - \tilde{\alpha}(n-1) > 0 \Leftrightarrow (n-2)\tilde{\alpha}(n-1) > 0,$$

as we desired to show. ■

By Lemma 1, consumer surplus and the time spent on the outside platform move in opposite directions. Thus, the above proposition also says that under symmetry, with platform entry, viewer spend less time with the outside option.

We have two main results for consumer surplus under entry. Proposition 4 states our first result: if platforms are symmetric, more entry must raise consumer surplus. Here, the variety effect outweighs the quality degradation on platforms. Our second result is that this no longer necessarily holds true with asymmetric platforms. As we show by example, in the presence of low-quality and high-quality platforms, entry of low-quality platforms can *reduce* consumer surplus.

To get to this result, we engage a Zero Profit Symmetric Entry Equilibrium (ZPSEE), following Anderson, Erkal, and Piccinin (2020). This is a free entry equilibrium at which profits are zero for marginal entrants, and such entrants have the same pay-off functions as each other (although infra-marginal firms may have different pay-off functions). In our current context, we let the marginal entrants all have low quality, s_L, while infra-marginal ones have higher quality, s_H (so we assume just two types). We take $\varepsilon \in (0, 1)$ constant, and will make some restrictions below. Here, we postulate that there are n_H high-quality platforms and that there is an unlimited supply of low-quality platforms

The key to determining the ZPSEE is to write the zero-profit condition of the marginal entrants (denoted by L subscripts because they are the lowest qualities around). Then we can uniquely determine their (common) ad-level. Let their entry cost be K. From the optimized profit (11) and using (8) we have

$$\phi(1 - \varepsilon \frac{1 - a_L}{1 - (1 + \tilde{\alpha})a_L})p(a_L) = K \tag{16}$$

which uniquely determines a_L because the LHS is the product of two terms that are positive and decrease in a_L. (Hence a larger K means lower ad levels across the board – for intuition, there are fewer fringe firms, they advertise less and the others come down with them, by strategic complementarity).

We can determine how many fringe firms there are once we know the qualities of other platforms. The solution is recursive: we illustrate with the case in hand when there are two types of platform. Indeed, now we know a_L from (16), we find a_H from the inclusive best replies. To see this, first write the inclusive best reply (10) as

$$\begin{aligned} \Psi &= \frac{\psi_i}{1-\varepsilon(a_i)\eta(\psi_i)} \\ &= \frac{a_i[s_i(1-a_i)]^{\tilde{\alpha}}}{1-\varepsilon\eta_i} \end{aligned} \tag{17}$$

where we recall that $\eta_i = \frac{1-a_i}{1-(1+\tilde{\alpha})a_i}$ from (8) and hence the RHS is equated across platforms. Note further that the RHS is an increasing function of a_i: both numerator and denominator are positive; the numerator is increasing in the relevant range; and the denominator is decreasing (as already argued above).

Notice that the equilibrium value of the aggregate, Ψ is found from (17) once we know a_L. Then the above relation tells us a_H. We treat the number of high-quality platforms, n_H, as exogenous (they earn more than low ones, and if their entry cost is the same, they would obliterate the low ones, so we restrict their number). Then the last parameter we need to find is the endogenous n_L. This is found from the aggregate fixed point condition, namely that

$$n_L\psi_L + n_H\psi_H = \Psi$$

or, rearranging this from (17) we find n_L from:

$$n_L(1-\varepsilon(a_L)\eta_L) + n_H((1-\varepsilon(a_H)\eta_H)) = 1. \tag{18}$$

Our objective is to compare consumer surplus with only high quality platforms present with the situation when both types are present. Notice that we can take a monotone transformation of the consumer surplus expression (4), and henceforth we use this transformation (with a slight abuse of notation).[17] When only high type platforms are present, we have

$$CS_H = n_H[s_H(1-a_H)]^{\tilde{\alpha}}, \tag{19}$$

and when both type are present we have

$$CS_B = n_H[s_H(1-a_H)]^{\tilde{\alpha}} + n_L[s_L(1-a_L)]^{\tilde{\alpha}}. \tag{20}$$

Define now

$$\Omega_i \equiv \frac{1-\varepsilon\eta_i}{a_i} = \frac{1-\varepsilon\frac{1-a_i}{1-(1+\tilde{\alpha})a_i}}{a_i}, \tag{21}$$

[17] We can thus ignore the power and the outside option (as long as these values are not changing in the comparison).

and note that Ω_i is a ratio of positive functions; the numerator is decreasing, while the denominator is increasing, so that Ω_i is decreasing in a_i. Using the relation between qualities from equation (17) above, we get

$$[s_L(1-a_L)]^{\tilde{\alpha}} \frac{\Omega_H}{\Omega_L} = [s_H(1-a_H)]^{\tilde{\alpha}}. \tag{22}$$

Then we can write

$$CS_B = \left(n_H \frac{\Omega_H}{\Omega_L} + n_L \right) [s_L(1-a_L)]^{\tilde{\alpha}}.$$

Using the fixed point condition $n_L \Omega_L a_L + n_H \Omega_H a_H = 1$, which condition determines the number of entrants (see (18)), we can rewrite this consumer surplus as

$$CS_B = \left(n_H \frac{\Omega_H}{\Omega_L} + \frac{1 - n_H \Omega_H a_H}{\Omega_L a_L} \right) [s_L(1-a_L)]^{\tilde{\alpha}}. \tag{23}$$

Since we tied down the a_L and a_H above, this expression then only depends on exogenous parameters (recall that we are treating n_H as exogenous).[18]

We can now use the above analysis to deliver the following result; the proof by example is relegated to the Appendix.

Proposition 5 *There are asymmetric markets in which the entry of an additional platform decreases consumer surplus.*

Entry is bad for consumers in this example because the ad-clutter degrades programs too much, even despite the extra variety. As we noted in the proof, we need a sufficiently low value for the low-quality types in order to overturn the result for symmetry that entry is beneficial. By Lemma 1 entry here has the effect that consumers spend more time with the outside option (i.e. λ_0 increases with entry).

So far we have been silent about the effect of platform entry on advertiser surplus. We now briefly discuss what happens to advertiser net surplus in a symmetric setting. As per Proposition 2, ad levels increase with entry. This tends to be good news for advertisers slightly above the marginal advertiser after entry. In particular, if inverse advertiser demand is strictly downward sloping, there is a positive mass of advertisers whose net surplus increases after entry. However, entry also increases congestion which reduces the probability that advertisers reach consumers. This tends to be bad news for top advertisers.

[18] Notice that if all platforms were low quality, then the term in parentheses is just $\frac{1}{\Omega_L a_L} = \frac{1}{1-\varepsilon\eta_L}$, which is the number of platforms, recalling that under symmetry $\frac{\psi}{\Psi} = \frac{1}{n} = 1 - \varepsilon\eta$ by (17).

In the special case of two types of advertisers, high-valuation advertisers and low-valuation advertisers (with valuations p_H and p_L, respectively), we obtain an unambiguous result on advertiser surplus. Suppose that market conditions are such that some low-valuation advertisers are active before and after entry. Their presence ties down the uncongested ad price as p_L and the per-viewer ad price will be $(\phi/A)p_L$. The per-viewer net surplus of each high-valuation advertiser is thus $(\phi/A)(p_H - p_L)$. By Proposition 3, A increases with entry and thus high-valuation advertisers are worse off after entry (for given λ_0). This shows that when viewers do not spend time on the outside option net advertiser surplus decreases with entry in this two-type model.

5 Extensions

We here consider two extensions to the model. First we contrast our findings with limited attention to when consumers have unlimited attention. Second we assume that consumers obtain a share of the surplus in the advertiser-consumer interaction and endogenize viewers' advertising digestion, ϕ.

