Advertising congestion, time use, and media variety[1]

Simon P. Anderson[2]
University of Virginia

Martin Peitz[3]
University of Mannheim

First version: July 2011. This version: September 2016.
still preliminary and incomplete

[1]We received helpful comments from Emilio Calvano, Michele Polo, Markus Reisinger, and Yossi Spiegel. Anderson gratefully acknowledges financial support from the National Science Foundation, Peitz from the German Science Foundation.

[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 more variety still leads to higher consumer benefits. However, entry of less attractive platforms can increase ad nuisance levels so much that consumers are worse off. Advertiser surplus can also fall with platform entry, though a sufficiently large externality from diversity of viewpoints may warrant encouraging variety.

JEL Classifications: D43, L13

Keywords: media diversity, advertising clutter, limited attention, information congestion, two-sided platforms

1 Introduction

Commercial media typically rely exclusively or predominantly on advertising for revenue and compete for viewers. Since viewers typically dislike advertising, one might think that more competition between media platforms should reduce ad levels. However, this is not what is often observed.

Critics of mass media decry advertising clutter, which implies that many messages are lost 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. We also link advertising choices of media to media diversity. An increase in media diversity (platform entry) leads advertising to replace more content. Advertising becomes more congested making it more difficult for high-quality advertisers to reach consumers and reducing advertiser surplus. Furthermore, despite a positive gain from variety, consumers may 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, 2015) 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 their viewers. The model proposed in this paper 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, Foros, Kind, and Peitz (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, as we noted above. 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 to eliminate competitive bottleneck. Because of multi-homing, no media platform provides sole access to a viewer. A viewer's attention can be seen as a common property resource to which multiple media platforms have access. Therefore platforms "compete" directly with each other for advertisers.

The upshot is to reverse the standard outcome quite radically. The mechanism which reverses standard findings is due to negative externalities due to congestion arising from limited attention. 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 internalization of the negative congestion effect less. 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). 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 model applies to a number of media markets. A special feature of the model is that although advertisers and viewers multi-home, an ad by a particular advertiser is seen at most once by a viewer. We call this type of advertising "synchronized advertising."[1] Synchronized advertising avoids partly wasteful multiple exposures. Radio and television markets as well as internet markets may endogenously lead to this property. To the extent that advertisers can rely on a consumer tracking across channels, our model also applies to internet media platforms.

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). This paper brings in 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 prices in the presence of congestion, as well as dealing with multiple platforms competing for attention.

Our analysis of asymmetric platform oligopoly uses aggregative game tools (Acemoglu and Jenssen, 2013, and Anderson, Erkal, and Picinnin, 2013).

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

[1] 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 (2012, 2013).

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 $\max\{1, \frac{\phi}{A}\}$ 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. Congestion can only arise in oligopoly because a monopoly media platform would never choose $a > \phi$.[2] Thus a monopolist would always price out congestion by 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. The demand price for ads on platform i is then determined as the product of the probability that the viewer is on the platform, that she retains the ad, and the conditional 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.[3]

[2] To see this point, note that with $a > \phi$, a monopolist would only be able to sell an ad at price $\frac{\phi}{a} p(a)$, 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.

[3] In our framework this assumption appears to be natural since viewers obtain zero surplus within an advertiser-viewer interaction and, thus, are ex post indifferent as to which ads they remembered.

If there are a_i ads on platform i, the ad price per 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.

We assume that the demand for ads is well-behaved, so it is not too convex.

Assumption 1 *$p(a)$ is log-concave and twice continuously differentiable in a.*

Even though viewers multi-home, advertisers do not waste impressions since ad placements are perfectly synchronized. In the context of television and radio advertising, one can think of advertising within a given time window. An advertiser who 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.[4] This is the most efficient use of an advertiser's advertising budget and thus the optimal choice of an advertiser. In the context of internet advertising we can think about perfect synchronization to arise when web sites place cookies and share this information with each other (perfect tracking). To extract most from advertisers, they may fill an ad space only with those advertisers which did not yet have contact with the same ad on a different platform. If visits to websites occur in random order, this is equivalent to perfect synchronized advertising in our model.

