How Does Advertising Influence Media Competition? A Two-Sided Market Perspective.*

David Henriques†

Ofcom – UK Regulator and Competition Authority

This version (preliminary): October 5, 2012

Abstract

In this paper we present a three-party model involving: content subscribers that differ in their sensitivity to advertisements, advertising firms that may invest in advertisements' quality, and media platforms competing simultaneously and non-cooperatively in the prices of contents and advertising slots. We investigate the determinants of advertising quality and the implications that the proportion of ad-sensitive consumers has on platforms' profits, pricing structure and agents' payoffs. We compare the market equilibrium levels of advertisements' airtime and quality to the socially optimal levels.

We show first that the proportion of ad-sensitive consumers and the number of competing media platforms are two determinant factors of advertising quality. Second, under some conditions, as the proportion of ad-sensitive consumers increases: (i) advertiser (subscriber) price increases (decreases), (ii) platforms may obtain higher profit levels, (iii) subscribers are better-off and, (iv) advertiser surplus increases iff ad-quality is sufficiently high. Third, we show that (i) the market solution under-provides advertisements' quality, but it might under- or over-provide ad-airtime, (ii) ad-quality regulation may work implicitly as ad-airtime regulation, and (iii) the ad-airtime permitted is longer with exclusive regulation on advertising airtime than if the regulator can set both the ad-quality standard and the ad-airtime.

Keywords: Access pricing, Advertising, Media competition, Two-sided markets.

JEL Classification: D62, L13, M37.

* I am grateful to Vasco Santos for extensive comments, suggestions and discussions on matters related to this paper. I also wish to thank Heski Bar-Isaac, Patrick Rey and Robin Lee for useful comments on early versions of this paper. All errors are my own. Financial support from the Amélia de Mello Foundation, the Foundation for Science and Technology (BD/47421/2008) and Nova School of Business and Economics is acknowledged. The analyses, opinions and findings in this paper represent the views of the author, they are not necessarily those of Ofcom.

† Contact: dthenriques@gmail.com. Home-page: `http://dthenriques.orgfree.com`.

1 Introduction

The media industry. The commercial media industry differs from traditional industries because it simultaneously serves two groups of agents mutually linked by cross-group network externalities: the subscribers (consumers)[1] who may or may not be sensitive to the number and quality of advertisements, and the advertising firms whose profits increase with the number of subscribers watching commercials.

In developed countries consumers spend a significant share of their leisure time connected to mass media platforms such as television, the Internet and radio. For instance, the Americans spend on average around four hours per day watching TV and more than one hour and twenty minutes using and watching videos on the Internet.[2] In addition, the Americans aged 15 or older spent on average 12 minutes a day more watching television in 2009 than they did in 2007, before the recession hit.[3] According to the Nielsen Company (July, 2009), the average American household in 2009 had 2.86 TV sets but only 2.5 individuals.[4] In Japan, the average time spent watching TV is three hours and thirty minutes per day. In Europe, slightly more television is watched. By subtracting hours of sleep, work, commuting, and hours eating from the daily total of 24 hours, one may argue that leisure time is mostly devoted to watching or interacting with media platforms such as TV, the Internet and radio.[5] Also, extensive scientific study[6] of media and entertainment reveals its increasing significance in the current information society.

Mass media platforms offer an opportunity for firms to advertise to a large pool of consumers. In fact, it is estimated that the average American is exposed to 61 minutes of TV ads per day.[7] Some firms spend billions of dollars per year in advertising,[8] an industry that is expected to reach a revenue of over US $780 billion worldwide in 2010, with the largest share of it going to TV broadcasting.[9]

[1] We will use the terms "subscriber" and "consumer" interchangeably.

[2] *Television, Internet and Mobile Usage in the U.S., Three Screen Report Volume 7 - 4th Quarter 2009*, The Nielsen Company, 2010.

[3] "Americans are spending more time watching TV and sleeping as unemployment rises, survey finds" in *What Would You Do With an Extra Hour?,* Wall Street Journal, June 23, 2010.

[4] Article *More than Half the Homes in U.S. Have Three or More TVs* available at `http://blog.nielsen.com/nielsenwire/media_entertainment/more-than-half-the-homes-in-us-have-three-or-more-tvs`.

[5] "TV is the dominant medium for media consumption and advertising. Computer usage has supplanted radio as the second most common media activity and print ranks fourth," The New York Times, *8 Hours a Day Spent on Screens, Study Finds*, March 27, 2009.

[6] For example, business schools such as NYU-Stern or IESE in NYC and the Center for Media Design at Ball State University have programs dealing with entertainment, media and technology. The CRE, created by the Nielson Company, has the Media Consumption and Engagement Committee with the mission "to improve and evolve audience measurement through comprehensive and ongoing study of media consumption." See `http://www.researchexcellence.com`.

[7] The New York Times, *8 Hours a Day Spent on Screens, Study Finds*, March 27, 2009.

[8] "*For example, Advertising Age (2005) reports that, in 2003 in the U.S., General Motors spent $3.43 billion to advertise its cars and trucks; Procter and Gamble devoted $3.32 billion to the advertisement of its detergents and cosmetics; and Pfizer incurred a $2.84 billion dollar advertising expense for its drugs. Advertising is big business indeed.*" (Bagwell, 2005).

[9] See Karawang Business, Information, Tips and Solutions for Business and Finance, *World Cup 2010: World Advertising Expenditures, Translucent US $780 Billion*, June 13, 2010.

Advertising plays a significant role in the TV broadcasting business model in most western countries. In the US, the frequency and length of commercial breaks are generally unregulated. Nonetheless, in most countries the regulatory authorities limit the advertising airtime on TV channels. As an example, advertising is limited to an average of six minutes per hour in France; the limit goes up to nine minutes in Germany while English regulators impose a seven-minute ceiling.

Description of the paper. We utilize a model of subscriber-advertiser supported broadcasting in a two-sided market[10] framework that yields predictions on how advertising quality is determined by firms and how consumer ad-sensitiveness affects media platforms' competition (access prices and profits) and the platforms' business model. We also address possible market failures in the media industry, i.e., whether the market provision of advertising airtime and quality levels differs from the socially-optimal values.

We consider a three-party model with content subscribers, advertising firms and media platforms. The main features of the model are as follows.
(i) Subscribers, who are also consumers in the goods market, extract a benefit from the content of media platforms, e.g., information or entertainment, and differ in their attitudes towards the number and quality of advertisements. We assume that a proportion λ of subscribers are *ad-sensitive*, while the remaining $(1-\lambda)$ are *ad-indifferent*. This assumption is crucial in the paper: the main results are tied to the proportion of ad-sensitive subscribers.[11]

The existence of a proportion of ad-sensitive subscribers is supported by the advertising economics literature, namely the persuasive and the informative views on ads. The persuasive view states that advertisements alter consumers' preferences and augment product differentiation and brand loyalty. As a result, advertising boosts firms' profits. The informative view holds that many markets suffer from imperfect consumer information because searching costs may prevent consumers from learning of a product's existence, quality and price. Advertising comes out as one of the endogenous answers that markets present in the face of imperfect information, supplying consumers with further information at low cost, e.g., regarding firm location, product description or prices. Both advertising views will be considered in our model.[12]

The ad-sensitive subscribers are averse to advertising airtime, or equivalently are averse to the number of commercials, while also appreciating advertisements of superior quality.

[10] In a two-sided market, two different groups of agents relate to each other through a platform. The latter sets access prices taking into account the cross-group externalities. For a general introduction to the theory of two-sided markets, see the seminal papers of Rochet and Tirole (2003, 2006) and Armstrong (2006).

[11] Ad-sensitive subscribers are sensitive to ad-quality but simultaneously dislike ad-airtime. Without a fraction of ad-sensitive subscribers, advertising quality would not play any role and would be optimally set at zero.

[12] Another theory holds that advertising is a complement to the consumption of the advertised good. According to this perspective advertising does not transform consumers' preferences and need not supply any information. For example, this happens if the consumption of a good generates more prestige to consumers when the good is advertised. See Bagwell (2005) on the economic analysis of advertising.

For example, ad-sensitive subscribers enjoy the participation in ads of famous performers or athletes.[13] The nuisance perceived by ad-sensitive subscribers is related to the duration or number of commercials. These affect negatively the "consumption" of platform contents. In particular, this negative effect of commercials may be understood as the boredom and wasted time that the ad-sensitive subscribers bear each time there is a commercial break on TV or the advertising pops up on the screen when browsing the Internet implying slower navigation and extra mouse clicks to close unwanted windows. It is implicitly assumed in our framework that ad-sensitive subscribers have no way to receive the media platform contents while skipping advertisements. There is empirical evidence that TV subscribers attempt to avoid the advertising time. For example, Wilbur (2008) estimated a two-sided model of the TV industry and found that viewers tend to be averse to commercials.

Ad-indifferent subscribers are insensitive both to the number of ads and their quality. Hence, this type of subscribers may be seen as having access to an ad-avoidance technology, e.g., TiVo or a pop-up stopper. Furthermore, regardless of type, every individual subscriber has an idiosyncratic preference for his favorite media platform, i.e., his favorite type of programming.

(ii) Advertising firms obtain an indirect benefit from contacting and informing potential customers (the informative view). For example, advertising products and promotions to consumers expands the demand for the advertised goods. Media platforms can earn revenues by charging firms for advertising airtime. Advertising firms have idiosyncratic preferences regarding the type of programming that best matches the product they wish to publicize. Firms may upgrade their advertisements by investing in ad-quality. On the one hand, our model assumes that quality is the tool by which firms counteract the nuisance that ad-airtime inflicts on ad-sensitive media subscribers. On the other hand, ad-quality increases the willingness-to-pay for the advertised products of ad-sensitive consumers (the persuasive view).

(iii) Media platforms compete non-cooperatively in prices by setting them simultaneously and independently, selling ad-airtime to firms and content to subscribers. Media platforms are profit maximizers whose equilibrium access prices are determined by balancing end-users' demands while taking into consideration cross-group externalities. The number of advertising firms and subscribers on each media platform is determined endogenously by the model in a two-sided framework. The number of firms willing to advertise on a platform depends positively on the number of subscribers, while the number of ad-sensitive subscribers on the platform is negatively affected by the number of advertisements and positively affected by ad-quality.

We show first that the proportion of ad-sensitive consumers and the number of com-

[13] Nike's *Write the Future* commercial campaign during Fifa World Cup 2010 had the participation of some of the best soccer players in the world and is available at `http://www.youtube.com/watch?v=idLG6jh23yE` since May, 17th of 2010. The commercial hit almost 20 million views in only two months after its release on YouTube.

peting media platforms are two determinant factors of advertising quality. Second, under some conditions, as the proportion of ad-sensitive consumers increases: (i) advertisers' (subscribers') price increases (decreases), (ii) platforms may achieve a higher profit level, (iii) subscribers are better-off, and (iv) advertisers' surplus increases if and only if ad-quality is sufficiently high. In fact, if firms improve advertising quality, media platforms will adjust by obtaining more (less) revenue on the advertising (subscription) side of the market. Third, we divide the welfare analysis into two parts: the first-best, in which the market regulator may choose both the ad-airtime and the ad-quality, and the second-best, in which the regulator decides only on ad-airtime or only on ad-quality. We show that in comparison to the first-best analysis, the market solution under-provides ad-quality, but it might under- or over-provide advertising airtime. The second-best analysis shows that ad-quality regulation may work implicitly as ad-airtime regulation, but the converse does not hold true. Moreover, if the regulator decides only on ad-airtime, then the airtime allowed to commercials will be longer than in the first-best solution.

Related literature. This paper intends to contribute to the economic analysis of advertising, namely the market provision of advertising quality, and the two-sided markets' literature. Seminal normative work on advertising, such as Steiner (1952) and Spence and Owen (1977), tended to focus on the benefits that commercials generate to the audience but ignored the surplus obtained by the advertising firms. The assumptions of fixed levels of advertising airtime and prices prevent the analysis of whether market under- or over-provision of advertisements took place.