5.1 Limited vs. unlimited viewer attention

Our base model features ad congestion. In this extension, we allow for some viewers with unlimited attention. For tractability, we confine ourselves to symmetric platforms. We first analyze the effect of entry in the two models with either limited or unlimited attention under symmetry (so that $a_i^* = a^*$ and $\lambda_i^* = \lambda^*$) and then compare the solutions for a given number of platforms. We show that markets with congested viewers behave markedly different from those with uncongested viewers: Platform entry leads to more advertising with congested viewers, while it leads to less advertising with uncongested viewers. As we then argue these findings continue for sufficiently asymmetric shares of congested and uncongested viewers.

Unlimited attention: the effect of entry. Profits without congestion are $\lambda_i a_i p(a_i) = \lambda_i R(a_i)$; see (5). Using symmetry, the first-order condition can be written as

$$\frac{R'(a^*)}{R(a^*)} = \frac{(1-\lambda^*)\tilde{\alpha}}{1-a^*}.$$

Equilibrium market share is

$$\lambda^* = \frac{[s(1-a^*)]^{\tilde{\alpha}}}{n[s(1-a^*)]^{\tilde{\alpha}} + v_0^{\tilde{\alpha}}}.$$

We have that $R'(a)/R(a) = (1-\varepsilon)/a$ where $\varepsilon = -ap'(a)/p$.

Under full coverage this becomes $\lambda^* = 1/n$ and the first-order condition simplifies to

$$\frac{1-\varepsilon}{a^*} = \frac{n-1}{n}\frac{\tilde{\alpha}}{1-a^*}. \tag{24}$$

This can be rewritten as

$$\frac{1-a^*}{a^*}(1-\varepsilon) = \tilde{\alpha}\frac{n-1}{n}. \tag{25}$$

The inverse price elasticity ε is upward sloping in a since $p(.)$ is assumed to be log-concave. This implies that the left-hand side is decreasing in a. The right-hand side is increasing in n. As a result the equilibrium ad level must be decreasing in n. The standard intuition of the competitive bottleneck model applies: after entry there is fiercer competition for viewers' time on a platform leading to less ad nuisance.

Under partial coverage we have

$$\frac{1-\varepsilon}{a^*} = \left(\frac{n-1}{n} + \frac{v_0^{\tilde{\alpha}}}{[s(1-a^*)]^{\tilde{\alpha}}}\right)\frac{\tilde{\alpha}}{1-a^*}. \tag{26}$$

This can be rewritten as

$$\frac{1-a^*}{a^*}(1-\varepsilon) - \frac{\tilde{\alpha}v_0^{\tilde{\alpha}}}{[s(1-a^*)]^{\tilde{\alpha}}} = \tilde{\alpha}\frac{n-1}{n}.$$

Compared to (25) the left-hand side has an additional term. This term is also decreasing in a. As a result the equilibrium ad level with partial coverage also must be decreasing in n.

We now turn to the model with congestion in which platforms maximize $\lambda_i \phi a_i p(a_i)/A$. Under symmetry the first-order condition (10) simplifies to

$$\frac{1}{n} = 1 - \varepsilon\eta \tag{27}$$

which uniquely determines a^* as a function of n. Since ε and η are upward-sloping in a (see Assumption 2 and Lemma 2), the right-hand side is decreasing in a. An increase in n therefore implies that ad level a^* is increasing in n. This result is an implication of Proposition 2 which covers symmetric platforms as a special case. It illustrates our finding that entry has the opposite effect in the model with congestion compared to the standard media model without congestion. Our finding tells us that with limited attention (i.e. with ad congestion) there is a trade-off between media diversity and media quality. Such a trade-off does not exist with unlimited attention.

Limited vs. unlimited attention: comparison of ad levels. Recalling that $\eta = \frac{1-a_i}{1-(1+\widetilde{\alpha})a_i}$, can write (27) as

$$\begin{aligned}\varepsilon &= \frac{n-1}{n}\frac{1-(1+\widetilde{\alpha})a^*}{1-a^*} \\ &= \frac{n-1}{n} - \frac{n-1}{n}\frac{a^*\tilde{\alpha}}{1-a^*}. \end{aligned} \tag{28}$$

Rewriting (24), the inverse price elasticity ε must satisfy without congestion and with full coverage

$$\varepsilon = 1 - \frac{n-1}{n}\frac{a^*\tilde{\alpha}}{1-a^*}. \tag{29}$$

We observe that the right-hand side of (29) takes larger values than the right-hand side of (28) for all admissible values for a and thus there is less advertising with advertising congestion than without.[19]

This may not seem obvious because with congestion attention ϕ is a common property resource and multiple platforms will exploit it excessively. Without ad congestion, any watched ad raises the attention of viewers. This allows the platform to extract the surplus of the marginal advertiser $p(a_i)$. By contrast, with ad congestion, the platform can only extract $(\phi/A)p(a_i)$. A higher ad level puts further downward pressure on the ad price (through ϕ/A), and the platform has an incentive to set a lower ad level with congestion.

The platform's profit per time unit is $a_i p(a_i) = R(a_i)$ without congestion and $a_i(\phi/A)p(a_i)$ with congestion. For given viewing time λ_i, the platform would maximize these expressions with respect to a_i. Without congestion the solution satisfies $ap'(a_i) + p(a_i) = 0$ which is equivalent to $\varepsilon = 1$; with congestion it satisfies

$$\frac{\phi}{A}\left[ap'(a_i) + p(a_i)\right] - \frac{\lambda_i \phi}{A^2} a_i p(a_i) = 0$$

which can be written as $\varepsilon = 1 - \lambda_i a_i / A$. Since ε is upward sloping this shows that ad levels are lower with congestion than without congestion if we treat viewer numbers as exogenous. With congestion, the platform takes into account that a higher ad level increases the degree of congestion. The associated drop in the ad price reduces the incentive to increase the ad level.

Platforms of course do not maximize profits for a given viewing time but take into account that viewers allocate their viewing time depending on the net quality of the platform. An ad-congested platform also takes into account that total ingestion A increases

[19] The right-hand side of (28), (29), and (30) is downward sloping and therefore in all specifications any solution a^* must be unique.

by less than λ_i as it marginally increases its ad level because its share λ_i decreases in the ad level, but this does not overturn the result for given viewing time.

Rewriting (24), the inverse price elasticity ε must satisfy without congestion and partial coverage

$$\varepsilon = 1 - \frac{n-1}{n} \frac{\frac{n}{n-1}(1-a^*) - a^*\tilde{\alpha}}{1-a^*} - \frac{v_0^{\tilde{\alpha}}}{[s(1-a^*)]^{\tilde{\alpha}}} \frac{a^*}{1-a^*}\tilde{\alpha}. \tag{30}$$

Since the right-hand side of (30) takes smaller values than (29) this is not clear with partial coverage. Here, the ad level with congestion is actually larger than without congestion if and only if

$$\frac{1}{n} < \frac{v_0^{\tilde{\alpha}}}{[s(1-a^*)]^{\tilde{\alpha}}} \frac{a^*}{1-a^*}\tilde{\alpha}$$

which is equivalent to

$$\frac{1}{n^2} < \frac{\lambda_0}{1-\lambda_0} \frac{a^*}{1-a^*}\tilde{\alpha}.$$

A mix of consumers with limited and unlimited attention. It is possible to extend the model to allow for a fraction $1-\kappa$ of viewers with unlimited attention. The profit function of media platform i is

$$\begin{aligned} \Pi_i &= \kappa \frac{\lambda_i \phi a_i p(a_i)}{A} + (1-\kappa)\lambda_i a_i p(a_i) \\ &= \lambda_i R(a_i) \left(1 - \kappa + \kappa \frac{\phi}{A}\right). \end{aligned}$$

Under full coverage and using symmetry, the first-order condition can be written as

$$\varepsilon = 1 - \frac{\kappa\phi}{n[(1-\kappa)a^* + \kappa\phi]} - \frac{n-1}{n} \frac{\tilde{\alpha} a^*}{1-a^*}. \tag{31}$$

In the special case $\kappa = 0$ we obtain (29) and in the special case $\kappa = 1$ we obtain (28). For a given number of platforms, the ad level is smaller for $\kappa \in (0,1)$ than when no viewer has limited attention ($\kappa = 0$). Regarding the comparative statics with respect to n we have to evaluate how the left-hand side varies with n for given a^*. By continuity, ad levels are increasing in the number of platform for κ sufficiently close to 1 and decreasing for κ sufficiently close to zero.[20]

[20] Here, we implicitly assume that there is a unique solution to the first-order condition and that this solution is an equilibrium.