Viewers

We propose a time-use model of media consumption with identical viewers who mix between media.[5] The outside option has index 0 and gives utility v_0^α per unit of

More generally, one may want to allow for some correlation between product characteristics and the likelihood to recall an ad.

[4]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).

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

time. 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 $(1-a_i)\lambda_i$ 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 advertising sections gives a benefit normalized to zero.

Define $\tilde{\alpha} = \frac{\alpha}{1-\alpha} > 0$. The fraction of time spent on platform i is

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

while the time spent on the outside option is

$$\lambda_0(\mathbf{a}) = \frac{v_0^{\tilde{\alpha}}}{v_0^{\tilde{\alpha}} + \Sigma(s_j(1-a_j))^{\tilde{\alpha}}}.$$

This fractional demand system is in the vein of Luce (1959).

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

$$CS = \left(v_0^{\tilde{\alpha}} + \sum_{i=1}^{n} [s_i(1-a_i)]^{\tilde{\alpha}} \right)^{1-\alpha}. \tag{3}$$

Under symmetry and full coverage ($\lambda_i = 1/n$, $i \neq 0$), the consumer surplus is $n^{1-\alpha}(s(1-a))^\alpha$. For given a, this surplus is increasing in the number of platforms since viewers are variety-loving.

Platforms

We analyze the platform balance problem of delivering reluctant viewers to advertisers. Define $A = \sum_{j=1}^{m} \lambda_j a_j$. The profit function depending on whether there is congestion takes the form

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

For $\phi < A$, platform 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. Interdependence on the viewer side comes from the assumption that consumers decide how to allocate their viewing times λ_i.

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 log-concave in a. We do not develop this case, as it resembles standard models of media platforms (e.g., Anderson and Coate, 2005, Anderson and Peitz, 2015).[6]

3 Analysis

The structure of the model enables us to cast the oligopoly interaction as an aggregate game. This construct was introduced by Selten (1971), and further developed by Acemoglu and Jensen (2013) and Anderson, Erkal, and Piccinin (2016) *inter alia*. For our purpose, an aggregate 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 aggregate 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

[6] Note that 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.,

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

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, 2016: an important distinctive feature of the current situation is that consumer surplus cannot be written as a function of the aggregate).

3.1 Media markets with ad congestion as aggregative games

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 $\Psi = \sum_i \psi_i$. 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 is somewhat challenging. We use the action variable $\psi_i = a_i[s_i(1-a_i)]^{\widetilde{\alpha}}$ defined on 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:

$$\begin{aligned}\Pi_i &= \frac{a_i[s_i(1-a_i)]^{\widetilde{\alpha}}}{\sum_{j=1}^n a_j[s_j(1-a_j)]^{\widetilde{\alpha}}}\phi p(a_i) \\ &= \frac{\psi_i}{\Psi}\phi p(a_i(\psi_i)). \qquad (4)\end{aligned}$$

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.

Notice that the function $a_i[s_i(1-a_i)]^{\widetilde{\alpha}}$ in the profit function (from which we have drawn the aggregator) 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 $\overline{a} = \arg\max_{a_i} a_i[s_i(1-a_i)]^{\widetilde{\alpha}}$ 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: then

$$\frac{d\psi_i}{da_i} = \frac{\psi_i(1-(1+\widetilde{\alpha})a_i)}{a_i(1-a_i)} > 0, \tag{5}$$

and hence we have the action function elasticity as $\eta(\psi_i) = \frac{\psi_i}{a_i}\frac{da_i}{d\psi_i}$, which simplifies to

$$\eta_i \equiv \eta(\psi_i) = \frac{1-a_i}{1-(1+\widetilde{\alpha})a_i} > 0. \tag{6}$$