More recently, Anderson and Coate (2005) explored the market failure in the broadcasting industry by modeling how media platforms fulfill their role of providing contents to subscribers and simultaneously supplying eye-balls to advertising firms. Their work connects the goods market to the advertising market and analyzes the trade-off between the nuisance stemming from commercial breaks during the broadcasts and the informational gains generated by the content of these commercials. Nonetheless, the authors ignored the possibility of firms investing in ad-quality. They show that the market equilibrium may under- or over-provide advertising airtime, depending on the nuisance cost to viewers, the substitutability of programs, and the expected benefits to advertising firms from contacting viewers. Gabszewicz, Laussel and Sonnac (2005) studied whether advertising subsidizes the newspaper prices charged to readers. They show that in a two-sided market framework with advertisers on one side and readers on the other, the answer depends on the readership's attitude towards advertising, i.e., it depends on the proportion of readers that are ad-lovers or ad-avoiders. Dukes (2004) shows that less product differentiation or more media differentiation leads to higher market levels of advertising. In particular, if media is sufficiently differentiated, the advertising levels will surpass the socially optimal solution. Dukes (2006) investigates how competition in the media market shapes decisions about advertising and program quality. Dukes shows that product differentiation using advertising is more effective when media markets are less competitive, increasing the prices for advertised products. Gantman and Shy (2007) use an advertising-supported media model (free-to-air broadcasting) to study the firms' incentives to improve the quality of their advertisements. They show that if improving ads' quality is profitable to firms, then it will be unprofitable to broadcasters.

This paper is also related to the two-sided markets literature. The seminal articles by Rochet and Tirole (2003) and Armstrong (2006) investigate the determinants of the price balance between two groups of end-users when each group exerts an externality over the other, and both are intermediated by a platform. Some of the discussed determinants of the price balance are: (i) possibility of multi-homing (i.e., some end-users subscribe or use more than one platform), (ii) platform differentiation, (iii) presence of same-side externalities, (iv) platform compatibility, (v) per-transaction (or lump-sum) pricing and relative size of cross-group externalities.

Our paper differs from the above literature in at least three aspects. First, the choice of ad-sensitive subscribers on media platforms is moulded not only by the subscription price and the advertising airtime in each platform but also by the average quality of commercials. Second, advertisers may invest in ad-quality. Third, the number of advertisers and subscribers on each media platform are determined endogenously in a two-sided market environment.

2 The Subscriber-Advertiser Supported Media Model

This section characterizes the participating agents in the subscriber-advertiser supported media industry and describes how they interact.

Consider a model with N media platforms indexed by $i = 1, 2, ..., N$ competing simultaneously and independently in two markets: (i) content subscription to subscribers and (ii) advertising airtime to firms whose profit level increases in the number of subscribers (potential customers). We assume that media platforms charge a fixed price to agents in each side of the market, e.g. a monthly flat rate. Hence, the pricing scheme does not depend explicitly on the number of agents on the other side of the market. For an illustration, think of TV broadcasters. We now characterize each set of agents.

Subscribers. There is a mass one of subscribers. Each subscriber may choose one media platform among N. Subscribers are heterogeneous in two dimensions: (i) with respect to programming in each platform, and (ii) regarding the attitude towards advertising. In particular, a proportion λ, $0 \leq \lambda \leq 1$, of subscribers is ad-sensitive, i.e., its indirect utility depends on the ad-airtime and the average quality of ads. The remaining $1 - \lambda$ are ad-indifferent, i.e., their indirect utility does not change with either the duration, or the quality of advertisements. We will refer to ad-sensitive consumers as S-type consumers and to ad-indifferent ones as being I-type.

Table 1: Notation for subscribers

λ	proportion of ad-sensitive subscribers
$\hat{x}_i$	advertising volume on platform i
y	index for ad-sensitive subscribers
$\hat{y}_i$	mass of ad-sensitive subscribers on platform i
z	index for ad-indifferent subscribers
$\hat{z}_i$	mass of ad-indifferent subscribers on platform i
v	subscriber's gross utility of accessing an ad-free platform
q_i	average ad-quality of platform i
$T_S(y, i)$	ad-sensitive subscriber's disutility from not being able to subscribe to his preferred programme at platform i
$T_I(z, i)$	ad-indifferent subscriber's disutility from not being able to subscribe to his preferred programme at platform i
w_i	subscription price charged by platform i

Formally, the indirect utility derived by an ad-sensitive subscriber, indexed by y, from subscribing to network i is given by

$$U_y \equiv v\left(1 - \hat{x}_i + q_i\right) - w_i - T_S(y, i) \tag{1}$$

where $\hat{x}_i$ denotes the advertising airtime in platform i. The indirect utility of an ad-indifferent subscriber, indexed by z, is defined by

$$U_z \equiv v - w_i - T_I(z, i) \,. \tag{2}$$

Functions $T_S(y, i)$ and $T_I(z, i)$ denote the subscribers' disutility of being prevented from watching their preferred programs.[14] For simplicity, it is assumed that not watching any program yields a zero net utility.

The key differences between the utility functions of ad-sensitive and ad-indifferent subscribers are: (i) the effect that advertising volume exerts on S-type subscribers (but not on I-type subscribers), and (ii) the average advertising quality level $q_i \equiv \int_0^{\hat{x}_i} q_x dx \,/\hat{x}_i$, where q_x denotes advertiser x's ad-quality, that positively affects U_S.

Remark 1 It is implicitly assumed that subscribers receive no other benefits in the goods market than those inherent to the product itself from purchasing advertised products. In other words, firms extract all the incremental surplus that advertising generates for their goods. This simplification allows us to focus on the media market without further

[14]The model presented here with N media platforms also includes the duopoly model à la Hotelling (1929) as a special case. The T_S and T_I functions represent the degree to which platforms are substitutes for ad-sensitive and ad-indifferent subscribers, respectively. For example, in the standard duopolistic Hotelling model with platforms located at the extremes of the unit interval, $T(x, 0) = \mu x$ and $T(x, 1) = \mu(1 - x)$ where μ is the transportation cost. The model assumes that subscribers are symmetrically distributed regarding their programming preferences (see properties P.1 and P.1$'$ below).

concerns about an endogenous distribution of informational gains between subscribers and firms in the market for the advertised good. Hence, subscribers allocate themselves across media platforms so as to maximize net subscription benefits according to (1) and (2). ■

Let $\hat{y}_i \equiv \hat{y}_i(w_i, \hat{x}_i, q_i, \mathbf{w}_{-i}, \hat{\mathbf{x}}_{-i}, \mathbf{q}_{-i})$ and $\hat{z}_i \equiv \hat{z}_i(w_i, \mathbf{w}_{-i})$ denote the measure of S-type and I-type consumers, respectively, subscribing to platform i's service, such that

$$\begin{aligned}
\hat{\mathbf{x}}_{-i} &\equiv (x_1, ..., x_{i-1}, x_{i+1}, ..., x_N), \\
\mathbf{q}_{-i} &\equiv (q_1, ..., q_{i-1}, q_{i+1}, ..., q_N), \\
\mathbf{w} &\equiv (w_1, ..., w_N), \\
\mathbf{w}_{-i} &\equiv (w_1, ..., w_{i-1}, w_{i+1}, ..., w_N).
\end{aligned}$$

I assume that $\hat{z}_i$ satisfies the following properties:

P.1 (symmetry) For any vector $\mathbf{w}$ with $w_i = w_j$ for any platform i and j, then $\hat{z}_i(\mathbf{w}) = \hat{z}_j(\mathbf{w})$.

P.2 (monotonicity) For any pair of platforms $i, j = 1, ..., N$ and $i \neq j$, $\hat{z}_i(w_i, \mathbf{w}_{-i})$ is twice differentiable with respect to w_i and each $w_j \in \mathbf{w}_{-i}$ and decreases with w_i and increases with w_j. Moreover, it strictly decreases with w_i and strictly increases with w_j for $\hat{z}_i \in (0, 1)$.

More rigorously, P.2 can be defined as follows. Given $\mathbf{w}_{-i}$, let $\underline{w}_i$ be the maximum w_i making $\hat{z}_i(w_i, \mathbf{w}_{-i}) = 1$ and let $\overline{w}_i$ be the minimum $w_i \in \mathbb{R}_+$ making $\hat{z}_i(w_i, \mathbf{w}_{-i}) = 0$. Then, $\hat{z}_i$ strictly decreases with w_i for $w_i \in [\underline{w}_i, \overline{w}_i]$. Similarly, given $\mathbf{w}_{-j}$ with $j \neq i$, let $\underline{w}_j$ be the maximum $w_j \in \mathbb{R}_+$ making $\hat{z}_i(w_i, \mathbf{w}_{-i}) = 0$ and let $\overline{w}_j$ be the minimum $w_j \in \mathbb{R}_+$ making $\hat{z}_i(w_i, \mathbf{w}_{-i}) = 1$. Then, $\hat{z}_i$ strictly increases with w_j for $w_j \in [\underline{w}_j, \overline{w}_j]$.

P.3 (full coverage and single-homing) $\sum_{i=1}^{N} \hat{z}_i(w_i, \mathbf{w}_{-i}) = 1$ for all $\mathbf{w} \in \mathbb{R}_+^N$.

Similarly, I assume that each $\hat{y}_i$ satisfies the following properties:

P.1′ (symmetry) For any vector $(\mathbf{w}, \mathbf{x}, \mathbf{q})$ with $w_i = w_j$, $\hat{x}_i = \hat{x}_j$ and $q_i = q_j$ for any platforms $i, j = 1, ..., N$, and $i \neq j$ then $\hat{y}_i(w_i, \hat{x}_i, q_i, \mathbf{w}_{-i}, \hat{\mathbf{x}}_{-i}, \mathbf{q}_{-i}) = \hat{y}_j(w_j, \hat{x}_j, q_j, \mathbf{w}_{-j}, \hat{\mathbf{x}}_{-j}, \mathbf{q}_{-j})$.

P.2′ (monotonicity) For any platform $i = 1, ..., N$, $\hat{y}_i(w_i, \hat{x}_i, q_i, \mathbf{w}_{-i}, \hat{\mathbf{x}}_{-i}, \mathbf{q}_{-i})$ is twice differentiable with respect to w_i, $\hat{x}_i$, q_i and to each component of $\mathbf{w}_{-i}$, $\hat{\mathbf{x}}_{-i}$, $\mathbf{q}_{-i}$ and is decreasing in w_i, $\hat{x}_i$, $\mathbf{q}_{-i}$ and increasing in $\mathbf{w}_{-i}$, $\hat{\mathbf{x}}_{-i}$, q_i. It strictly decreases with w_i, $\hat{x}_i$, $\mathbf{q}_{-i}$ and strictly increases in each component of $\mathbf{w}_{-i}$, $\hat{\mathbf{x}}_{-i}$, q_i for $\hat{y}_i \in (0, 1)$. For the sake of clarity, when we say that a function is (de)increasing in a vector, we mean in each component of the vector.

P.3′ (full coverage and single-homing) $\sum_{i=1}^{N} \hat{y}_i(w_i, \hat{x}_i, q_i, \mathbf{w}_{-i}, \hat{\mathbf{x}}_{-i}, \mathbf{q}_{-i}) = 1$ for all $(\mathbf{w}, \mathbf{x}, \mathbf{q}) \in \mathbb{R}_+^N \times [0, 1]^N \times \mathbb{R}_+^N$.