5.2 Surplus sharing in the advertiser-consumer relationship and endogenous ad digestion

In the base model, consumers did not care about the composition of advertisers that come to their attention because advertisers extracted all the surplus generated from their interaction with consumers. Here we extend the analysis to have consumers obtain a positive fraction of the realized gains from trade. We show that consumer surplus can decrease with entry even under symmetry. We also engage this extension to endogenize viewer attention ϕ. Since viewer attention is equal to the number of digested ads, both variables affecting ad congestion will then be endogenous (ϕ and A).

Surplus sharing between advertisers and consumers. In the main part, we assumed that advertisers extract all gains from trade in the advertiser-consumer relationship, while now some fraction of this benefit goes to consumers. We will analyze situations with symmetric firms and a symmetric equilibrium, and determine equilibrium properties with special attention to the effects of entry. We first look at the simplest case, when the ad demand is perfectly elastic. In that case, we show that entry has no effect on equilibrium ad levels for platforms: including ad benefits to consumers has no impact compared to the model with no such benefits. This observation leads us to introduce a simple but rich extension, namely to have two advertiser types, one with high willingness to pay for ad exposure (and so high benefit to consumers too) and one type with low willingness to pay. With enough platforms, equilibrium has the marginal advertiser on the lower step of the inverse advertiser demand curve. Then entry raises ad levels, due to the common property resource feature of the main model. This effect implies that entry reduces consumer benefits from ads because there are more worse ads in the mix. Consumers get less enjoyment from each platform as platform effective quality is reduced both due to more ads (and less content) and less direct benefit from the ads due to the composition effect. Nonetheless, consumers still benefit directly from having more program variety. We show that the net effect can be that entry reduces consumer welfare even in the symmetric case (in contrast to the main model). Thus the introduction of consumer ad benefits into the model can cause consumers to suffer from more variety even absent platform quality asymmetries.

To introduce viewer ad benefits, we continue to denote the per-viewer gains from trade that go to any advertiser a by $p(a)$ and we assume now that the consumer obtains $\beta p(a)$

(where before we had $\beta = 0$).[21]

We write the utility maximization problem as

$$\max_{\lambda_0,\lambda_1,\ldots,\lambda_n} \sum_{i=1}^{n}[s_i(1-a_i)\lambda_i]^\alpha + (\lambda_0 v_0)^\alpha + \beta \frac{\phi \sum_{i=1}^{n} \lambda_i \int_0^{a_i} p(a)da}{\sum_{i=1}^{n} \lambda_i a_i}$$

subject to $\Sigma_{i=0}^{n}\lambda_i = 1$.[22] Here we add the surplus benefits from product consumption to the time-use utility. Notice that this structure gives consumers an additional incentive to spend time on platforms with more ads, and therefore for platforms to carry more ads, increasing congestion. Moreover, the resulting overfishing will be exacerbated the more platforms are present, for the individual platform internalizes less of the congestion cost. Thus, one may suspect that platform entry might decrease consumer surplus.

We first (briefly) deploy the analysis with a covered market ($v_0 = 0$) and a single advertiser type such that $p(a) = p$. Recall from Remark 1 that absent ad benefits and with flat inverse advertiser demand, platforms each set the maximal level of ads, $a = 1-\alpha$. The calculus of ad setting does not change with (constant) ad benefits because the consumer utility is $n^{1-\alpha}s^\alpha(1-a)^\alpha$ plus ad benefits $\beta\phi p$, which are constant for flat ad demand. So consumer surplus simply rises with n.

So consider the two-step advertiser demand. Let there be a_H advertisers with willingness to pay p_H for a contact, and then an infinitely elastic demand from advertisers willing to pay $p_L < p_H$. At a symmetric equilibrium with $a(n) > a_H$ ads per platform,

[21] One way to micro-found β is as follows. If there is a linear conditional product demand per consumer $(a - t_a)$ for an advertiser of type a charging product price t_a and zero marginal production cost, a's profit maximizing (uniform) product price would be $t_a = a/2$ and its profit $p(a) = a^2/4$. Consumer surplus in the advertiser-consumer relationship is $p(a)/2 = a^2/8$, corresponding to $\beta = 1/2$. Other values of β (below 1) can be generated in like manner from the class of ρ-linear demand functions.

Another example highlights price discrimination based on the consumers' disclosure decision of their personal characteristics. Consumers have unit demand and draw their willingness to pay after having decided how to allocate their time and being exposed to advertising. Advertisers operating as sellers have zero marginal costs of production. With a uniform distribution of consumer willingness to pay, consumers draw their valuation from the $[0,a]$-interval for advertiser a if they digested that ad. In the consumer-optimal perfect Bayesian equilibrium of the game in which consumers truthfully communicate some information about their willingness to pay and then the seller sets the price, the consumer communicates the interval $(a(1/2)^k, a(1/2)^{k-1}]$ her draw belongs to (see Nageeb, Lewis, and Vasserman, 2019). All potential gains from trade materialize and are equal to $a^2/2$. Consumer surplus is $a^2/6$ and thus $\beta = 1/2$.

[22] The consumer internalizes the expected consumption net benefit from being exposed to ads when deciding how to allocate her time. As in the main model, with the price of the outside consumption good normalized to one, utilities are expressed in units of this outside consumption good. Income is assumed large enough that the consumer can afford to buy all advertised products.

consumer surplus is (under symmetry with common per platform viewership λ):

$$\begin{aligned} CS(n) &= n^{1-\alpha}s^{\alpha}\left(1-a\left(n\right)\right)^{\alpha}+\phi\beta\frac{n\lambda\left[p_{H}a_{H}+p_{L}\left(a\left(n\right)-a_{H}\right)\right]}{n\lambda a\left(n\right)} \\ &= n^{1-\alpha}s^{\alpha}\left(1-a\left(n\right)\right)^{\alpha}+\phi\beta p_{L}+\phi\beta\frac{\left(p_{H}-p_{L}\right)a_{H}}{a\left(n\right)}. \end{aligned} \tag{32}$$

To find equilibrium ad levels we must first characterize consumer choice in the face of a deviation from symmetric ad choices by rivals. If the market is covered ($\sum_{i>0}\lambda_i = 1$) and all but platform i set a common ad level $a > a_H$, the consumer problem is (for $a_i > a_H$)

$$\begin{aligned} &\max_{\{\lambda_i,\lambda\}}\left[s\left(1-a_i\right)\lambda_i\right]^{\alpha}+\left(n-1\right)\left[s\left(1-a\right)\lambda\right]^{\alpha}+\phi\beta\frac{p_{H}a_{H}+\left(n-1\right)p_{L}\left(a-a_{H}\right)\lambda+p_{L}\left(a_i-a_{H}\right)\lambda_i}{\lambda_i a_i+\left(n-1\right)\lambda a} \\ &\text{s.t. } \lambda_i+\left(n-1\right)\lambda=1 \end{aligned}$$