From (4), the first-order condition defining the inclusive best reply is (recalling that ψ_i enters Ψ)

$$p\left(\frac{1}{\Psi} - \frac{\psi_i}{\Psi^2}\right) + \frac{\psi_i}{\Psi} p' \frac{da_i}{d\psi_i} = 0, \tag{7}$$

where $\frac{da_i}{d\psi_i}$ is given as the reciprocal of (5).[7] We can rewrite this expression as

$$\begin{aligned} 1 - \frac{\psi_i}{\Psi} &= -\psi_i \frac{p'}{p}\frac{da_i}{d\psi_i} \\ &= \varepsilon(a_i)\,\eta(\psi_i), \end{aligned}$$

where we have defined the advertising demand elasticity as $\varepsilon(a_i) \equiv -\frac{p(a_i)'}{p(a_i)} a_i > 0$ and $\eta(\psi_i) = \frac{\psi_i}{a_i}\frac{da_i}{d\psi_i}$. Hence the first-order condition in elasticity form is

$$\frac{\psi_i}{\Psi} = 1 - \varepsilon(a_i)\,\eta(\psi_i), \tag{8}$$

[7] 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 $\overline{a} = \arg\max_a a(1-a)^{\widetilde{\alpha}} = 1/(1+\widetilde{\alpha}) \in (0,1)$. In this case, inclusive best replies are flat.

and thence equilibrium profit can be written as

$$\phi(1-\varepsilon(a_i)\,\eta(\psi_i))p(a_i). \tag{9}$$

We can now show the following result.

Lemma 1 *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 (6) that

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

Furthermore, we observe that the assumption that p is log-concave implies that $\varepsilon(a_i)$ is non-decreasing.

To determine the slope of the inclusive best reply, we differentiate the right-hand side of (8) with respect to ψ_i.

$$\frac{d\Psi}{d\psi_i}=\frac{1-\varepsilon(a_i)\,\eta(\psi_i)+\psi_i[\varepsilon(a_i)\,\eta(\psi_i)]'}{[1-\varepsilon(a_i)\,\eta(\psi_i)]^2}. \tag{10}$$

Hence, the inclusive best reply is upward sloping if $[\varepsilon(a_i)\,\eta(\psi_i)]'>0$. Since (i) $[\varepsilon(a_i)\,\eta(\psi_i)]'=\varepsilon'(a_i)\,\frac{da_i}{d\psi_i}\eta(\psi_i)+\varepsilon(a_i)\,\eta'(\psi_i)$, (ii) the function $\varepsilon(a_i)$ is non-decreasing in a_i, (iii) $\frac{da_i}{d\psi_i}$ and $\frac{da_i}{d\psi_i}$ are increasing, and (iv) $\varepsilon(a_i)$ and $\eta(\psi_i)$ are positive, the result follows.

Using (8) we can rewrite equation (10)

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

As shown in the previous paragraph, $\psi_i[\varepsilon(a_i)\,\eta(\psi_i)]'>0$. 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 the Lemma 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.

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

Proof. Inserting the expression for $da_i/d\psi_i$ from (5) into the inclusive best reply (7) and simplifying yields

$$\left(1-\frac{\psi_i}{\Psi}\right)=-\frac{p'}{p}\frac{a_i(1-a_i)}{(1-(1+\widetilde{\alpha})a_i)}. \tag{11}$$

The left-hand side is decreasing in ψ_i. We now show that the right-hand side is increasing in a_i. The slope of the right-hand side is

$$-\left(\frac{p'}{p}\right)'\frac{a_i(1-a_i)}{1-a_i-\widetilde{\alpha}a_i}-\frac{p'}{p}\frac{(1-2a_i)(1-a_i-\widetilde{\alpha}a_i)+a_i(1-a_i)(1+\widetilde{\alpha})}{(1-a_i-\widetilde{\alpha}a_i)^2}.$$