Properties 1, 1′, 2, 2′ and 3, 3′ are satisfied by the standard fully-served Hotelling duopoly and the circular city model with $N = 2$ or 3 (Salop, 1979), respectively. For

$N > 3$, the model here is more natural than the circular city model since in the latter, a (slight) price change of platform i affects only the demands of its direct neighbors (platform $i-1$ and platform $i+1$), but does not affect other platforms. Within the context of broadcasting or Internet search engines, all platforms compete directly with each other for all customers, and not only with the two neighboring platforms for a specific subset of subscribers. The properties of symmetry and full coverage with single-homing together imply that $\hat{z}_i = \hat{y}_i = 1/N$ for all $i = 1, ..., N$ if $w_i = w_j$, $x_i = x_j$, $q_i = q_j$ for all $i, j = 1, ..., N$, and $i \neq j$. Full coverage means that the constant utility v from subscribing to one of the networks is sufficiently high such that all potential subscribers end up joining at least one platform, while single-homing induces subscribers to choose only one platform.[15] Since the total measure of subscribers is equal to one, under the full coverage and single-homing assumptions the measure of subscribers on platform i corresponds to its market share $D_i \equiv \lambda \hat{y}_i + (1-\lambda)\hat{z}_i$ on the subscription side of the market.

Advertising firms. There is a mass one of advertising firms. Each advertising firm has access only to the subscribers of one of the platforms, i.e., there is no multihoming. Firms use media platforms as an advertising outlet to reach consumers and thus increase profits.

TABLE 2: Notation for advertising firms

q_x	firm x's chosen ad-quality
p_i	advertising price charged by platform i
$T(x, i)$	firm x's surplus loss when it advertises on platform i
D_i	total mass of consumers on platform i
β	cost of advertising quality
ϕ	advertising effectiveness regarding the informative effect

Let firm x's surplus of advertising on platform i be defined as follows:

$$\Pi_x \equiv \phi D_i + q_x \lambda \hat{y}_i - \beta q_x^2/2 - p_i - T(x, i). \tag{3}$$

Function (3) measures the extra profit to firm x of advertising with quality q_x on platform i. If the firm does not advertise, no surplus will be generated.[16]

Firm x's gross surplus of advertising on platform i with quality q_x is measured by $\phi D_i + q_x \lambda \hat{y}_i$.[17] In particular, the term ϕD_i represents the increase in willingness-to-pay for the advertised product, ϕ, by all subscribers on platform i or alternatively the

[15] We assume that subscribers cannot access more than one media platform, e.g., simultaneously watch more than one TV channel.

[16] In other words, the costs and benefits in equation (3) are net of firm x's profit when it does not advertise.

[17] Note that (3) is compatible with both classical views of advertising: informative and persuasive. The term ϕD_i corresponds to the informative role of advertising since all subscribers of platform i learn about

extra profit arising from sales due to the information generated by the advertisement. Additionally, if x's advertising quality is strictly positive, $q_x > 0$, then $q_x \lambda \hat{y}_i$ captures an additional demand expansion arising from ad-sensitive consumers.[18] Firm x pays a cost of implementing quality level q_x equal to $\beta q_x^2/2$, with $\beta > 0$. The higher β, the more expensive the ad-quality technology. The incentive for firm x to spend resources to improve ad-quality is driven by the increase in the willingness-to-pay (perceived quality) for the advertised product displayed by ad-sensitive consumers. Moorthy and Zhao (2000) found evidence that advertising expenditure and perceived quality are in general positively correlated for both durable and nondurable goods, even after accounting for objective quality, price and market share.

Function (3) also tells us that on the one hand firm x's surplus is increasing in the number of subscribers of platform i but on the other hand it is decreasing in the distance to that platform. More specifically, advertising firms have idiosyncratic preferences regarding media platforms. They are concerned with matching the type of product to be advertised with the scope of programming broadcast by the platform. The distance from firm x's preferred type of programming to the actual type of programming broadcast by platform i is measured by $T(x, i)$. We assume in the model, for the sake of simplicity, that firms are symmetrically distributed regarding the media programming preferences (see property P.1$''$ below).

Remark 2 (a simple illustration) Consider two TV broadcasters: a sport station (entertainment content) and a news station (information content) corresponding to the first and second extremes of a broadcasting unit interval on $[0, 1]$, respectively. Suppose that both broadcasters charge the same price and there are few sport lovers and many viewers with a strong preference for the news station, i.e., both ad-indifferent and ad-sensitive groups of consumers obey a nonsymmetric negatively-skewed distribution on $[0, 1]$ regarding their idiosyncratic preferences for sports and news. In this case, a sport club (firm), located at 0, will benefit from a larger pool of subscribers if it advertises its events on the news station. Nonetheless, by advertising on the news station, the club would face a strong negative effect measured by $T(0, 1)$. This distance represents the low efficacy of sports advertising on news platforms, since news subscribers do not have a particular interest in sport ads whereas sports fans do. Hence, despite there only being a few sport lovers, the club may still prefer to advertise on the sports platform instead of

the existence and features of x's product upon watching the advertisement. The term $q_x \lambda \hat{y}_i$ captures the persuasive effect of advertising. The persuasive effect results in the augmented willingness-to-pay of S-type subscribers on platform i that firm x captures. Recall that only S-type subscribers are liable to be persuaded.

[18]Suppose that firm x's commercial had been made with a well-known public figure rather than an anonymous performer. Then, S-type subscribers would have derived higher utility from watching the media channel where the advertisement was screened and would have been willing to pay more for x's advertised product or buy more units of it. However, I-type subscribers would not have been willing to pay more for the product and would not have bought more units regardless of the performer who did the ad. Recall that I-type subscribers are only concerned with the product's features (objective information) and not the way those features are presented.

the news one, in order to avoid the low efficacy that sports ads have on news subscribers. ■

Let $\hat{x}_i \equiv \hat{x}_i\left(p_i, \hat{y}_i, \hat{z}_i, \mathbf{p}_{-i}, \hat{\mathbf{y}}_{-i}, \hat{\mathbf{z}}_{-i}\right)$ denote the measure of advertisers that subscribe to platform i, where

$$\begin{aligned} \mathbf{p}_{-i} &\equiv \left(p_1, ..., p_{i-1}, p_{i+1}, ..., p_N\right), \\ \hat{\mathbf{y}}_{-i} &\equiv \left(\hat{y}_1, ..., \hat{y}_{i-1}, \hat{y}_{i+1}, ..., \hat{y}_N\right), \\ \hat{\mathbf{z}}_{-i} &\equiv \left(\hat{z}_1, ..., \hat{z}_{i-1}, \hat{z}_{i+1}, ..., \hat{z}_N\right). \end{aligned}$$

I assume that $\hat{x}_i$ satisfies the following properties:

P.1$''$ (symmetry) For any vector $(\mathbf{p}, \mathbf{y}, \mathbf{z})$ with $p_i = p_j$, $y_i = y_j$ and $z_i = z_j$ for any pair of platforms i and j, $i \neq j$, then $\hat{x}_i\left(p_i, \hat{y}_i, \hat{z}_i, \mathbf{p}_{-i}, \hat{\mathbf{y}}_{-i}, \hat{\mathbf{z}}_{-i}\right) = \hat{x}_j\left(p_j, \hat{y}_j, \hat{z}_j, \mathbf{p}_{-j}, \hat{\mathbf{y}}_{-j}, \hat{\mathbf{z}}_{-j}\right)$.

P.2$''$ (monotonicity) For any platforms $i, j = 1, ..., N$ and $i \neq j$, $\hat{x}_i\left(p_i, \hat{y}_i, \hat{z}_i, \mathbf{p}_{-i}, \hat{\mathbf{y}}_{-i}, \hat{\mathbf{z}}_{-i}\right)$ is twice differentiable in p_i, $\hat{y}_i$, $\hat{z}_i$ and in each component of $\mathbf{p}_{-i}$, $\hat{\mathbf{y}}_{-i}$, $\hat{\mathbf{z}}_{-i}$ and decreases in p_i, $\hat{\mathbf{y}}_{-i}$, $\hat{\mathbf{z}}_{-i}$ and increases in $\mathbf{p}_{-i}$, $\hat{y}_i$, $\hat{z}_i$. It strictly decreases in p_i, $\hat{\mathbf{y}}_{-i}$, $\hat{\mathbf{z}}_{-i}$ and strictly increases in $\mathbf{p}_{-i}$, $\hat{y}_i$, $\hat{z}_i$ for $\hat{x}_i \in (0,1)$.

P.3$''$ (full coverage and single-homing) $\sum_{i=1}^{N} \hat{x}_i\left(p_i, \hat{y}_i, \hat{z}_i, \mathbf{p}_{-i}, \hat{\mathbf{y}}_{-i}, \hat{\mathbf{z}}_{-i}\right) = 1$ for all $(\mathbf{p}, \mathbf{y}, \mathbf{z}) \in \mathbb{R}_+^{3N}$.

Note that we are implicitly restraining the analysis to a set of transportation costs $T(x,i)$, $T_S(y,i)$ and $T_I(z,i)$ whose functional forms ensure the properties of symmetry, monotonicity and full coverage regarding the demand functions.

Media platforms. Platforms provide horizontally differentiated contents and each individual platform has the capacity to fully cover both sides of the market. Platforms have two revenue sources, namely advertising firms and subscribers. The profit function[19] of platform i is defined as follows,

$$\Pi_i = p_i.\hat{x}_i + w_i.D_i, \tag{4}$$

where $D_i = \lambda \hat{y}_i + (1-\lambda)\hat{z}_i$. Platform i chooses a pair of access prices (p_i, w_i) that maximizes (4).

3 The Subgame-perfect Nash Equilibrium

This is a three-party model with media platforms, subscribers, and advertising firms. These agents interact according to the following three-stage game.

[19] As a matter of simplicity, we assume that platforms have zero costs in providing their services or alternatively p_i and w_i may be interpreted as markups over constant marginal costs. Hence, platforms face the same level of costs regardless of the broadcast mix of advertising and regular programming. This simplification allows us to disregard eventual complications that costs may introduce in the analysis.

Table 3: The timing of the model

I.	**Media platforms** choose simultaneously and independently the pair of prices (p_i, w_i). Each platform i chooses (p_i, w_i) such that it maximizes (4).
II.	**Firms** decide to buy advertising airtime from one of the media platforms depending on the advertising airtime prices and the expectations of how many subscribers there will be in each platform. Firms choose the quality level of their advertising that maximizes (3).
III.	**Subscribers**, indexed by y and z, maximize (1) and (2), respectively, choosing among N media platforms according to idiosyncratic preferences regarding the programming type, the subscription prices, the advertising airtime and the average ad-quality in each platform.

The model is solved by backward induction in order to find a subgame-perfect Nash equilibrium (SPNE). All computations are relegated to an appendix. Stage III's solution is simply defined by subscribers' demand functions,

$$\hat{y}_i \equiv \hat{y}_i \left(w_i, \hat{x}_i, q_i, \mathbf{w}_{-i}, \hat{\mathbf{x}}_{-i}, \mathbf{q}_{-i} \right), \tag{5}$$

$$\hat{z}_i \equiv \hat{z}_i \left(w_i, \mathbf{w}_{-i} \right), \tag{6}$$

where $\hat{z}_i$ satisfies P.1, P.2 and P.3 and $\hat{y}_i$ satisfies P.1$'$, P.2$'$ and P.3$'$.

In stage II, assuming rational expectations on $\hat{x}_i$ and $\hat{y}_i$, the following solution emerges for each firm's problem.

$$q_x^* = \frac{\lambda}{\beta N} \tag{7}$$

Since $q_i \equiv \frac{\int_0^{\hat{x}_i} q_x dx}{\hat{x}_i}$, in a symmetric equilibrium the average ad-quality is $q_i^* = q_x^*$, for every i. Proposition 1 presents the main determinants of advertising quality in a symmetric equilibrium.

Lemma 0 *In a symmetric equilibrium, the advertising quality chosen by advertisers is increasing in the mass of sensitive consumers, λ, and decreasing in the advertising technological quality cost, β, and the number of competing platforms, N.*
Proof All proofs are in an appendix. $\square$

Lemma 0 underscores the three determinants of the advertisers' decision to spend resources in advertising quality. First, as the proportion of subscribers liable to be persuaded by ads increases, the return to persuasive advertising also increases. As a consequence, firms have more incentive to invest in ad-quality. Second, the quality increasing technology is crucial since it affects costs. Therefore, the incentive to improve quality increases with cheaper technologies. Third, under the single-homing hypothesis subscribers only watch one platform and firms only choose to advertise on one platform. Therefore, the number of media platforms determines the audience that each advertising firm reaches, thereby affecting the return from releasing better quality advertisements.