Solving and differentiating i's viewing time at the symmetric times, $\lambda = 1/n$ yields

$$\left.\frac{d\lambda_i}{da_i}\right|_{\lambda=\frac{1}{n}} = -\left(n-1\right)\frac{a^2\alpha^2\left(1-a\right)^{\alpha-1}n^{1-\alpha}+\phi\beta s^{-\alpha}a_H\left(p_H-p_L\right)}{a^2\alpha\left(1-\alpha\right)\left(1-a\right)^{\alpha}n^{3-\alpha}} < 0.$$

Armed with this leakage effect, we can solve the deviant platform's problem and hence the equilibrium. Platform i solves (for $a_i > a_H$)

$$\max_{a_i}\pi_i = \phi p_L\frac{\lambda_i a_i}{\sum_j \lambda_j a_j},$$

which yields the first-order condition under symmetry as $n\lambda a\left(\lambda + a\frac{d\lambda_i}{da_i}\right) - \lambda^2 a = 0$. Substituting and simplifying delivers the implicit form of the symmetric equilibrium as

$$f\left(a\right) \equiv a\left(1-\alpha-a\right)\alpha\left(1-a\right)^{\alpha-1}n^{1-\alpha} = \frac{\phi\beta}{s^{\alpha}}a_H\left(p_H-p_L\right) \tag{33}$$

where the RHS is a constant. Notice that for $p_H = p_L$ there is a single advertiser type and the solutions are $a = 0$ and the equilibrium one previously identified, $a = (1-\alpha)$. Note that $f(0) = f(1-\alpha) = 0$, and $f(a)$ is quasi-concave (as argued below), so that there are two solutions (as long as the RHS is not too large); the larger one is readily selected as the equilibrium one for it corresponds to the maximal profit for each individual platform's best reply. Therefore the equilibrium ad level is at or past the peak of $f(a)$ and is below the equilibrium level $a = 1-\alpha$ of the single-advertiser case. This is because platforms recognize that airing more ads brings in more low advertiser types which are less attractive to viewers and demeans the viewer surplus from ads, so platforms rein back ads. Another feature to note is that this latter effect is diminished when more platforms are present: in tune with the main model, the common property resource effect gets more acute. To see this effect, just note that $f(a)$ is increasing in n and recall the equilibrium is at the larger solution for a.

Returning to the shape of $f(a)$, write

$$\begin{aligned} f'(a) &= \alpha (1-a)^{\alpha-2} n^{1-\alpha} g(a) \\ \text{where } g(a) &\equiv a^2 (1+\alpha) + a\left(\alpha^2 - \alpha - 2\right) + 1 - \alpha \end{aligned} \tag{34}$$

so the quadratic convex form of $g(a)$ implies that $f(a)$ is quasiconcave: there are two roots to $g(a)$, but the larger one is above $1-\alpha$ and so out of bounds.[23]

What does this imply for consumer surplus? The main take-away of this extension is that the consumer surplus effect of entry can be reversed.

Proposition 6 *For a symmetric covered market with ad benefits to consumers and two advertiser types, there are circumstances under which consumer surplus decreases with entry.*

Proof. Differentiate $CS(n)$ in (32) and substitute the equilibrium condition (33) to give

$$\frac{dCS(n)}{dn} = \frac{(1-\alpha) n^{-\alpha} s^{\alpha} (1-a^{eq})^{\alpha}}{g(a^{eq})} \left(g(a^{eq}) + a^{eq}\alpha (1-\alpha-a^{eq}) + \alpha (1-\alpha-a^{eq})^2 \right)$$

where we here understand a^{eq} to be the equilibrium ad level from (33). Recalling that $g(a^{eq}) < 0$, then $\frac{dCS(n)}{dn} < 0$ if and only if $a^{eq}\alpha(1-\alpha-a^{eq}) + \alpha(1-\alpha-a^{eq})^2 > -g(a^{eq})$. The LHS of this expression is positive on $a^{eq} \in (0, 1-\alpha)$. Recall that the function $f(a)$ is maximized where $g(a) = 0$.[24] Therefore, if $\frac{\phi\beta}{s^{\alpha}} a_H (p_H - p_L)$ (which equals $f(a^{eq})$: see (33)) is close to the maximal $f(a)$ then $g(a^{eq})$ is close to zero. In such cases then necessarily *consumer surplus decreases with entry.* ■

The variety benefit is overwhelmed by the lost surplus from ads stemming from exacerbating the common property problem and inducing too many low-surplus advertisers into the mix. Advertiser surplus is decreasing with entry as in the main model without

[23] We thus see that $f(a)$ is maximized on $[0, 1-\alpha]$ at

$$a^* = \frac{2 + \alpha - \alpha^2 - \left(\alpha(1+\alpha)\left(\alpha^2 - 3\alpha + 4\right)\right)^{1/2}}{2(1+\alpha)},$$

where we note that $\alpha^2 - 3\alpha + 4 > 0$ for $\alpha \in [0, 1]$.

[24] It can be shown that the maximum is attained for $a > (1-\alpha)/2$ though simulations show it is not much larger.

consumer ad benefits. The reason is the same: entry leads to higher equilibrium advertising levels. This implies that fewer high-value ads are digested by consumers and this decreases advertiser surplus. To summarize, if consumers obtain a fraction of the realized gains from trade in the advertiser-consumer relationship entry may reduce both consumer *and* advertiser surplus.

Endogenous ad digestion. In the main model, consumers had no incentive to adjust their ad digestion to the prevailing market structure since all realized gains from trade in the consumer-advertiser relationship accrued to advertisers. Consider now the possibility that consumers adjust their attention to equate marginal benefit from increased attention to marginal costs.[25] We see this as a long-term behavioral pattern that responds to costs and benefits (as in Anderson and de Palma, 2009). The cost of attention is denoted by $C(\phi)$ which is strictly convex with the standard boundary properties. Then there is a unique solution to the utility maximization problem with respect to ϕ for any given a^e. This solution $\phi^*(a^e)$ is implicitly defined by the average viewer benefit per ad equal to the marginal cost of increasing attention,[26]

$$\frac{\beta}{a^e}\int_0^{a^e} p(a)da = C'(\phi). \tag{35}$$

Because average ad benefits decrease in a^e, the solution ϕ^* is decreasing in a^e. Since a^e is increasing with entry, viewers respond to the expected increase in ad ingestion due to entry by digesting even fewer ads. The reason is that advertisers come from a worse selection which makes paying attention to ads less attractive. Thus, with entry, an increase in ingestion A is accompanied by a reduction of digestion ϕ. Hence, also when endogenizing ϕ, ad congestion $(A-\phi)/A$ is increasing in the number of platforms.

5.3 Subscription pricing

We here introduce subscription pricing into the model. We have in mind that a subscription enables the viewer to access the product, like buying a magazine or subscribing to a pay-TV channel, a web-site or video service. We exclude (for simplicity) that time usage can be metered and charged, so our subscriptions are all-or-nothing.

[25] In spirit, this relates to work on the use of ad blockers (Anderson and Gans, 2011; Johnson, 2013). In those papers, consumers decide whether or not to block ads; i.e. whether to have unlimited attention to ads or zero attention. In our model, consumers choose how many ads to digest.

[26] To be precise, there is a threshold $\hat{a}$ such that for $a^e < \hat{a}$ we must have $\phi(a^e) = a^e$ (because the solution to eq. (35) has the property that $\phi > a^e$) and $\phi(a^e) < a^e$ for $a^e > \hat{a}$. We consider environments in which the resulting a^e satisfies $a^e > \hat{a}$.

We denote the subscription price for platform i by σ_i and we modify the viewer utility to read

$$V = \left[\sum_{i \in I} [s_i(1-a_i)\lambda_i^I]^\alpha + (\lambda_0^I v_0)^\alpha\right] - \sum_{i \in I} \sigma_i \tag{36}$$

when the viewer chooses to pay the subscription price for a subset $I \subseteq N = \{1, ..., n\}$ from the set of all possible offerings: the superscript I denotes the optimal time choices corresponding to this set.[27] Denote the term in the square bracket $CS(\mathbf{a}; I)$. Platform payoffs are zero if not selected, and otherwise ad revenues given time use plus revenue from subscription.