Since the numerator of this last fraction reduces to $(1-a_i)^2+\widetilde{\alpha}a_i^2$, the second term is positive. Log-concavity of p implies that $-\left(\frac{p'}{p}\right)'$ is positive. Hence, the first term is also positive and so right-hand side of (11) is increasing 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. Thus, a lower equilibrium value a_i is observed for a higher-quality platform. ■

As the previous Lemma has established, ψ_i and a_i are negatively related in equilibrium. Since $\psi_i = a_i\lambda_i$, 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 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, $\frac{\phi}{A}p(a_i)$.[8] 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 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 (2016), the single-homing model (and, by extension, the version of the current time-use model without congestion) exhibits a

[8] 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 writes $\Pi_i = \phi\frac{\psi_i(\Psi)}{\Psi}p(a_i\psi_i(\Psi))$. This function is decreasing in Ψ so that larger values of the aggregate constitute greater competition, and hurt profit: see Anderson, Erkal and Piccinin, 2016, for more on the competitiveness property. Also, higher values of s_i entail higher profit.)

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 single-homing case has more ads and a consequent 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, 2000), higher qualities 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, 2016) 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 ITV premium discussed above), casual evidence on the ad/viewership relation seems quite mixed. There are clearly high quality publications with few ads (such as the Economist), 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 Diversity of platforms

Here we look at the effects of adding more varieties, i.e., platform entry. The consumer surplus analysis is more intricate, so we defer it to a later 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 market shares $\lambda_i = \psi_i/\Psi$ depend on the aggregate. 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 1, inclusive best replies slope up, and so individual actions of incumbent platforms rise. Ad levels rise too because they vary directly with actions (see (5)). Hence we have:

Proposition 2 *Entry of platforms raises advertising levels on all channels.*

This is 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. As we discuss further in the conclusions, softer forms of congestion function than unit elastic one we use could yield ambiguous effects (and a softer transition into a congestion regime from the uncongested case).

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 viewer, $p(a_i)$, goes down. Because A rises, then each incumbent's full price per ad per viewer, $\frac{\phi}{A}p(a_i)$, 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. This is an example of a see-saw effect in two-sided markets (see Rochet and Tirole, 2003, and also Anderson and Peitz, 2016, for a detailed analysis of a variety of see-saw effects in media markets).

With congestion, matters are much different – the see-saw effect works in the opposite direction completely. Indeed, with congestion effects, we have seen in Proposition 2 above that ad levels rise too, meaning that the price paid by consumers on each channel goes up. In turn, this implies that 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*.[9] 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. The effect of entry on consumer surplus is not obvious a priori. 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 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, 2016), this is not the case in our current setting with advertising congestion, as has been shown in the previous proposition.

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 (contrast the central

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

CES/Logit examples in Anderson, Erkal, and Piccinnin, 2016). 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 for a constant elasticity of advertiser demand.

Proposition 3 *In a symmetric market with a constant elasticity of advertiser demand, 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

$$(1-a) - n\tilde{\alpha}\frac{da}{dn} > 0. \tag{12}$$

Using (8), (6), and the fact that, under symmetry, $\psi_i/\Psi = 1/n$, the equilibrium advertising level as a function of firms n is

$$a(n) = \frac{n-1-\varepsilon n}{(\tilde{\alpha}+1)(n-1)-\varepsilon n}. \tag{13}$$

Hence (after simplifying)

$$\frac{da}{dn} = \frac{\varepsilon\tilde{\alpha}}{[(\tilde{\alpha}+1)(n-1)-\varepsilon n]^2},$$

and, using this expression in (12), we want to show that

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

which is equivalent to

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

or, equivalently,

$$(n-1)[(\tilde{\alpha}+1)(n-1)-\varepsilon n]-\tilde{\alpha}n\varepsilon>0. \tag{14}$$

For $a\left(n\right)>0$ to hold in (13), we must have $\varepsilon n<(n-1)$. Given this restriction, (14) 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. ■

We have two main results for consumer surplus under entry. The previous Proposition 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 Piccinnin (2016). 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 (9) and using (6) we have

$$\phi(1-\varepsilon\frac{1-a_L}{1-(1+\widetilde{\alpha})a_L})p(a_L)=K \tag{15}$$