Substituting q_x for its equilibrium value in (7) into $\hat{y}_i(w_i, \hat{x}_i, q_i, \mathbf{w}_{-i}, \hat{\mathbf{x}}_{-i}, \mathbf{q}_{-i})$ the system of demand equations faced by media platform i is as follows:

$$\begin{cases} \hat{x}_i = \hat{x}_i(p_i, \hat{y}_i, \hat{z}_i, \mathbf{p}_{-i}, \hat{\mathbf{y}}_{-i}, \hat{\mathbf{z}}_{-i}) \\ \hat{y}_i = \hat{y}_i(w_i, \hat{x}_i, \mathbf{w}_{-i}, \hat{\mathbf{x}}_{-i}) \qquad \text{for } i \in \{1, 2, ..., N\}. \\ \hat{z}_i = \hat{z}_i(w_i, \mathbf{w}_{-i}) \end{cases} \tag{8}$$

Let

$$A \equiv \begin{bmatrix} \frac{\partial \hat{x}_1}{\partial \hat{y}_1} & \cdots & \frac{\partial \hat{x}_1}{\partial \hat{y}_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial \hat{x}_N}{\partial \hat{y}_1} & \cdots & \frac{\partial \hat{x}_N}{\partial \hat{y}_N} \end{bmatrix} \text{ and } B \equiv \begin{bmatrix} \frac{\partial \hat{y}_1}{\partial \hat{x}_1} & \cdots & \frac{\partial \hat{y}_1}{\partial \hat{x}_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial \hat{y}_N}{\partial \hat{x}_1} & \cdots & \frac{\partial \hat{y}_N}{\partial \hat{x}_N} \end{bmatrix}.$$

We will assume from now on that $|A| \cdot |B| \neq 1$.[20] The demand functions in system (8) can be expressed as depending exclusively on platform prices, as claimed in Lemma 1. The demand reformulation suggested in Lemma 1 is particularly useful to simplify the platforms' problem. Specifically, the reformulation allows us to write platforms' profit as a function of access prices only.

Lemma 1 *If* $|A| \cdot |B| \neq 1$ *the system of structural equations* (8) *can be re-written as*

$$\begin{cases} \hat{x}_i = \hat{x}_i(p_i, w_i, \mathbf{p}_{-i}, \mathbf{w}_{-i}) \\ \hat{y}_i = \hat{y}_i(w_i, p_i, \mathbf{w}_{-i}, \mathbf{p}_{-i}) \quad \text{for } i \in \{1, 2, ..., N\}, \\ \hat{z}_i = \hat{z}_i(w_i; \mathbf{w}_{-i}) \end{cases} \tag{9}$$

defining the reduced form of the system of structural equations faced by platform i.

We assume that the reduced form of $\hat{x}_i$ is decreasing in p_i and increasing in p_j and $\hat{y}_i$ is decreasing in w_i and increasing in w_j, analogously to properties P.2$'$ and P.2$''$ regarding the system of structural equations.[21] Lemma 2 characterizes the sign of the remaining derivatives of system (9).

Lemma 2 *The demands system* (9) *exhibits the following properties: (i) each advertising firm's demand function for platform i is decreasing in* w_i *and increasing in all components of* $\mathbf{w}_{-i}$ *and, (ii) the sensitive consumers' demand function for platform i is increasing in* p_i *and decreasing in each component of* $\mathbf{p}_{-i}$.

[20] This technical assumption is sufficient to guarantee that the demand functions can be expressed as depending only on prices.

[21] Note that without this assumption, the sign of the derivatives could be different. By constructing an example with a system of structural linear functions, we can conclude that the sign of the partial derivatives may change when the original system is re-written as a reduced form. The intuition for this possible sign disparity is as follows. In the structural form $\hat{x}_i$ is decreasing in p_i. However, the reduced form takes into account that $\hat{y}_i$ increases in p_i and $\hat{x}_i$ increases in $\hat{y}_i$. Hence, our assumption is required to impose that the former pricing effect dominates the later (the externality effect on advertisers due to having more ad-sensitive viewers on platform i).

Lemma 2 (i) states that if platform i decides to charge higher subscription prices (higher w_i) there will be a reduction in the number of advertisers in i because subscribers will switch to other platforms. On the other hand, if a platform $j \neq i$ decides to charge higher prices to subscribers (higher w_j) some of them will move from j to i. Part (ii) of Lemma 2 discusses how ad-sensitive viewers' demand for platform i varies with changes in advertisers' access prices. In particular, if platform i increases the access price to advertising firms, then fewer advertisers will place ads in platform i, making it more attractive to ad-sensitive consumers. On the other hand, if j increases its access price to advertisers, j will be less attractive to advertisers and thus more attractive to ad-sensitive subscribers.

For the sake of technical simplicity we will assume the following regarding the demand functions of system (9).

Assumption 0 (demand functions) The demand functions of system (9) present the following features: (i) constant first-order derivatives, (ii) $\frac{\partial \hat{y}_i}{\partial w_i} = \frac{\partial \hat{z}_i}{\partial w_i}$, and (iii) the own-price effect $\frac{\partial \hat{x}_i}{\partial p_i} \times \frac{\partial \hat{y}_i}{\partial w_i}$ exceeds the externality effect $\left(\frac{\partial \hat{x}_i}{\partial w_i} + \lambda \frac{\partial \hat{y}_i}{\partial p_i}\right)^2$.

Parts (i) and (iii) of Assumption 0 mean that our analysis is restricted to the case of linear demand functions with relatively low cross-group externalities. The low cross-group externality is a sufficient condition for the concavity in (p_i, w_i) of the platforms' profit function. Assumption 0 (ii) says that the impact of a change in the subscription price is identical on both the demands of ad-sensitive and ad-indifferent viewers. As a consequence, since $\frac{\partial D_i}{\partial w_i} = \lambda \frac{\partial \hat{y}_i}{\partial w_i} + (1-\lambda)\frac{\partial \hat{z}_i}{\partial w_i}$, if $\frac{\partial \hat{y}_i}{\partial w_i} = \frac{\partial \hat{z}_i}{\partial w_i}$, then $\frac{\partial D_i}{\partial w_i}$ will be independent of λ.

In stage I platforms choose simultaneously and independently the subscriber and advertiser access prices that maximize (4). The equilibrium price definition used is as follows.

Definition 1 (Nash equilibrium prices) The *Nash equilibrium prices* are an N-tuple of pairs (p_i^*, w_i^*), $i = 1, ..., N$, such that $\forall i$, (p_i^*, w_i^*) solves $\text{Max}_{\{p_i, w_i\}} \Pi_i$, as defined in (4), given $\left(\mathbf{p}_{-i}^*, \mathbf{w}_{-i}^*\right)$.

The reformulated problem of a representative media platform i is

$$\underset{\{p_i, w_i\}}{\text{Max}}\ \Pi_i = p_i \hat{x}_i + w_i \left(\lambda \hat{y}_i + (1-\lambda)\hat{z}_i\right) \quad \text{s.t.} \ \ (9).$$

From the first-order conditions of i's problem and appealing to platform pricing symmetry, the following optimality conditions emerge

$$p_i^* = \frac{w_i^* \lambda \frac{\partial \hat{y}_i}{\partial p_i} + \hat{x}_i}{-\frac{\partial \hat{x}_i}{\partial p_i}}, \tag{10}$$

$$w_i^* = \frac{p_i^* \frac{\partial \hat{x}_i}{\partial w_i} + D_i}{-\frac{\partial D_i}{\partial w_i}}. \tag{11}$$

Equation (10) describes the optimal price that platform i should charge to advertisers, while equation (11) represents the optimal price that i should charge to subscribers.

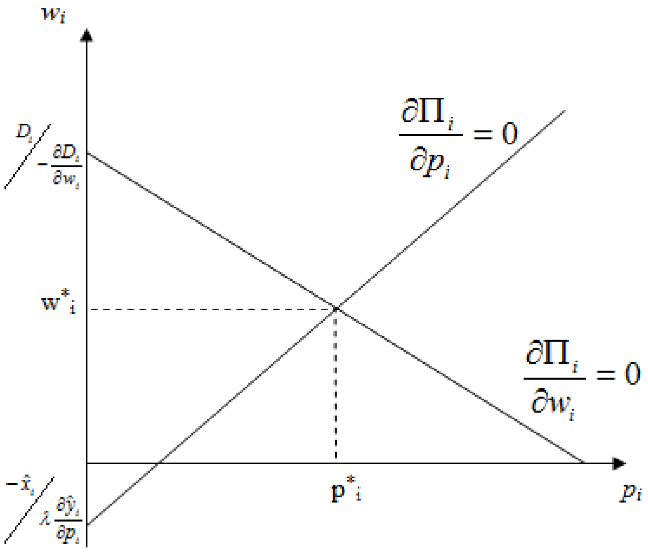


Figure 1: Representation of platform i's reaction functions with linear demand functions.

Since $\partial \hat{y}_i / \partial p_i > 0$ by Lemma 2 (ii), equation (10) reflects a positive relation between p_i and w_i. Intuitively, it means that if platform i charges more to advertisers then it will broadcast fewer ads and so should charge more to subscribers. Since $\partial \hat{x}_i / \partial w_i < 0$ by Lemma 2 (i), equation (11) reflects a negative relation between p_i and w_i. If platform i charges more to subscribers, then fewer will subscribe to it and the platform will be worth less to advertisers. At the optimum both effects are taken into account.

The following result presents the Nash equilibrium prices arising from the f.o.c.

Lemma 3 *The Nash equilibrium prices are defined by*

$$p_i^* = \frac{\dfrac{D_i}{-\frac{\partial D_i}{\partial w_i}} \lambda \frac{\partial \hat{y}_i}{\partial p_i} + \hat{x}_i}{-\frac{\partial \hat{x}_i}{\partial p_i} + \dfrac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial D_i}{\partial w_i}}}, \tag{12}$$

$$w_i^* = -\frac{D_i - \hat{x}_i . \dfrac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}}{\frac{\partial D_i}{\partial w_i} - \lambda \frac{\partial \hat{y}_i}{\partial p_i} \dfrac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}}. \tag{13}$$

Note that in equilibrium the optimal price charged to advertisers is strictly positive without further assumptions. However, the sign of w_i^* depends on the sign of the numerator, $D_i - \hat{x}_i . \frac{\partial \hat{x}_i}{\partial w_i} \Big/ \frac{\partial \hat{x}_i}{\partial p_i}$, which, in a symmetric equilibrium, has the same sign as $\frac{1}{N}\left(1 - \frac{\partial \hat{x}_i}{\partial w_i} \Big/ \frac{\partial \hat{x}_i}{\partial p_i}\right)$. In order to ensure the non-negativity of w_i^* we assume the following.

Assumption 1 (advertisers' demand) The advertisers' demand function in (9) satisfies $\partial \hat{x}_i / \partial p_i \leq \partial \hat{x}_i / \partial w_i$.

Remark 3 (Assumption 1's intuition) By definition, $\partial \hat{x}_i / \partial p_i < 0$ and by Lemma 2 (i) $\partial \hat{x}_i / \partial w_i < 0$. Hence, under Assumption 1 we have $|\partial \hat{x}_i / \partial p_i| \geq |\partial \hat{x}_i / \partial w_i|$. Intuitively, Assumption 1 (i) states that a one-dollar increase in the price charged by the platform to advertisers has a stronger negative impact on advertisers' demand than a one-dollar increase in the price charged to subscribers. A one-dollar increase in the access price charged by i to subscribers will reduce i's audience. This will reduce advertisers' demand but not as much as a one-dollar increase in the advertising price.