There are several different ways to configure the strategic game. We suppose that the platform first sets subscription prices and then ad levels. Consumers may or may not observe ad levels when making their subscription decisions. This leads to two alternative games which we will analyze.

5.3.1 Subscription decisions prior to observing ad levels

The timing of the game is as follows. First the n platforms simultaneously choose their subscription prices. Second, the viewer observes these subscription prices and makes her subscription decisions. Third, the platforms simultaneously choose their ad levels. Finally, the viewer chooses her time allocation across platforms. In any such subgame, the analysis of the main model applies: time use is allocated to maximize utility across the programs accessed by paying their subscriptions.

If the viewer chooses not to subscribe to a platform, the platform would do better by reducing its subscription price so that it would be included. Note that each platform has a positive value when added to a portfolio, by dint of the positive extra surplus it contributes (for any ad level below 1), so it can always charge some positive subscription price and be chosen. Given this property, *all platforms must be active in equilibrium*.

Along the equilibrium path, each platform sets its subscription price at the *incremental value* it adds to the viewer and so the price for platform j is

$$\sigma_j = CS(\mathbf{a}^*(N); N) - CS(\mathbf{a}^*(N \backslash j); N \backslash j).$$

Thus subscription prices are set so as to just keep viewers on board each platform, while a reduction in own subscription price has no impact on other platforms (and an increase

[27] We implicitly assume that the viewer has a sufficient income to afford subscriptions and product purchases. Regarding the time budget, we must have $\Sigma_{i=0}^n \lambda_i = 1$ with $\lambda_i = 0$ for all $i \notin I \cup \{0\}$.

is suicidal).[28]

Under symmetry, in an abuse of notation, we write $CS(n)$ instead of $CS\left(\mathbf{a}^{*}(N);N\right)$. The incremental value is $CS(n)-CS(n-1)$ and so the viewer spends $n(CS(n)-CS(n-1))$ for subscriptions. In other words, the subscription price is pinned down by the loss from dropping one program. The viewer does not have an incentive to drop more than one program if CS in concave in n, which will be assumed in the sequel.[29]

Consider now entry of an additional platform. The incremental subscription price after entry is $CS(n+1)-CS(n)$ and viewers spend $(n+1)(CS(n+1)-CS(n))$. Because of the concavity of CS the incremental price is decreasing with entry and thus a viewer pays less for each individual subscription. However, with entry a viewer buys an additional subscription. Total spending on subscriptions is larger with entry if and only if $(n+1)(CS(n+1)-CS(n))>n(CS(n)-CS(n-1))$, which is equivalent to $n[CS(n+1)-2CS(n)+CS(n-1)]+CS(n+1)-CS(n)>0$. Concavity of CS in n implies that the term in square brackets is negative, while the fact that CS is increasing in n implies that $CS(n+1)-CS(n)$ is positive.

For example, suppose that there is a single advertiser type with $p(a)=p$ and that the market is covered ($v_0=0$). Again recall from Remark 1 that with flat inverse advertiser demand, platforms each set the maximal level of ads, $a=1-\alpha$. Consumer surplus is $n^{1-\alpha}s^{\alpha}\left(1-a\right)^{\alpha}$ which is concave in n. Thus, $CS(n)$ is a constant times $n^{1-\alpha}$. In this case, total spending on subscriptions goes up with entry if and only if $n\left((n+1)^{1-\alpha}-2n^{1-\alpha}+(n-1)^{1-\alpha}\right)+(n+1)^{1-\alpha}-n^{1-\alpha}>0$. Even in this special case it depends on the value of α whether total spending on subscriptions goes up or down with entry. For n large it always increases, for the subscription price does not fall much with entry. For high α and low n it decreases. This is because $CS\left(n\right)$ is highly concave at first for low n and so subscription prices drop precipitously.

In the main model we showed that viewers are better off with entry, i.e., $CS(n+1)>CS(n)$ because the variety effect dominates the quality effect. When the viewer pays less for subscriptions in total, entry is a fortiori beneficial. By contrast, a higher subscription

[28] The term Incremental Value Pricing was coined by Anderson, Foros, and Kind (2018) to describe platform pricing to advertisers when some viewers multi-home so platforms can only charge the value of contacting their exclusive consumers.

[29] If CS were convex in n, the best consumer deviation under incremental value pricing is to drop all programs. The n platforms jointly offer the benefit $CS(n)-CS(0)$. In a symmetric pricing equilibrium, each firm sets a subscription fee of $[CS(n)-CS(0)]/n$ and the full gross surplus generated through the media platforms would be extracted through fees independent of the number of platforms in the industry. In this case, net consumer surplus would be neutral to entry.

bill counteracts the gross benefits $CS(n+1) - CS(n)$ of entry. Consumers continue to be better off with entry if

$$CS(n+1) - (n+1)(CS(n+1) - CS(n)) > CS(n) - n(CS(n) - CS(n-1))$$

which is equivalent to $2CS(n) > CS(n+1) + CS(n-1)$. The next result follows immediately.

Proposition 7 *For a symmetric market with consumers choosing subscriptions before observing ads, consumer surplus increases with entry if and only if $CS(n)$ is strictly concave, where $CS(n)$ is the (gross) consumer surplus from the main model, evaluating at the equilibrium advertising levels there.*

Thus, Proposition 4 continues to hold if the equilibrium viewer benefit gross of payments for subscriptions, $CS(n)$, is strictly concave in n. This is the condition for each platform to price its subscription at its incremental value, where it internalizes that its absence from the consumer portfolio will induce the equilibrium ad levels corresponding to one fewer platform. As noted in the preceding footnote, if instead $CS(n)$ is convex, the symmetric subscription price equilibrium involves each platform pricing at the average gross consumer surplus from the full palette of platforms. In this case, consumer surplus is fully extracted and so entry does not affect viewer welfare.

Returning to the concave case, for which subscriptions are priced at incremental values, entry is pro-competitive in the sense that it reduces each platform's subscription price, along with improving gross surplus directly. If total subscription payments also drop with entry, introducing subscription prices renders higher benefits from entry compared to the main model where there are no subscription prices, and there is just the gross surplus effect (which is just the same as in the main model). Conversely though, if total subscription payments rise with entry, benefits from entry are lower than for the main model because they come with higher subscription payments. Both cases can arise for flat advertiser demand, for which $CS(n)$ is indeed concave. Notice that the result of Proposition 4 shows that $CS(n)$ is increasing, but does not depend upon its concavity.

5.3.2 Subscription decisions after observing ad levels

The timing of the game is now as follows. First the n platforms simultaneously choose their subscription prices. These are observed, then the platforms simultaneously choose their ad levels. Finally, viewers choose to which platforms they subscribe, and their time allocation across them. The first point to note in general is that the ad levels for the second stage game are those of the main model for whatever selection of subscriptions the viewer makes. That is, given the (binary) selection decisions, the platforms just choose ad levels along the equilibrium path as before. This though belies the key role of the subscription price choices. A platform with too large a subscription price is vulnerable to being pried from the market by a platform which can topple it by adjusting its ad level so as to render subscribing unattractive and benefitting from fewer active competitors in the advertising market. This vulnerability disciplines the subscription prices according to a novel *topple-free condition*. The insight here is that subscription prices will be set below incremental values. They are tied down by being as large as possible subject to being topple-free.[30]

To illustrate, assume a covered market, symmetric platforms, and perfectly elastic (flat) advertiser demand. For this case we have from the main model without subscription prices $a = 1 - \alpha$ as the equilibrium ad level, independent of n, along the equilibrium path (see Remark 1). This is the advertising sub-game equilibrium for *any* vector of σ's (that induce at least one subscription chosen) as long as no platform prefers to topple another. We now turn to the topple-free condition.