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 (15), we find a_H from the inclusive best replies. To see this, first write the inclusive best reply (8) as

$$\begin{aligned}\Psi &= \frac{\psi_i}{1-\varepsilon\left(a_i\right)\eta\left(\psi_i\right)} \\ &= \frac{a_i\left[s_i(1-a_i)\right]^{\tilde{\alpha}}}{1-\varepsilon\eta_i}\end{aligned} \tag{16}$$

where we recall that $\eta_i=\frac{1-a_i}{1-(1+\tilde{\alpha})a_i}$ from (6) 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 (16) 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 (16) we find n_L from:

$$n_L\left(1-\varepsilon\left(a_L\right)\eta_L\right)+n_H\left(\left(1-\varepsilon\left(a_H\right)\eta_H\right)\right)=1. \tag{17}$$

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 (3), and henceforth we use this transformation (with a slight abuse of notation).[10] When only high type platforms are present, we have

$$CS_H=n_H\left[s_H(1-a_H)\right]^{\tilde{\alpha}}; \tag{18}$$

and when both type are present we have

$$CS_B=n_H\left[s_H(1-a_H)\right]^{\tilde{\alpha}}+n_L\left[s_L(1-a_L)\right]^{\tilde{\alpha}}. \tag{19}$$

Define now

$$\Omega_i\equiv\frac{1-\varepsilon\eta_i}{a_i}=\frac{1-\varepsilon\frac{1-a_i}{1-(1+\tilde{\alpha})a_i}}{a_i}, \tag{20}$$

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 (16) above, we get

$$\left[s_L(1-a_L)\right]^{\tilde{\alpha}}\frac{\Omega_H}{\Omega_L}=\left[s_H(1-a_H)\right]^{\tilde{\alpha}}. \tag{21}$$

Then we can write

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

Using the fixed point condition $n_L\Omega_La_L+n_H\Omega_Ha_H=1$, which condition determines the number of entrants (see (17)), we can rewrite this consumer surplus as

$$CS_B=\left(n_H\frac{\Omega_H}{\Omega_L}+\frac{1-n_H\Omega_Ha_H}{\Omega_La_L}\right)\left[s_L(1-a_L)\right]^{\tilde{\alpha}}. \tag{22}$$

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

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).[11]

We can now use the above analysis to deliver the following result.

Proposition 4 *In an asymmetric market, the entry of an additional platform can decrease consumer surplus.*

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

First simplify by setting $\tilde{\alpha} = 1$ (i.e., $\alpha = 1/2$) and set $\phi = 1$.[12] 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 (16) 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 (22), 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 $a_L = 1/3 >$

[11] 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 (16).

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

$a_H = 1/6$. From the ZPSEE for the low types (15), 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 (16) 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

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

Using the definition of the Ω's from (20) above, we can write them as

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

Inserting these values into the consumer surplus expression when both types are present, (22), we get

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

Now, we know too that consumer surplus before the wave of entry induced by the reduction in entry cost to K is from (18)

$$CS_H = n_H s_H (1-a)$$

and we know that n_H satisfies $\frac{\psi_i}{\Psi} = 1 - \varepsilon(a_i)\eta_H$, 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

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

We can now eliminate the qualities by using (21), 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

$$\frac{(1-a)}{1-\varepsilon\frac{1-a}{1-2a}}\frac{8}{5}\frac{1-\varepsilon\frac{5}{4}}{1-2\varepsilon} > \left(\left(1-\varepsilon\frac{5}{4}\right)n_H + 1\right)\frac{2}{3(1-2\varepsilon)}. \tag{23}$$

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 (23) 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. ■

Entry is bad even 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.

5 Further considerations

We here consider a couple of extensions to the model. The first is to add the possibility of subscription pricing. The second addresses the possibility of advertisers choosing more than one ad per platform.