Assumption 1 (i) guarantees the non-negativity of w_i^* defined by (13). In most cases, negative prices are unrealistic and create perverse incentives (see Armstrong (2006) for a discussion of this issue). If negative prices were allowed, subscribers could for instance subscribe to a TV platform, never watch it and get paid for leaving the TV on. More generally, agents will have incentives, when they get paid, to make multiple purchases, behavior that platforms will hardly be able to prevent. ■

The symmetric Nash equilibrium in prices is then characterized by

$$\forall i, j,\ p_i^* = p_j^* > 0,\ w_i^* = w_j^* \geq 0.$$

This implies, together with properties P.1, P.1′ and P.1″ of symmetry and P.3, P.3′ and P.3″ of full coverage,

$$\forall i, j,\ z_i^* = z_j^* = \frac{1}{N},\ y_i^* = y_j^* = \frac{1}{N},\ x_i^* = x_j^* = \frac{1}{N}.$$

In the symmetric Nash equilibrium in prices each subscriber and advertising firm chooses the platform nearest to his preference. Each viewer chooses the platform that minimizes his disutility from not being able to watch his preferred program. Similarly, each firm chooses the platform that minimizes its profit loss from not being able to advertise on the program that best matches the product it sells. From the Nash equilibrium prices in Lemma 3 and the ad-quality symmetric equilibrium in (7) the following characterization emerges in equilibrium.

$$\begin{aligned} p_i^* &= \frac{1}{N} \frac{1 + \frac{\lambda \frac{\partial \hat{y}_i}{\partial p_i}}{-\frac{\partial D_i}{\partial w_i}}}{\frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial D_i}{\partial w_i}} - \frac{\partial \hat{x}_i}{\partial p_i}},\ i = 1, ..., N \\ w_i^* &= \frac{1}{N} \frac{1 - \frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}}{\lambda \frac{\partial \hat{y}_i}{\partial p_i} \frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}} - \frac{\partial D_i}{\partial w_i}},\ i = 1, ..., N \end{aligned}$$

$$q_x^* = \frac{\lambda}{\beta N},$$

$$
\begin{aligned}
\Pi_i^* &= \frac{1}{N^2}\left(\frac{1+\frac{\lambda\frac{\partial \hat{y}_i}{\partial p_i}}{-\frac{\partial D_i}{\partial w_i}}}{\frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial D_i}{\partial w_i}}-\frac{\partial \hat{x}_i}{\partial p_i}}+\frac{1-\frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}}{\lambda\frac{\partial \hat{y}_i}{\partial p_i}\frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}-\frac{\partial D_i}{\partial w_i}}\right),\ i=1,...,N \\
\Pi_x^* &= \frac{\phi}{N}+\left(\frac{\lambda}{\beta N}\right)^2\frac{\beta}{2}-\frac{\frac{1}{-N\frac{\partial D_i}{\partial w_i}}\lambda\frac{\partial \hat{y}_i}{\partial p_i}+\frac{1}{N}}{\frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial D_i}{\partial w_i}}-\frac{\partial \hat{x}_i}{\partial p_i}}-\min_i T\left(x,i\right),
\end{aligned}
$$

$$
\begin{aligned}
U_y^* &= v\left(1-\frac{1}{N}+\frac{\lambda}{\beta N}\right)+\frac{\frac{1}{N}\left(1-\frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}\right)}{\frac{\partial D_i}{\partial w_i}-\lambda\frac{\partial \hat{y}_i}{\partial p_i}\frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}}-\min_i T_S\left(y,i\right), \\
U_z^* &= v+\frac{\frac{1}{N}\left(1-\frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}\right)}{\frac{\partial D_i}{\partial w_i}-\lambda\frac{\partial \hat{y}_i}{\partial p_i}\frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}}-\min_i T_I\left(z,i\right).
\end{aligned}
$$

The following comparative static results on the sensitiveness of consumers to the quality of advertising are of interest.

Proposition 1 *In a symmetric Nash equilibrium, an increase in the proportion of ad-sensitive subscribers, λ, will increase advertisers' access price and decrease the subscription price.*

A media platform exhibiting more ads than its competitors will lose more subscribers, the larger the proportion of ad-sensitive viewers. It is clear from the platforms' f.o.c. on p_i that the incentive to raise the advertising price increases with the proportion of ad-sensitive subscribers. This happens because selling less advertising airtime, i.e., charging a higher p_i, expands the ad-sensitive subscribers' demand and that positive effect on the platform's profit is multiplied by the proportion of ad-sensitive subscribers.[22]

Due to the positive cross-group externality from viewers on advertisers, an increase in w_i will reduce both the number of viewers and advertisers. Note that as p_i increases, the cost of losing advertisers increases as well. Hence, if the proportion of ad-sensitive subscribers increases, platforms will choose to decrease the subscription price in order to expand the ad-sensitive audience and "sell" it at a higher price to advertising firms.

Proposition 2.1 *In a symmetric Nash equilibrium, an increase in the proportion of ad-sensitive subscribers, λ, will increase media platforms' profits if* $|\partial \hat{x}_i/\partial w_i|$ is sufficiently

[22] In Lemma A.1 in the appendix, we show that an expansion in the proportion of ad-sensitive subscribers increases the advertisers' price even if the market is partially-served insofar as advertisers are concerned.

high compared to $|\partial\hat{x}_i/\partial p_i|$.

Under symmetry and full-coverage imposed by properties P.3, P.3′ and P.3″, advertisers' and subscribers' market shares of each platform are both equal to $1/N$. Each individual platform serves, in equilibrium, the same number of subscribers and advertisers regardless of λ. However, the optimal access prices p_i^* and the subscription fees w_i^* both depend on λ. According to Proposition 1, when the proportion of ad-sensitive viewers expands, platforms will increase profits from the advertising side of the market and decrease profits from subscriptions. The net effect on platforms' profits is unclear without further assumptions. Nonetheless, if the external effect from subscribers on advertisers (measured by $|\partial\hat{x}_i/\partial w_i|$) is sufficiently high relatively to the own-price effect of advertisers (measured by $|\partial\hat{x}_i/\partial p_i|$), the platforms' main profit source will shift from subscriptions to advertisements. An example of an extreme case implying $w_i^* = 0$ is when $|\partial\hat{x}_i/\partial p_i| = |\partial\hat{x}_i/\partial w_i|$. This corresponds to a free-to-air business model, fully supported by advertising revenues. Hence, if the share of the total profits derived from the advertising side of the market is sufficiently high, an increase in the proportion of ad-sensitive subscribers will increase the media industry profits.

Proposition 2.2 *In a symmetric Nash equilibrium, an increase in the proportion of ad-sensitive subscribers, λ, will increase the indirect utility of both ad-sensitive and ad-indifferent subscribers.*

Subscribers will be better off as the proportion of ad-sensitive subscribers expands, for two reasons. First, all subscribers pay less for platform access, and, second, commercials improve in quality, thus benefiting ad-sensitive subscribers.

Proposition 2.3 *In a symmetric Nash equilibrium, an increase in the proportion of ad-sensitive subscribers, λ, will increase advertisers' profits iff $q_x^*/N > \partial p_i^*/\partial\lambda$.*

If the proportion of ad-sensitive subscribers increases, two opposite effects will hit advertisers' profits. On the one hand, Proposition 1 shows that the price of advertising airtime will increase. On the other hand, more subscribers will be liable to be persuaded and willing to pay more for the advertised goods. The net effect on advertisers' profits is ambiguous without further information. However, we can claim that if ads quality is sufficiently high (i.e., $q_x^* > N.\partial p_i^*/\partial\lambda$) then the (positive) persuasive effect will dominate the (negative) pricing effect on advertisers' profits.

The economic intuition of the condition in Proposition 2.3 is as follows. On the left-hand side of the inequality, the term q_x^*/N represents the benefit captured by firm x of an increase in the proportion λ of ad-sensitive subscribers. The quality level q_x^* represents the extra value captured by advertisers from the sales to each one of the ad-sensitive consumers. Since in equilibrium each platform serves λ/N ad-sensitive viewers, the marginal benefit of an increase in λ, for firm x, equals $q_x^* \times \partial\left(\lambda/N\right)/\partial\lambda = q_x^*/N$.

On the right-hand side of the inequality, given that the access cost p_i^* increases with the proportion of ad-sensitive viewers, the term $\partial p_i^*/\partial\lambda$ represents the marginal cost for firm x of an increase in λ. Hence, when the proportion of ad-sensitive subscribers increases, firms must present a sufficiently high advertising quality in order to cover the increase in the access cost.

4 Welfare Analysis

In this section we derive the socially-optimal advertising airtime (i.e., the mass of advertising firms) and ad-quality level and compare them to the market equilibrium conditions obtained in the SPNE. This analysis allows us to assess how well the market supplies advertising airtime and determines ad-quality. We start with the first-best analysis in which a regulator, whose objective is to maximize social welfare, defines both advertising airtime and ad-quality. In the second-best analysis, we restrict the intervention instruments to one. Specifically, the regulator maximizes social welfare choosing either ad-airtime or ad-quality.

Social welfare, W, is defined as the sum of the aggregate consumer surplus, the firms' aggregate surplus due to advertisements and the total profit earned by media platforms. Formally,

$$
\begin{aligned}
W &\equiv \sum_{i=1}^{N} \left[\begin{array}{c} \lambda \left(\int\limits_0^{\hat{y}_i} v\left(1-\hat{x}_i+q_i\right) - T_S\left(y,i\right) dy \right) + (1-\lambda) \left(\int\limits_0^{\hat{z}_i} v - T_I\left(z,i\right) dz \right) + \\ + \int\limits_0^{\hat{x}_i} \phi D_i + q_x \lambda \hat{y}_i - \beta q_x^2/2 - T\left(x,i\right) dx \end{array} \right] \\
&= \sum_{i=1}^{N} \left[\begin{array}{c} \lambda \left(v\left(1-\hat{x}_i+q_i\right)\hat{y}_i - \int\limits_0^{\hat{y}_i} T_S\left(y,i\right) dy \right) + (1-\lambda) \left(v\hat{z}_i - \int\limits_0^{\hat{z}_i} T_I\left(z,i\right) dz \right) + \\ + \left(\phi D_i + q_x \lambda \hat{y}_i - \beta q_x^2/2\right)\hat{x}_i - \int\limits_0^{\hat{x}_i} T\left(x,i\right) dx \end{array} \right]
\end{aligned} \tag{14}
$$

Property P.3$''$ is relaxed, that is, advertising firms may be only partially-served,

$$\sum_{i=1}^{N} \hat{x}_i \leq 1.$$

The case of fully-served advertising firms is part of the second-best analysis with exclusive regulation of ad-quality. Imposing full-coverage on the advertising side of the market would prevent the regulator from using advertising airtime regulation. This is why, in the welfare analysis, we treat of the case of partial-coverage as well.

We make the following assumption regarding the transportation costs throughout the rest of the analysis.

Assumption 2 (transportation costs) For each platform i, transportation costs $T_S(\hat{y}_i, i)$, $T_I(\hat{z}_i, i)$ and $T(\hat{x}_i, i)$ are differentiable and increasing in the the first argument.

Mathematically, transportation costs must satisfy

$$\frac{\partial T_S}{\partial \hat{y}_i} > 0, \ \frac{\partial T_I}{\partial \hat{z}_i} > 0 \text{ and } \frac{\partial T}{\partial \hat{x}_i} > 0, \tag{15}$$

implying that the marginal agent (viewer or advertiser) has a lower willingness to pay for the service (programming subscription or advertising slot) than the inframarginal agents. Therefore, agents subscribe to a platform in order of proximity to that platform. In other words, as the agents' demand for platform i increases, the location of the marginal agent is farther away.

4.1 First-Best Analysis: Airtime and Quality Regulation

In the first-best analysis the regulator fully controls advertising airtime and ad-quality, solving

$$\operatorname*{Max}_{\hat{x}_i, q_i} W,$$

where W was defined in (14).