We wish to find the equilibrium subscription prices; doing so entails finding the deviation ad level, $\hat{a}$, by any platform that just renders the viewer indifferent between dropping a platform and keeping it. For a platform to be indifferent between eliminating a rival and not doing so, it must earn the same amount with all rivals subscribed and choosing $a = 1 - \alpha$ and setting $\hat{a}$ with one less rival present. If the incremental value of another platform goes down so much that it becomes priced (at σ) above its incremental value then it is dropped by the viewer. And if the deviating platform's profit rises when another is eliminated, the deviation is profitable. We leverage this argument to find equilibrium σ's such that no such deviation is possible.

We first determine what the viewer does when one platform chooses some $\hat{a}$, and then find profits to the deviant when it succeeds in ejecting a platform (so it now faces $n - 2$ competitors). So suppose that a platform sets some ad level $\hat{a}$ and consider the viewer problem, with n platforms in total (the viewer having activated her subscription to all).

[30] Moreover, along the equilibrium path, the toppling strategy will always involve a drop in a platform's ads, for even if a platform can drive another's incremental value below its subscription price by raising its ad level, it will drive its own incremental value below its subscription price and get itself unsubscribed. This feature is borne out in the example that follows.

Her time-use utility, extracting the common s terms,[31] is

$$U = (1 - \hat{a})^{\alpha} \hat{\lambda}^{\alpha} + (n-1)(1-a)^{\alpha} \lambda^{\alpha}, \tag{37}$$

and with a covered market, $\hat{\lambda} + (n-1)\lambda = 1$, which we can substitute to get

$$U = (1 - \hat{a})^{\alpha} (1 - (n-1)\lambda)^{\alpha} + (n-1)(1-a)^{\alpha} \lambda^{\alpha}.$$

The viewer's optimal λ entails the first-order condition

$$-\alpha (n-1)(1-\hat{a})^{\alpha} (1-(n-1)\lambda)^{\alpha-1} + \alpha (n-1)(1-a)^{\alpha} \lambda^{\alpha-1} = 0,$$

which rearranges to

$$\omega \equiv \left(\frac{(1-a)}{(1-\hat{a})} \right)^{\frac{\alpha}{1-\alpha}} = \frac{\lambda}{(1-(n-1)\lambda)}.$$

Hence $(1-(n-1)\lambda)\omega = \lambda$ and we obtain

$$\frac{\omega}{1+\omega(n-1)} = \lambda \text{ and } \hat{\lambda} = \frac{1}{1+\omega(n-1)}.$$

We now determine platform profits. First, recall profits are $\pi_i = \phi p_L \frac{\lambda_i a_i}{\sum_j \lambda_j a_j} + \sigma_i$, and we can normalize $\phi p_L = 1$ for it makes no difference to the incremental calculus. When all ads are the same (symmetric equilibrium), the profit per platform is $\frac{1}{n} + \sigma$. We now check what happens to deviant i's profit if it eliminates a rival by successfully moving $\hat{a}$ in the direction that reduces rivals' incremental values to the consumer below their subscription fees, and inducing the consumer to not subscribe to one.[32] It will get

$$\pi_i^D = \frac{\hat{\lambda}_O \hat{a}}{\hat{\lambda}_O \hat{a} + (n-2)\lambda_O a} + \sigma_i,$$

where the O signifies that a rival is Out. Hence

$$\frac{\omega}{1+\omega(n-2)} = \lambda_O \text{ and } \hat{\lambda}_O = \frac{1}{1+\omega(n-2)}.$$

Substituting,

$$\pi_i^D = \frac{\hat{a}}{\hat{a} + (n-2)\omega a} + \sigma_i,$$

[31] That is, we can effectively normalize $s = 1$ for transparency: all equilibrium σ's would be just scaled appropriately.

[32] We show below that it will decrease its ads. Note there is a free-rider problem to levering out a rival insofar as other platforms benefit.

and so this should equal profit along the equilibrium path, namely $\pi_i = \frac{1}{n} + \sigma_i$.

Therefore we determine $\hat{a}$ by

$$\frac{\hat{a}}{\hat{a} + (n-2)\,\omega a} = \frac{1}{n} \text{ or } \hat{a} = \frac{n-2}{n-1}\omega a.$$

With $\hat{a}$ thus determined, we can determine σ as the extra profit that can be protected against a deviation to $\hat{a}$. That is, platform i will deviate to $\hat{a}$ if doing so will cause the viewer to drop one of the other products (and we need to ensure that the viewer does not want to drop i). So we need to return to the viewer benefit expression (37), and see when she is just indifferent to dropping some $j \neq i$. Her benefit with all products is:[33]

$$U_I = (1-\hat{a})^\alpha \left(1-(n-1)\,\lambda_I\right)^\alpha + (n-1)\,(1-a)^\alpha\,\lambda_I^\alpha - n\sigma,$$

with $\frac{\omega}{1+\omega(n-1)} = \lambda_I$; dropping a (non-deviant) platform yields benefit

$$U_O = (1-\hat{a})^\alpha \left(1-(n-2)\,\lambda_O\right)^\alpha + (n-2)\,(1-a)^\alpha\,\lambda_O^\alpha - (n-1)\,\sigma$$

with $\frac{\omega}{1+\omega(n-2)} = \lambda_O$ and $\hat{\lambda}_O = \frac{1}{1+\omega(n-2)}$.[34] The equilibrium σ equalizes these two ($U_O = U_I$) so it is given by

$$\begin{aligned}\sigma &= (1-\hat{a})^\alpha \left[\left(\frac{1}{1+\omega\,(n-1)}\right)^\alpha - \left(\frac{1}{1+\omega\,(n-2)}\right)^\alpha\right] \\ &\quad + (1-a)^\alpha \left[(n-1)\left(\frac{\omega}{1+\omega\,(n-1)}\right)^\alpha - (n-2)\left(\frac{\omega}{1+\omega\,(n-2)}\right)^\alpha\right].\end{aligned}$$

We also need to check that the viewer does not want to ditch the deviant. Because the optimal U decreases with $\hat{a}$, which is true by the envelope theorem, the deviant will not be ditched if the deviant has a lower $\hat{a} < a$ $(= 1-\alpha)$.[35]

[33] Again setting the common $s = 1$ for this, but reinserting will just scale the answer accordingly.

[34] Inserting the λ values gives

$$\begin{aligned}U_I &= (1-\hat{a})^\alpha \left(1-(n-1)\frac{\omega}{1+\omega\,(n-1)}\right)^\alpha + (n-1)\,(1-a)^\alpha \left(\frac{\omega}{1+\omega\,(n-1)}\right)^\alpha - n\sigma, \text{ and} \\ U_O &= (1-\hat{a})^\alpha \left(1-(n-2)\frac{\omega}{1+\omega\,(n-2)}\right)^\alpha + (n-2)\,(1-a)^\alpha \left(\frac{\omega}{1+\omega\,(n-2)}\right)^\alpha - (n-1)\,\sigma.\end{aligned}$$

[35] Recall that $\hat{a} = \frac{n-2}{n-1}\omega a$, and that $\omega \equiv \left(\frac{(1-a)}{(1-\hat{a})}\right)^{\frac{\alpha}{1-\alpha}}$. Hence $\hat{a}\,(1-\hat{a})^{\frac{\alpha}{1-\alpha}} = \frac{n-2}{n-1} a\,(1-a)^{\frac{\alpha}{1-\alpha}}$.
The LHS of this equality is quasi-concave on $\hat{a} \in (0,1)$ and zero at each endpoint. Moreover, it peaks at $\hat{a} = 1-\alpha = a$, so one root is always above a and the other is always below (and note that there are always two solutions in $(0,1)$ because the RHS is positive and below the peak of the LHS, for the factor $\frac{n-2}{n-1} < 1$). Recall we select the lower root, or else the viewer ditches the deviant. So necessarily the pertinent $\hat{a} < a$.