5.1 Subscription pricing

We now argue that allowing for platforms to charge access prices to consumers does not affect the main results.

Subscription prices can be readily incorporated into the model. Assume that platforms first set such prices, as the first stage in a two-stage game, where the second stage follows the analysis of the main paper for any set of prices chosen such that n platforms are active (meaning that consumers pay the price charged by each of the n platforms). If the price charged by a platform is too high, no consumer pays for that platform. Hence, it is excluded from the advertising sub-game.

The optimal price charged by a platform is then simply the Incremental Value that it offers to consumers over and above the utility they get were they not to pay the subscription price. Indeed, we now modify the consumer utility to read

$$V = \sum_{i=1}^{n} [s_i(1 - a_i^e)\lambda_i^e]^\alpha + (\lambda_0^e v_0)^\alpha - \sum_{i=1}^{n} \sigma_i$$

when the consumer chooses to pay the subscription price for the n platforms from the set of all possible offerings: the superscript e denotes the equilibrium values corresponding to the set chosen.

The incremental value calculation delivers the equilibrium price for platform j as

$$\sigma_j = \left(\sum_{i=1}^{n} [s_i(1 - a_i^e)\lambda_i^e]^\alpha + (\lambda_0^e v_0)^\alpha \right) - \left(\sum_{i=1; i \neq j}^{n} [s_i(1 - a_i^e)\lambda_i^e]^\alpha + (\lambda_0^e v_0)^\alpha \right)$$

where the equilibrium values in the second expression are understood to be those corresponding to the advertising game with j absent. Thus subscription prices are set so as to just keep viewers on board.

There are two confounding effects in the equilibrium price of a platform. First, the incremental value of a high-quality platform is high because it directly contributes a strong value to utility, which, in equilibrium, is further strengthened because it wants a low ad level. Second, though, is the impact on the other platforms' ad levels. From our earlier results, removing a platform decreases Ψ and other ψ's and a's (see Proposition 2 and the preceding analysis). All effects are stronger when a higher quality platform exits, because a platform with a larger inclusive best reply is removed from the total. Thus $s_i (1 - a_i^e)$ is higher when a high quality platform is pulled out than when a low quality one is. Insofar as the first (direct) effect dominates, then equilibrium prices are higher for higher quality firms.

The analysis here is facilitated by the assumption that all consumers have the same utility function. This means that there is no interaction between the subscription price and the advertising side of the analysis. Quality-dependent fixed costs might upset this result, although if these costs did not rise too fast with quality, then only the highest quality platforms would enter because they get both highest subscription revenue (higher prices) and higher ad revenues too (through the higher view-time base). Note that subscription pricing transfers surplus to platforms from consumers and therefore enables more platforms to serve the market.

Heterogeneity of consumer tastes would entail a composition effect that would yield potentially interesting interaction effects.

5.2 Multiple impressions

We have assumed above that advertisers post at most one ad per platform. When there is no congestion, and the market is "fully covered" (meaning that the outside option of non-purchase is not exercised) they have no advantage to placing 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 a benefit from a second ad, or more. 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 (2016), and Athey, Calvano, and Gans (2016). 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 (2016) assume a fixed number of advertisers with the same willingness to pay for impressions and only obliquely allow for advertising nuisance to consumers.

We first determine a sufficient condition for advertisers not to want to place second ads. 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). First consider the case when 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 2 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. 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.[13]

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 $v \geq \frac{p(a_i)}{\left(1 - \frac{\phi}{A} \Sigma_{j \in N} \lambda_j\right)}$. Again, the advertiser demand should not be too heterogeneous over the relevant range.

[13] 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.

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 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.

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 a reduction of 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, when the type of the marginal platform does not change then increasing the quality of incumbent platforms reduces diversity. However, it 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, this is good for consumers even though content is partially replaced by advertising. However, with asymmetric media platforms the latter effect may dominate the benefit from variety and consumers may be worse off when entry costs go down. With a covered market, then total surplus also goes down because, with a covered market, advertising efficiency always decreases without a corresponding increase in market base.