In a symmetric equilibrium,[23]

$$\begin{aligned} \hat{y}_i &= \hat{z}_i = 1/N, \text{ for all } i, \text{ and} \\ \hat{x}_i &= \hat{x}_j, \ q_x = q_i = q_j \text{ for any } x, i, j, \ i \neq j, \end{aligned}$$

and the f.o.c. system can be simplified to[24]

$$\begin{cases} \frac{\lambda v}{N} + \left(\frac{\lambda}{N} - \beta q_i\right)\hat{x}_i = 0 \\ \frac{\phi - \lambda v}{N} + \frac{q_i \lambda}{N} - \beta\frac{q_i^2}{2} - T(\hat{x}_i, i) = 0 \end{cases} \Leftrightarrow \begin{cases} q_i^o = \frac{\lambda}{\beta N}\left(1 + \frac{v}{\hat{x}_i^o}\right) > \frac{\lambda}{\beta N} = q_x^* \\ \frac{\phi - \lambda v}{N} + \frac{q_i^o \lambda}{N} - \beta\frac{\left(q_i^o\right)^2}{2} - T(\hat{x}_i^o, i) = 0 \end{cases}. \tag{16}$$

From the socially optimal system of equations in (16), the optimal ad-quality q_i^o exceeds the market outcome q_i^* given in (7). The difference arises from the subscribers' benefit from ad-quality that advertising firms ignore in their private maximization process. Recall that an advertising firm is unable to influence the average ad-quality on a platform and unable to internalize all the benefits from investment in ad-quality since part of it is captured by viewers. Hence, advertisers do not invest in ad-quality up to the socially optimal level. The regulator takes into account in his maximization problem the subscribers' surplus. Therefore, the socially optimal ad quality exceeds the free market level.

[23] Computations can be found in the appendix.

[24] Assume that $\frac{\partial T(\hat{x}_i, i)}{\partial \hat{x}_i} > \frac{(\lambda v)^2}{N^2 \beta \hat{x}_i^3}$ holds. In plain words, the platforms' differentiation from the subscribers' point of view is sufficiently high. This condition allows us to satisfy the second-order conditions.

It is important to note that we take the perspective that advertising quality is adding value rather than deceiving ad-sensitive subscribers. Otherwise, eventual perverse effects that ad-quality might have on consumers' surplus could reverse the conclusion that $q_i^o > q_x^*$.

From (16) we can derive the impacts of a variation in the proportion of ad-sensitive viewers on the optimal advertising airtime and quality.

Proposition 3.1 (comparative statics on λ) *If $v \geq q_i^o$, an increase in the proportion of ad-sensitive subscribers, λ, will (i) reduce the socially optimal advertising airtime and (ii) enhance the socially optimal advertising quality.*

Parameter v measures the benefit that ad-sensitive consumers forgo by watching one more advertisement, while q_i is the social benefit arising from ad-sensitive consumers watching one additional advertisement. Therefore, if $v \geq q_i^o$, the net social benefit of one more advertisement decreases in the proportion of ad-sensitive consumers, and so does the socially optimal advertising airtime. Given $\partial \hat{x}_i^o / \partial \lambda < 0$, from the first equation in (16) we conclude that the marginal benefit of ad-quality to society increases with the proportion of ad-sensitive consumers. Moreover, as the mass of advertisers decreases, attaining a given average of ad-quality is cheaper in terms of aggregate investment.

In order to compare the social optimum, where firms may not be fully-covered, with the market equilibrium we need to compute the advertising airtime market outcome $\hat{x}_i^*$ when advertising firms are partially-covered. Advertisers enter the market until profits are zero

$$\Pi_x^* = 0 \Leftrightarrow \frac{\phi}{N} + \frac{\beta}{2}\left(\frac{\lambda}{N\beta}\right)^2 - p_i^* = T\left(\hat{x}_i^*, i\right). \tag{17}$$

We solve (17) w.r.t. $\hat{x}_i^*$ to get the advertisers' demand. Note that the optimal access price p_i^* charged by platform i was defined in (12). The condition, from (16), that characterizes the socially-optimal advertising airtime level $\hat{x}_i^o$ is

$$\frac{\phi - \lambda v}{N} + \frac{\frac{\lambda^2}{\beta N}\left(1 + \frac{v}{\hat{x}_i^o}\right)}{N} - \beta \frac{\left(\frac{\lambda}{\beta N}\left(1 + \frac{v}{\hat{x}_i^o}\right)\right)^2}{2} = T\left(\hat{x}_i^o, i\right). \tag{18}$$

In order to compare the free market advertising airtime to the socially-optimal solution, we use the left-hand sides of (17) and (18). Since T is increasing in $\hat{x}_i$ by Assumption 2, the equation with the highest LHS value will be the one with the largest ad-airtime outcome. In general the two advertising levels, $\hat{x}_i^*$ and $\hat{x}_i^o$, will differ. The following proposition formalizes the conditions under which the market ad-airtime outcome is above (below) the socially-optimal solution.

Proposition 3.2 (market outcome vs socially-optimal solution) *If the advertisers'* *access price equilibrium* p_i^* exceeds (falls short) of

$$\frac{\lambda v}{N}\left(1+\frac{\lambda v}{2\beta N\left(\hat{x}_i^o\right)^2}\right)$$

then, the market provision of advertising airtime will be below (above) the socially-optimal solution.

The social and private ad-airtime outcomes will be equal if and only if

$$p_i^* = \frac{\lambda v}{N}\left(1+\frac{\lambda v}{2\beta N\left(\hat{x}_i^o\right)^2}\right).$$

An extreme case would be when the proportion of ad-sensitive consumers is nil. In that case, $q_i^o = 0$ since quality would only generate costs without benefits, and the socially optimal mass of advertisers $\hat{x}_i^o$ would be defined by $\phi/N = T\left(\hat{x}_i^o, i\right)$, where ϕ represents the informative effect of advertising.

Corollary to Proposition 3.2 *In the absence of ad-sensitive consumers,* $\lambda = 0$*, for any* $\hat{x}_i^* > 0$*, the market provision of advertising airtime will be below the socially optimal advertising airtime.*

From the Corollary to Proposition 3.2, we claim that for a sufficiently low proportion λ of ad-sensitive consumers there is no excessive advertising airtime. With no ad-sensitive consumers the persuasive effect of advertisements is nil and advertisers only obtain revenues due to the informative effect. It is socially optimal to have advertisers entering the market up to the level that the social benefits cover the cost $T\left(\hat{x}_i^o, i\right)$. However, in the market equilibrium, platforms will charge a price $p_i^* > 0$, inflating costs to advertisers and, thus, preventing entry of advertisers with higher $T\left(\hat{x}_i, i\right)$ costs that would have entered the market in the socially optimal solution, i.e., if $p_i^* = 0$. Note that when $\lambda = 0$ there only exist ad-indifferent consumers, advertisers do not impose any negative externality on consumers, and the threshold in Proposition 3.2 becomes zero.

As the proportion of ad-sensitive consumers increases, advertisers' revenues grow due to the persuasive effect. This attracts more advertisers to the market thus increasing the negative externality of advertisements on the surplus of ad-sensitive consumers. Since advertisers do not internalize the negative effect of advertisements on consumers, an excess of advertising airtime may occur, depending on the price p_i^* charged by media platforms to advertisers. In fact, in the free-to-air broadcasting case (in appendix we develop a simple illustration with explicit solution) if the proportion of ad-sensitive subscribers is above a given threshold λ^*, then the market provision of advertising airtime will be above the socially desirable level.

4.2 The Second-Best Analysis

In the second-best analysis we restrict the intervention instruments to one. Specifically, we treat the case where the regulator maximizes social surplus choosing either ad-airtime or ad-quality while the non-regulated variable is freely determined by the market.

In this section we show first that ad-quality regulation is an implicit instrument to regulate advertising airtime. The underlying rationale is that by imposing higher ad-quality standards the regulator inflates the advertisement costs, thus restraining firms' demand for advertising slots on media platforms. Second, if the regulator decides exclusively on advertising airtime, the airtime allowed for advertisements will be higher than in the first-best solution.

Advertising quality regulation. Consider the regulator's problem of choosing an advertising quality standard that maximizes social surplus, W, while the advertising airtime of media platforms is kept unregulated. The sequence of interactions among regulator, platforms, subscribers and firms is the following.

TABLE 4: The timing of interactions and choices

I.	**Regulator** chooses an ad-quality standard.
II.	**Media platforms** choose simultaneously and independently the pair of prices (p_i, w_i).
III.	**Firms** decide on purchasing advertising airtime, among N platforms, depending on the ad-quality standard imposed, the advertising airtime price, the (rational) expectation on the number of subscribers in each platform, and the idiosyncratic preference (tranportation cost).
IV.	**Subscribers**, indexed by y (ad-sensitive) and z (ad-indifferent), choose programming among N media platforms according to idiosyncratic preferences on the programming, the subscription price charged by platforms, and the advertising airtime and average ad-quality of each platform.

Stage IV is summarized by equations (5) and (6). Since the advertisers' market is partially-served, its demand towards platform i is derived in stage III by solving $\Pi_x^* = 0$ w.r.t. $\hat{x}_i^*$ (see equation (17)). By setting the ad-quality standard, the regulator changes equation (17) and may regulate advertising airtime without doing it directly.

Proposition 3.3 *Ad-quality standard regulation may implicitly serve the purpose of regulating advertising airtime.*

From the proof of Proposition 3.3 (in appendix) we can conclude that there exists a negative relation between ad-quality standards and advertising airtime if the ad-quality standard $q_x^S > \frac{\lambda}{\beta N} = q_x^*$. If the regulator imposes ad-quality standards above the level that one would expect from the free market equilibrium, this will increase advertising costs and will restrain advertising airtime.

In stage II, equilibrium prices (p_i^*, w_i^*) are derived as the solution of platform i's profit-maximizing problem (4), whose results are provided in (12) and (13). Finally, in stage I the regulator solves

$$\begin{aligned} &\operatorname*{Max}_{q_i} W \\ \text{s.t.} \quad &\phi D_i + q_i \lambda \hat{y}_i - \beta q_i^2/2 - p_i^* - T\left(\hat{x}_i^*, i\right) = 0. \end{aligned}$$

The restriction can be seen as defining advertising airtime as an implicit function of the quality standard, $\hat{x}_i^*(q_i)$, that can be replaced in the objective function W. Then, the f.o.c. equals

$$\frac{\partial W}{\partial q_i} + \frac{\partial W}{\partial \hat{x}_i}\frac{d\hat{x}_i^*}{dq_i} = 0.$$

Solving for a symmetric equilibrium, the second-best advertising quality standard q_i^s satisfies

$$\underbrace{\frac{\lambda v}{N} + \left(\frac{\lambda}{N} - \beta q_i\right)\hat{x}_i^*\left(q_i^s\right)}_{\partial W/\partial q_i} + \underbrace{\left(\frac{\phi - \lambda v}{N} + \frac{q_i^s \lambda}{N} - \beta\frac{\left(q_i^s\right)^2}{2} - T\left(\hat{x}_i^*\left(q_i^s\right), i\right)\right)}_{\partial W/\partial \hat{x}_i}\frac{d\hat{x}_i^*\left(q_i^s\right)}{dq_i} = 0. \tag{19}$$

The derivative $d\hat{x}_i\left(q_i^s\right)/dq_i$ measures how responsive advertisers' demand is to the ad-quality standard. This derivative can also be interpreted as the weight given by the regulator to the objective of achieving the first-best optimal ad-airtime *vis-à-vis* achieving the first-best optimal ad-quality. For example, if advertisers' demand displays a very small variation given a large variation in the ad-quality standard, i.e., $d\hat{x}_i^*\left(q_i^s\right)/dq_i$ is close to zero, then the regulator will choose q_i^s very close to the first-best ad-quality level. On the contrary, if a small change in the ad-quality standard produces a large change in the mass of advertisers, then, the regulator should choose q_i^s such that it induces the advertising airtime level to be closer to its first-best level.