With the model (implicitly) solved, we next find how viewer benefits vary with n given the equilibrium σ derived above. Given that $CS = n\left[s(1-a)\lambda\right]^{\alpha} - n\sigma$, and setting $s = 1$ and recalling that $\lambda = 1/n$ and that $1 - a = \alpha$, we have

$$CS(n) = \alpha^{\alpha} n^{1-\alpha} - n\sigma(n).$$

The behavior of this function is readily determined by inserting values — it is a little intricate for we must find the lower root for $\hat{a}$, and note that this depends on n. First of all, $CS(n)$ is increasing, so that more varieties benefit viewers, despite them being priced. Second, $\sigma(n)$ is decreasing, so that more competition does indeed diminish subscription prices through the intricate mechanism of the topple-free condition. Third though, $n\sigma(n)$ increases in n so that viewers spend more on subscriptions when there are more varieties. Nonetheless, in conjunction with the first result, the spending effect is dominated by the variety effect. Thus, the result in Proposition 4 continues to hold.[36]

In terms of the take-away, these results are qualitatively similar to those of Proposition 7. While subscription prices fall with entry, they entail more spending, so the benefits of entry are smaller than when there are no subscription prices. However, with the timing of events here (with consumers observing ad levels before subscribing) the subscription prices are lower (below incremental values).

6 Conclusion

Even though consumers dislike program content to be padded with advertising and even though some advertisers fail to sell because of ad clutter, we observe huge amounts of

[36] The reader may wonder how a game in which platforms choose both ad levels and subscription prices before viewers make their choices would play out. Any set of ad levels and subscription prices induces a set of subscriptions and time-use by the consumer. If the viewer chose not to subscribe to a platform, it would do better by reducing its subscription price so that the viewer would include it. Each platform i brings a positive incremental surplus to any subscription portfolio when choosing $a_i < 1$, and so it can always set a positive subscription fee and be chosen. Therefore any equilibrium entails that all platforms must be active. Indeed, for any vector of platform ad levels, each platform will price at its incremental value. The price for platform j given any vector of ad levels, $\mathbf{a}$, is thus $\sigma_j = CS(\mathbf{a}; N) - CS(\mathbf{a}; N/j)$. Subscription prices are set so as to just keep viewers on board each platform: a higher price evicts the platform and a lower one changes nothing for other platforms. However, changing own ad level *does* impact rivals' incremental values, via the differential CS terms. Therefore, any platform j can change its ad level by a small amount (in whatever direction is effective) and cause the consumer to eject some other platform or platforms. Such a second-order deviation from any candidate equilibrium $\mathbf{a}$ will cause a first-order jump up in j's profit. Consequently, no (pure strategy) equilibrium can exist. We eschew mixed strategy equilibria here and conclude that this game structure is bankrupt.

advertising in TV and other mass media. If neither advertisers nor consumers obtain a service they like, this begs the question why media platforms do not simply reduce the volume of ads and make everybody happier. The answer comes from thinking of viewer attention as a common property resource.

In this paper we propose a time-use model of media consumption and show that limited attention for advertising can explain a number of features that standard theory cannot, and delivers several novel results.

First, higher-quality platforms attract more consumer time and place *less* advertising. Lower-value advertisers post ads on lower-quality platforms only, whereas higher-value advertisers advertise more broadly.

Second, an increase in the variety of opinion (platform entry) causes more advertising on each platform, and thus reduces net content quality. In the presence of advertising clutter, the matching of advertisers to consumers becomes important – which ads get through? Matching is efficient if the advertisers with the highest willingness to pay get their messages to consumers. Advertising efficiency is diminished when higher-value advertisers are replaced by lower-value advertisers, and this happens when there are more media platforms vying for attention in the presence of clutter.

Third, under free entry, increasing the quality of some incumbent platforms reduces media diversity when the quality of the marginal platform does not change. However, this increases consumer surplus and advertising efficiency. Thus, consumer surplus and total surplus increase when media diversity is reduced. However, if society values variety of opinion more strongly than do consumers, society may well be better off under more diversity, despite consumers being worse off and advertising efficiency decreasing.

Fourth, lower entry costs result (as expected) in more diversity of opinion. As a benchmark with symmetric media platforms, entry is good for consumers even though more content is replaced by advertising — this continues to hold when platforms charge for subscriptions, but consumers may end up paying more in total for subscriptions. However, with asymmetric media platforms the negative indirect effect can dominate the direct benefit from more variety and consumers may be worse off when entry costs go down. With a covered market, then total surplus also goes down because advertising efficiency always decreases without a corresponding increase in market base. In settings in which consumers obtain a fraction of the surplus from the advertiser-consumer interaction, consumers may be worse off even under platform symmetry when entry costs go down.

Appendix.

Proof of Proposition 5. The proof is by example: we reverse engineer the result.

First, simplify by setting $\tilde{\alpha} = 1$ (i.e., $\alpha = 1/2$) and set $\phi = 1$.[37] Next, choose a pair of advertising levels with $a_L > a_H$ for the post-entry situation. These advertising levels are both below $1/2$ because with $\tilde{\alpha} = 1$, equilibrium actions (the ψ_i) cannot support higher ad levels.

We next use (17) to find the corresponding quality ratio that supports the specified advertising levels, and then use the free-entry condition for the low-quality platforms to find the value for K that supports zero profit at the chosen a_L. In (23), n_H is a parameter: we can choose its value as the number of high-quality platforms that would freely enter under some higher level of entry cost, and then we can suppose that the market has just those firms active initially. When the entry cost drops to K, new (low-quality) platforms come in, and we show that consumer surplus can go down. Notice that the quality degradation (between high and low qualities) needs to be severe enough to offset the earlier finding (for symmetry) that entry benefits consumers. We can set a pre-entry level of a for the high quality platforms alone (which is below a_H because we know ad levels rise with entry) and support that level of a with an initial level of entry cost.

For the example, we first determine the level of K which will support $a_L = 1/3 > a_H = 1/6$. From the ZPSEE for the low types (16), we must have ε and K combinations that deliver $a_L = 1/3$, so they must satisfy

$$\begin{aligned} K &= \left(1 - \varepsilon \frac{2/3}{1 - 2\left(\frac{1}{3}\right)}\right)\left(\frac{1}{3}\right)^{-\varepsilon} \\ &= (1 - 2\varepsilon)\, 3^{\varepsilon}. \end{aligned}$$

In particular, we can take $\varepsilon = 1/3$ to find $K = 3^{-\frac{2}{3}} = 0.48$. Next, we need to find the quality ratio that delivers $a_H = 1/6$: from (17) we have

$$\frac{1/6}{1/3}\frac{5/6}{2/3}\frac{s_H}{s_L} = \frac{1 - \varepsilon\frac{5/6}{2/3}}{1 - 2\varepsilon}$$

or

$$\frac{s_H}{s_L} = \frac{8}{5}\frac{1 - \varepsilon\frac{5}{4}}{1 - 2\varepsilon}.$$

[37] Or else ϕ can be folded into the entry cost.