In this paper we have used a congestion function which is unit elastic. This formulation gave rise to stark results that came through very cleanly with the aggregative game approach that we were able to engage as a consequence. One implication of the ϕ/A formulation for congestion is that there is a sharp discontinuity in the model equilibrium behavior on the two sides of the point where congestion kicks in: in particular, the see-saws and comparative statics work in completely opposite directions. Our formulation highlights this clearly. Milder congestion functions would be expected to draw from both sides; which would give us a more nuanced (ambiguous) set of results. The advantage of the current congestion function is that it gives strong possibility results for introducing system-wide congestion. The intuition for the results then come from thinking of attention as a common property resource.

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*, forthcoming.

[3] Ambrus, A., and M. Reisinger. 2006. Exclusive vs overlapping viewers in media markets. Mimeo, University of Munich.

[4] Anderson, S. 2012. Advertising on the Internet. In: Peitz and Waldfogel, *Oxford Handbook of the Digital Economy*, Oxford University Press.

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

[6] Anderson, S., N. Erkal, and D. Picinnin. 2016. Aggregative Oligopoly Games with Entry. CEPR Discussion Papers 9511.

[7] Anderson, S., O. Foros, and H. J. Kind. 2016. Competition, product quality, and multi-purchasing. *International Economic Review*, forthcoming.

[8] Anderson, S., O. Foros, and H. J. Kind. 2016. Competition for advertisers and for viewers in media markets. *Economic Journal*, forthcoming.

[9] 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.

[10] 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.

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

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

[13] Anderson, S. and A. de Palma. 2012b. Oligopoly and Luce's Choice Axiom. *Regional Science and Urban Economics*, 42, 1053-1060.

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

[15] Anderson, Simon, and Bruno Jullien. 2016. The advertising-financed business model in two-sided media markets. *Handbook of Media Economics*, vol. 1A: 41-90.

[16] Anderson and Peitz. 2015. Media See-Saws: Winners and Losers on Media Platforms. University of Mannheim Working Paper 15-16.

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

[18] Athey, S., E. Calvano, and J. Gans. 2016. The Impact of the Internet on Advertising Markets for News Media. Mimeo, Rotman School, University of Toronto.

[19] Corchón, L. 1994. Comparative Statics for Aggregative Games: The Strong Concavity Case. *Mathematical Social Sciences,* 28, 151-165.

[20] Cornes, R. and R. Hartley. 2012. Fully Aggregative Games. *Economics Letters,* 116, 631-633.

[21] Crampes, C., C. Haritchabalet, and B. Jullien. 2009. Advertising, competition and entry in media industries. *Journal of Industrial Economics,* 57, 7-31.

[22] Deneckere, R. and C. Davidson. 1985. Incentives to Form Coalitions with Bertrand Competition. *Rand Journal of Economics*, 16, 473-486.

[23] 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.

[24] 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.

[25] Luce, R. D. 1959. *Individual choice behavior: A theoretical analysis.* New York, Wiley.

[26] Peitz, Martin, and Markus Reisinger. 2016. The economics of internet media. *Handbook of Media Economics*, vol. 1A. Elsevier.

[27] Peitz, M. and T. M. Valletti. 2008. Content and advertising in the media: Pay-tv versus free-to-air. *International Journal of Industrial Organization,* 26(4), 949-965.

[28] Rochet, J. C. and J. Tirole. 2003. Platform competition in two-sided markets. *Journal of the European Economic Association,* 1(4), 990-1029.

[29] Salant, S. W., S. Switzer, and R. J. Reynolds. 1983. Losses from Horizontal Merger: The Effects of an Exogenous Change in Industry Structure on Cournot-Nash Equilibrium. *Quarterly Journal of Economics,* 98, 185-199.

[30] Schmalensee, R., 1983, Advertising and Entry Deterrence: An Exploratory Model. *Journal of Political Economy* 91, 636-653

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

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