Despite the fact that the regulator sets the ad-quality, this instrument works simultaneously as a way of controlling ad-airtime. Nonetheless, the opposite is not true, i.e., regulating advertising airtime does not influence the choice of ad-quality by firms since each of them is infinitesimal in determining the average ad-quality.

It is noteworthy that the case of fully-served advertisers (FS) is also included in (19). Since, under full-coverage, $\hat{x}_i$ is constant and equal to $1/N$, the socially optimal ad-quality will be defined by $\partial W / \partial q_i = 0$, i.e.,

$$\frac{\lambda v}{N} + \left(\frac{\lambda}{N} - \beta q_i^{FS}\right)\frac{1}{N} = 0$$
$$\Leftrightarrow q_i^{FS} = \frac{\lambda}{\beta N}\left(1 + vN\right).$$

We can conclude that

$$q_x^* = \frac{\lambda}{\beta N} < \frac{\lambda}{\beta N}\left(1 + vN\right) = q_i^{FS} \text{ and}$$
$$q_i^{FS} = \frac{\lambda}{\beta N}\left(1 + vN\right) \leq \frac{\lambda}{\beta N}\left(1 + \frac{v}{\hat{x}_i^o}\right) = q_i^o.$$

The intuition for the previous two inequalities is as follows. First, $q_i^{FS} > q_x^*$ because, unlike advertising firms, the regulator takes into account all the benefits from the investment in ad-quality, including those captured by viewers. Second, $q_i^{FS} \leq q_i^o$ because in the first-best the advertising firms may be partially-served, $\hat{x}_i^o \leq 1/N$. The derivative of each advertiser's profit w.r.t. q_i is negative at levels of ad-quality $q_i > q_x^*$. Maintaining an average quality above q_x^*, such as q_i^o, implies a higher cost to society as the mass of advertisers on each platform increases. When deciding on ad-quality standards (above q_x^*), the regulator equates the viewers' marginal gain with the advertisers' marginal loss. More advertisers imply more weight on the marginal loss, whilst the marginal benefit of ad-quality for viewers is unchanged. Hence, the optimal ad-quality standard under advertisers full-coverage cannot be higher than the optimal ad-quality standard in the first-best.

Advertising airtime regulation. Consider now that the regulator's problem is to choose the advertising airtime x_i^s such that it maximizes the social surplus leaving the ad-quality choice to be determined by the market. The timing of interactions and choices is the same as in Table 4 except that the regulator chooses ad-airtime instead of quality in stage I, and firms decide on ad-quality in stage III. Mathematically, the regulator solves

$$\underset{\hat{x}_i}{\text{Max}}\, W$$
$$\text{s.t.} \quad q_i^* = \frac{\lambda}{N\beta}$$

and the f.o.c. of the problem evaluated in a symmetric equilibrium can be re-written as

$$\frac{\partial W}{\partial \hat{x}_i} = 0 \Leftrightarrow \frac{\phi + \left(q_i^* - v\right)\lambda}{N} - \beta\frac{\left(q_i^*\right)^2}{2} = T\left(\hat{x}_i^s, i\right). \tag{20}$$

Proposition 3.4 *The ad-airtime permitted is longer with regulation exclusively on adver-*

tising airtime than in the first-best solution, i.e., $\hat{x}_i^o < \hat{x}_i^s$.

In this second-best solution the ad-quality is set by advertising firms that maximize profits and in equilibrium $q_i^* < q_i^o$ as shown in (16). Due to the lower ad-quality chosen by firms, making commercials is now cheaper than in the first-best and the social value of an advertisement has increased compared to the first-best solution. Thus, when an ad-airtime ceiling is the only regulatory tool, more firms should be allowed to advertise on media platforms as compared to the situation in which the regulator can set both the ad-airtime and the ad-quality.

5 Conclusions

A common business model may describe several media markets, including TV, radio and the Internet. Entertaining and informative contents are the bait to get prospective purchasers of consumer goods exposed to advertisements. In this paper we have described the economics of this business model. What makes broadcasting different from other goods is that the broadcast delivers two goods, the program to subscribers and the audience to the advertising firms. Thus, it is useful to take a two-sided market's perspective.

Whenever an audience is watching an event, there is an incentive to reach them with a message. Commercial placement is increasingly being seen in movies[25] and programs and will alter the movie scripts themselves, as writers will have to write in the sponsoring products.

In this paper we have shown that an increase in the proportion of ad-sensitive subscribers drives firms to invest more in ad-quality since the persuasive effect from advertising gets stronger and is thus more profitable to advertisers. Thus, media platforms adjust this business model by decreasing the subscription price in order to expand the base of subscribers and charge higher prices to the advertising side of the market. Subscribers will be better off as the proportion of ad-sensitive subscribers increases. First, because all subscribers pay less for the subscription and, second, ad-sensitive subscribers will have advertisements of better quality. The effect on advertising firms is ambiguous, namely an increase in the proportion of ad-sensitive subscribers will only be profitable for advertising firms if the equilibrium ad-quality is sufficiently high. In the fully-served market case and under the hypothesis that advertisers' demand reacts more to advertising prices than to subscription prices, media platforms may increase their overall profits under some circumstances.

This paper also contributes to the public goods' theory in the sense that media platforms, such as TV, Internet or radio may be considered as public goods "consumed" simultaneously by subscribers and advertising firms. In our welfare analysis, when the regulator

[25]For example, the Aston Martin in James Bond's films or the Nokia cellphone in the first Matrix movie.

is able to choose both advertising levels and ad-quality, we show that the socially-desirable quality is above the market outcome. The disparity relies on subscribers benefiting from ad-quality, a fact that advertising firms ignore in their private maximization problem. Moreover, each single advertising firm is infinitesimal and thus is unable to influence the average ad-quality on platforms. Therefore, advertisers are not encouraged to invest in ad-quality up to the socially-optimal amount of ad-quality. Regarding ad-airtime, we have concluded that for a sufficiently low proportion of ad-sensitive subscribers there is always under-provision of advertisements. To see the intuition, suppose that there are no ad-sensitive subscribers. Therefore, advertisers do not impose a negative externality on consumers and only make revenues due to the informative effect. It would be socially optimal having advertisers entering the market up to the level that social benefits cover the cost $T\left(\hat{x}_i^o, i\right)$ and $p_i^* = 0$. However, platforms charge some $p_i^* > 0$ to advertisers which creates distortions on advertising decisions. For higher proportions of ad-sensitive subscribers over-provision may be observed.

In the second best analysis we argue that if the regulator only has one instrument and chooses ad-quality, it may use it as an instrument to regulate ad-airtime. For example, the regulator may impose higher ad-quality standards as a means of reducing advertising airtime. Moreover, if regulation determines advertising airtime only, the ad-airtime allowed will be longer than in the first-best solution. Intuitively, since in the second-best solution the ad-quality is chosen by advertising firms and the free market outcome is below the first-best level, making each advertisement costs less. Hence, the social surplus generated by each advertising slot is higher, driving the regulator to tolerate a higher ad-airtime level than in the first-best solution.

As a future research path we point out the welfare implications of the use of a targeting technology by media platforms in order to increase the advertising effectiveness. For example, "during a commercial break for Lost, a young couple watching TV might see an ad for the latest cellphone, while at the same time their neighbors with children may see a diaper commercial".[26] Targeted advertising allowing specific video ads to be sent to particular viewers will play an important role in digital TV. In fact, targeted advertising has already come true in Internet services such as Gmail or Google and it is expected to be extensively used in digitally broadcast TV.

6 References

Anderson, S., and J. Gabszewicz (2006), *The media and advertising: a tale of two-sided markets*, Handbook on the Economics of Art and Culture (Victor A. Ginsburg and David Throsby, eds.), North-Holland, Vol. 1, pp. 567–614.

[26] "Cable TV Puts New Spin on Ads", Wall Street Journal, May 16, 2007.

Anderson, S., and S. Coate (2005), *Market Provision of Broadcasting: A Welfare Analysis*, Review of Economic Studies, Vol. 72, No. 4, pp. 947–972.

Armstrong, M. (2006), *Competition in two-sided markets*, RAND Journal of Economics, Vol. 37, No. 3, pp. 668–691.

Bagwell, K. (August 2005), *The Economic Analysis of Advertising*, Handbook of Industrial Organization, Vol. 3.

Cable TV Puts New Spin on Ads, Wall Street Journal, May 16, 2007.

Dukes, A. (2004), *The Advertising Market in a Product Oligopoly*, Journal of Industrial Economics, Vol. 52, No. 3, pp. 327–348.

Dukes, A. (2006), *Media Concentration and Consumer Product Prices*, Economic Inquiry, Vol. 44, No. 1, pp. 128–141.

Dukes, A. and E. Gal-Or (2003), *Negotiations and Exclusivity Contracts for Advertising*, Marketing Science, Vol. 22, No. 2, pp. 222–245.

Epstein, G. (1998), *Network Competition and The Timing of Commercials*, Management Science, Vol. 44, No. 3, pp. 370–387.

Gabszewicz, J., D. Laussel, and N. Sonnac (1999), *TV-Broadcasting Competition and Advertising*, CORE Discussion Papers, Université Catholique de Louvain, Center for Operations Research and Econometrics (CORE).

Gabszewicz, J., D. Laussel, and N. Sonnac (2001), *Press advertising and the ascent of the Pensée Unique*, European Economic Review, Vol. 45(4-6), pp. 641–651.

Gabszewicz, J., D. Laussel, and N. Sonnac (2005), *Does advertising lower the price of newspapers to consumers? A theoretical appraisal* Economics Letters, Vol. 87, No. 1, pp. 127–134.

Gal-Or, E., and A. Dukes (2003), *Minimum Differentiation in Commercial Media Markets,* Journal of Economics and Management Strategy, Vol. 12, No. 3, pp. 291–325.

Gal-Or, E., M. Gal-Or, J. May, and W. Spangler (2006), *Targeted Advertising Strategies on Television*, Management Science, Vol. 52, No. 5, pp. 713–725.

Gantman, N., and O. Shy (May 2, 2007), *Broadcasting Competition and Advertising Quality: A Two-sided Market Approach.* Working paper available at http://ozshy.50webs.com/broadad15.pdf

Hotelling, H. (1929), *Stability in Competition*, Economic Journal, Vol. 39, No. 153, pp. 41–57.

World Cup 2010: World Advertising Expenditures, Translucent US $780 Billion, Karawang Business, Information, Tips and Solutions for Business and Finance, June 13, 2010 in Business News.

Kind, H., T. Nilssen, and L. Sørgard (2007), *Competition for Viewers and Advertisers in a TV Oligopoly*, Journal of Media Economics, Vol. 20, No. 3, pp. 211–233.

Liu, Q. and Konstantinos Serfes (2007), *Price Discrimination in Two-Sided Markets*, NET Institute working paper #07–25.

Masson, R., R. Mudambi, and R. Reynolds (1990), *Oligopoly in Advertiser-Supported Media*, Quarterly Review of Economics and Business, Vol. 30, No. 2, pp. 3–16.

Moorthy, S. and Zhao, H. (2000), *Advertising Spending and Perceived Quality*, Marketing Letters, Vol. 11, No. 3, pp. 221–233.

More than half the homes in U.S. have three or more TVs, The Nielsen Company, July 20, 2009.

Nelson, P. (1974b), *Advertising as Information*, Journal of Political Economy, Vol. 82, No. 4, pp. 729–54.

Rochet, J-C and J. Tirole (2003), *Platform Competition in Two-Sided Markets*, Journal of the European Economic Association, Vol. 1, No. 4, pp. 990–1029.

Rochet, J-C and J. Tirole (2006), *Two-Sided Markets: A Progress Report*, RAND Journal of Economics, Vol. 37, No. 3, pp. 645–667.

Salop, S. (1979), *Monopolistic Competition with Outside Goods*, Bell Journal of Economics, Vol. 10, No. 1, pp. 141–156.

Spence, M. and Owen, B. (1977), *Television Programming, Monopolistic Competition and Welfare*, Quarterly Journal of Economics, Vol. 91, No. 1, pp. 103–126.