Using the definition of the Ω's from (21) above, we can write them as

$$\Omega_L = \frac{1-\varepsilon\frac{2/3}{1-2/3}}{1/3} = 3\left(1-2\varepsilon\right);$$
$$\Omega_H = \frac{1-\varepsilon\frac{5/6}{1-2/6}}{1/6} = 6\left(1-\varepsilon\frac{5}{4}\right).$$

Inserting these values into the consumer surplus expression when both types are present, (23), we get

$$\begin{aligned} CS_B &= \left(n_H\Omega_H + \frac{1-n_H\Omega_H a_H}{a_L}\right)\frac{s_L(2/3)}{\Omega_L} \\ &= \left(\frac{n_H\Omega_H}{2}+3\right)\frac{2s_L}{3\Omega_L} \\ &= \left(3\left(1-\varepsilon\frac{5}{4}\right)n_H+3\right)\frac{2s_L}{9\left(1-2\varepsilon\right)} \\ &= \left(\left(1-\varepsilon\frac{5}{4}\right)n_H+1\right)\frac{2s_L}{3\left(1-2\varepsilon\right)}. \end{aligned}$$

Now, we know too that consumer surplus before the wave of entry induced by the reduction in entry cost to K is from (19)

$$CS_H = n_H s_H\left(1-a\right)$$

and we know that n_H satisfies $\frac{\psi_i}{\Psi} = 1-\varepsilon\left(a_i\right)\eta_H$, so that under symmetry (recalling that $a_i\Omega_i = 1-\varepsilon\eta_i$), this means that $n_H = 1/a\Omega$.

Therefore, $CS_H > CS_B$ as

$$\frac{s_H\left(1-a\right)}{1-\varepsilon\frac{1-a}{1-2a}} > \left(\left(1-\varepsilon\frac{5}{4}\right)n_H+1\right)\frac{2s_L}{3\left(1-2\varepsilon\right)}.$$

We can now eliminate the qualities by using (22), which here simplifies to $s_L(1-a_L)\frac{\Omega_H}{\Omega_L} = s_H(1-a_H)$: then $CS_H > CS_B$ as

$$\frac{\left(1-a\right)}{1-\varepsilon\frac{1-a}{1-2a}}\frac{8}{5}\frac{1-\varepsilon\frac{5}{4}}{1-2\varepsilon} > \left(\left(1-\varepsilon\frac{5}{4}\right)n_H+1\right)\frac{2}{3\left(1-2\varepsilon\right)}. \tag{38}$$

Here we can take a value for a and a prior entry cost to find a value for n_H. If we take $a = 1/8$, the above surplus comparison condition (38) reduces to $n_H < \frac{663}{385}$. Setting now

$K = 1$ (note this is above the value we had that supported both types) and $\varepsilon = 1/3$, we use the Zero-Profit condition $n_H = 1/(a\Omega) = 1/\left(1 - \varepsilon\frac{1-a}{1-2a}\right)$ to find

$$n_H = 1/\left(1 - \frac{1}{3}\frac{7}{6}\right) = \frac{18}{11}.$$

This value is below the critical value $n_H < \frac{663}{385}$ we found above, so that indeed the surplus falls with entry of the low quality types. ■

References

[1] Acemoglu, D. and M. K. Jensen. 2013. Aggregate Comparative Statics. *Games and Economic Behavior*, 81, 27-49.

[2] Ambrus, A., E. Calvano, and M. Reisinger. 2016. Either or Both Competition: A "Two-Sided" Theory of Advertising with Overlapping Viewerships. *American Economic Journal: Microeconomics*, 8, 189-222.

[3] Ambrus, A., and M. Reisinger. 2006. Exclusive vs. Overlapping Viewers in Media Markets. Mimeo, University of Munich.

[4] Anderson, S. and S. Coate. 2005. Market Provision of Broadcasting: A Welfare Analysis. *Review of Economic Studies*, 72, 947-972.

[5] Anderson, S., N. Erkal, and D. Piccinin. 2020. Aggregative Games and Oligopoly Theory: Short-run and Long-run Analysis. *Rand Journal of Economics* 51, 470-495.

[6] Anderson, S., O. Foros, and H. J. Kind. 2018. Competition for Advertisers and for Viewers in Media Markets. *Economic Journal*, 128, 34-54.

[7] Anderson, S., O. Foros, H.-J. Kind, and M. Peitz. 2012. Media Market Concentration, Advertising Levels and Ad Prices. *International Journal of Industrial Organization*, 30, 321-325.

[8] Anderson, S. and A. de Palma. 2001. Product Diversity in Asymmetric Oligopoly: Is the Quality of Consumer Goods Too Low? *Journal of Industrial Economics,* 49, 113-135.

[9] Anderson, S. and A. de Palma. 2009. Information Congestion. *Rand Journal of Economics*, 40, 688-709.

[10] Anderson, S. and A. de Palma. 2012. Competition for Attention in the Information Overload Age. *Rand Journal of Economics*, 43, 1-25.

[11] Anderson, S. and J. S. Gans. 2011. Platform Siphoning: Ad-Avoidance and Media Content. *American Economic Journal: Microeconomics*, 3, 1-34.

[12] Anderson, S., A. de Palma and J. F. Thisse. 1992. *Discrete Choice Theory of Product Differentiation.* MIT Press: Cambridge, MA.

[13] Anderson, S. and M. Peitz. 2020. Media See-Saws: Winners and Losers in Platform Markets. *Journal of Economic Theory* 186.

[14] Armstrong, M. 2006. Competition in Two-Sided Markets. *Rand Journal of Economics*, 37, 668-691.

[15] Athey, S., E. Calvano, and J. Gans. 2018. The Impact of Consumer Multi-homing on Advertising Markets and Media Competition. *Management Science*, 64, 1574-1590.

[16] Crampes, C., C. Haritchabalet, and B. Jullien. 2009. Advertising, Competition and Entry in Media Industries. *Journal of Industrial Economics,* 57, 7-31.

[17] Fisher, F. M., J. J. McGowan, and D. S. Evans. 1980. The Audience-Revenue Relationship for Local Television Stations. *Bell Journal of Economics*, 11, 694-708.

[18] Gabszewicz, J. J., D. Laussel, and N. Sonnac. 2004. Programming and Advertising Competition in the Broadcasting Industry. *Journal of Economics & Management Strategy,* 13, 657-669.

[19] Hefti, A. and S. Heinke. 2015. On the Economics of Superabundant Information and Scarce Attention, *Oeconomia*, 5, 37-76.

[20] Hefti, A. and S. Liu. 2019. Targeted Information and Limited Attention. Unpublished manuscript, *Rand Journal of Economics*, forthcoming.

[21] Johnson, J. 2013. Targeted Advertising and Advertising Avoidance. *Rand Journal of Economics*, 44, 128-144.

[22] Luce, R. D. 1959. *Individual Choice Behavior: A Theoretical Analysis.* New York, Wiley.

[23] Mankiw, G. and M. Whinston. 1986. Free Entry and Social Efficiency. *Rand Journal of Economics*, 17, 448-58.

[24] Nageeb, A., G. Lewis, and S. Vasserman. 2019. Voluntary Disclosure and Personalized Pricing. Unpublished manuscript.

[25] Peitz, M. and T. Valletti. 2008. Content and Advertising in the Media: Pay-tv versus Free-to-air. *International Journal of Industrial Organization,* 26, 949-965.

[26] Rochet, J. C. and J. Tirole. 2003. Platform Competition in Two-Sided Markets. *Journal of the European Economic Association,* 1, 990-1029.

[27] Selten, R. 1970. *Preispolitik der Mehrproduktenunternehmung in der Statischen Theorie*, Springer Verlag, Berlin.

[28] Suzumura, K. and K. Kiyono. 1987. Entry Barriers and Economic Welfare. *Review of Economic Studies*, 54, 157-167.

[29] Tan, G. and J. Zhou. 2020. The Effects of Competition and Entry in Multi-Sided Markets, *Review of Economic Studies*, forthcoming.

[30] van Zandt, T. 2004. Information Overload in a Network of Targeted Communication. *Rand Journal of Economics*, 35, 542-560.

[31] von Weiszäcker C. 1980. A Welfare Analysis of Barriers to Entry. *Bell Journal of Economics*, 11, 399-420.