Steiner, P. (1952), *Program patterns and Preferences, and the Workability of Competition in Radio Broadcasting*, Quarterly Journal of Economics, Vol. 66, No. 2, pp. 194–223.

Targeted advertising holds promise, Financial Times, August 31 2009.

Television, Internet and Mobile Usage in the U.S., Three Screen Report Vol. 7 – 4th Quarter 2009, The Nielsen Company, 2010.

What would you do with an extra hour?, Wall Street Journal, June 23, 2010.

Wilbur, K. C. (2008), *A Two-Sided, Empirical Model of Television Advertising and Viewing Markets*, Marketing Science, Vol. 27, No. 3, pp. 356–378.

Weyl, E. Glen (March 2009), *The Price Theory of Two-Sided Markets*, working paper at Harvard University.

Weyl, E. Glen (2010), *A Price Theory of Multi-sided Platforms*, American Economic Review, Vol. 100, No. 4, pp. 1642–72.

8 Hours a Day Spent on Screens, Study Finds, The New York Times, March 27, 2009.

7 Appendix

7.1 The SPNE

Stage III yields

$$\begin{aligned} \hat{y}_i &\equiv \hat{y}_i\left(w_i, \hat{x}_i, q_i, \mathbf{w}_{-i}, \hat{\mathbf{x}}_{-i}, \mathbf{q}_{-i}\right) \\ \hat{z}_i &\equiv \hat{z}_i\left(w_i, \mathbf{w}_{-i}\right). \end{aligned}$$

In stage II, an advertiser x's problem is

$$\underset{q_x}{\text{Max}}\, \Pi_x = \phi D_i + q_x \lambda \hat{y}_i - \frac{\beta q_x^2}{2} - p_i - T_x(x,i),$$

yielding the f.o.c.

$$\phi \frac{\partial D_i}{\partial q_x} + \lambda \hat{y}_i + q_x \lambda \frac{\partial \hat{y}_i}{\partial q_x} - \beta q_x = 0,$$

where $\frac{\partial D_i}{\partial q_x} = \lambda \frac{\partial \hat{y}_i}{\partial q_x}$. Hence, $q_x = \frac{\phi \frac{\partial \hat{y}_i}{\partial q_x} + E(\hat{y}_i)}{\frac{\beta}{\lambda} - \frac{\partial \hat{y}_i}{\partial q_x}}$.

Note that $\frac{\partial \hat{y}_i}{\partial q_x} = \frac{\partial^2 \hat{y}_i}{\partial q_x^2} = 0$ since $\frac{\partial \hat{y}_i}{\partial q_x} = \frac{\partial \hat{y}_i}{\partial q_i} \cdot \frac{\partial q_i}{\partial q_x}$, and $q_i \equiv \frac{\int_0^{\hat{x}_i} q_x dx}{\hat{x}_i}$ which implies $\frac{\partial q_i}{\partial q_x} = 0$ due to the infinitesimal value of advertising firms. Firm x has rational expectation, $E(\hat{y}_i)$, on $\hat{y}_i$. In a symmetric, fully-served and single-homing equilibrium, $E(\hat{y}_i) = \frac{1}{N}$. Hence, the s.o.c. equals

$$\lambda(\phi + q_x) \frac{\partial^2 \hat{y}_i}{\partial q_x^2} + 2\lambda \frac{\partial \hat{y}_i}{\partial q_x} - \beta < 0 \Leftrightarrow \beta > 0, \text{ which holds by assumption.}$$

Advertiser x's chosen quality is then,

$$q_x^* = \frac{\phi \frac{d\hat{y}_i}{dq_x} + \frac{1}{N}}{\frac{\beta}{\lambda} - \frac{d\hat{y}_i}{dq_x}} = \frac{\frac{1}{N}}{\frac{\beta}{\lambda}} = \frac{\lambda}{\beta N} > 0.$$

In stage I, platform i's problem is

$$\underset{\{p_i, w_i\}}{\text{Max}}\, \Pi_i = p_i \hat{x}_i + w_i\left(\lambda \hat{y}_i + (1-\lambda)\hat{z}_i\right) \quad \text{s.t. system (9)}$$

$$\text{f.o.c.} \quad : \quad \begin{cases} \frac{\partial \Pi_i}{\partial p_i} = 0 \Leftrightarrow \hat{x}_i + p_i \frac{\partial \hat{x}_i}{\partial p_i} + w_i \lambda \frac{\partial \hat{y}_i}{\partial p_i} = 0 \\ \frac{\partial \Pi_i}{\partial w_i} = 0 \Leftrightarrow p_i \frac{\partial \hat{x}_i}{\partial w_i} + D_i + w_i \left(\lambda \frac{\partial \hat{y}_i}{\partial w_i} + (1-\lambda) \frac{\partial \hat{z}_i}{\partial w_i} \right) = 0 \end{cases}$$

$$\begin{aligned}
\text{s.o.c.} \quad &: \\
H &= \begin{bmatrix} \begin{array}{c} 2\frac{\partial \hat{x}_i}{\partial p_i} + p_i \frac{\partial^2 \hat{x}_i}{\partial p_i^2} + \\ + w_i \lambda \frac{\partial^2 \hat{y}_i}{\partial p_i^2} \end{array} & \begin{array}{c} \frac{\partial \hat{x}_i}{\partial w_i} + p_i \frac{\partial^2 \hat{x}_i}{\partial p_i \partial w_i} + \\ + \lambda \left(\frac{\partial \hat{y}_i}{\partial p_i} + w_i \frac{\partial^2 \hat{y}_i}{\partial p_i \partial w_i} \right) \end{array} \\ \begin{array}{c} \frac{\partial \hat{x}_i}{\partial w_i} + p_i \frac{\partial^2 \hat{x}_i}{\partial p_i \partial w_i} + \\ + \lambda \left(\frac{\partial \hat{y}_i}{\partial p_i} + w_i \frac{\partial^2 \hat{y}_i}{\partial p_i \partial w_i} \right) \end{array} & \begin{array}{c} p_i \frac{\partial^2 \hat{x}_i}{\partial w_i^2} + 2 \left(\lambda \frac{\partial \hat{y}_i}{\partial w_i} + (1-\lambda) \frac{\partial \hat{z}_i}{\partial w_i} \right) + \\ + w_i \left(\lambda \frac{\partial^2 \hat{y}_i}{\partial w_i^2} + (1-\lambda) \frac{\partial^2 \hat{z}_i}{\partial w_i^2} \right) \end{array} \end{bmatrix} \\
|H_1| &= 2\frac{\partial \hat{x}_i}{\partial p_i} + p_i \frac{\partial^2 \hat{x}_i}{\partial p_i^2} + w_i \lambda \frac{\partial^2 \hat{y}_i}{\partial p_i^2} \\
|H_2| &= \left(2\frac{\partial \hat{x}_i}{\partial p_i} + p_i \frac{\partial^2 \hat{x}_i}{\partial p_i^2} + w_i \lambda \frac{\partial^2 \hat{y}_i}{\partial p_i^2} \right) \left(\begin{array}{c} p_i \frac{\partial^2 \hat{x}_i}{\partial w_i^2} + 2 \left(\lambda \frac{\partial \hat{y}_i}{\partial w_i} + (1-\lambda) \frac{\partial \hat{z}_i}{\partial w_i} \right) + \\ + w_i \left(\lambda \frac{\partial^2 \hat{y}_i}{\partial w_i^2} + (1-\lambda) \frac{\partial^2 \hat{z}_i}{\partial w_i^2} \right) \end{array} \right) + \\
&\quad - \left(\begin{array}{c} \frac{\partial \hat{x}_i}{\partial w_i} + p_i \frac{\partial^2 \hat{x}_i}{\partial p_i \partial w_i} + \\ + \lambda \left(\frac{\partial \hat{y}_i}{\partial p_i} + w_i \frac{\partial^2 \hat{y}_i}{\partial p_i \partial w_i} \right) \end{array} \right)^2 .
\end{aligned}$$

By Assumption 1,

$$\frac{\partial^2 \hat{x}_i}{\partial p_i^2} = \frac{\partial^2 \hat{x}_i}{\partial w_i^2} = \frac{\partial^2 \hat{x}_i}{\partial p_i \partial w_i} = \frac{\partial^2 \hat{y}_i}{\partial p_i^2} = \frac{\partial^2 \hat{y}_i}{\partial w_i^2} = \frac{\partial^2 \hat{y}_i}{\partial p_i \partial w_i} = \frac{\partial^2 \hat{z}_i}{\partial w_i^2} = 0 \text{ and}$$

$$\frac{\partial \hat{x}_i}{\partial p_i} \left(\lambda \frac{\partial \hat{y}_i}{\partial w_i} + (1-\lambda) \frac{\partial \hat{z}_i}{\partial w_i} \right) > \left(\frac{\partial \hat{x}_i}{\partial w_i} + \lambda \frac{\partial \hat{y}_i}{\partial p_i} \right)^2$$

we get

$$\begin{aligned}
|H_1| &= 2\frac{\partial \hat{x}_i}{\partial p_i} < 0, \\
|H_2| &= 4\frac{\partial \hat{x}_i}{\partial p_i} \left(\lambda \frac{\partial \hat{y}_i}{\partial w_i} + (1-\lambda) \frac{\partial \hat{z}_i}{\partial w_i} \right) - \left(\frac{\partial \hat{x}_i}{\partial w_i} + \lambda \frac{\partial \hat{y}_i}{\partial p_i} \right)^2 > 0.
\end{aligned}$$

From the f.o.c. system, the pricing solution equals

$$
\left\{ \begin{array}{l} p_i^* = \dfrac{w_i^* \lambda \frac{\partial \hat{y}_i}{\partial p_i} + \hat{x}_i}{-\frac{\partial \hat{x}_i}{\partial p_i}} \\ w_i^* = \dfrac{p_i^* \frac{\partial \hat{x}_i}{\partial w_i} + D_i}{-\frac{\partial D_i}{\partial w_i}} \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} p_i^* = \dfrac{\frac{D_i}{-\frac{\partial D_i}{\partial w_i}} \lambda \frac{\partial \hat{y}_i}{\partial p_i} + \hat{x}_i}{-\frac{\partial \hat{x}_i}{\partial p_i} + \frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial D}{\partial w_i}}} \\ w_i^* = -\dfrac{D_i - \hat{x}_i \cdot \frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}}{\frac{\partial D_i}{\partial w_i} - \lambda \frac{\partial \hat{y}_i}{\partial p_i} \frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}} \end{array} \right. .
$$

7.2 Welfare Analysis

7.2.1 The First-Best

$$
\begin{aligned}
\underset{q_i, \hat{x}_i}{\operatorname{Max}} W &= \sum_{i=1}^{N} \left[\begin{array}{c} \lambda \left(\int\limits_0^{\hat{y}_i} v \left(1 - \hat{x}_i + q_i\right) - T_S \left(y, i\right) dy \right) + (1-\lambda) \left(\int\limits_0^{\hat{z}_i} v - T_I \left(z, i\right) dz \right) + \\ + \int\limits_0^{\hat{x}_i} \phi D_i + q_x \lambda \hat{y}_i - \beta q_x^2/2 - T\left(x, i\right) dx \end{array} \right] \\
&= \sum_{i=1}^{N} \left[\begin{array}{c} \lambda \left(v \left(1 - \hat{x}_i + q_i\right) \hat{y}_i - \int\limits_0^{\hat{y}_i} T_S \left(y, i\right) dy \right) + (1-\lambda) \left(v \hat{z}_i - \int\limits_0^{\hat{z}_i} T_I \left(z, i\right) dz \right) + \\ + \left(\phi D_i + q_x \lambda \hat{y}_i - \beta q_x^2/2\right) \hat{x}_i - \int\limits_0^{\hat{x}_i} T\left(x, i\right) dx \end{array} \right]
\end{aligned}
$$

The f.o.c. consists of

$$
\left\{ \begin{array}{l} \frac{\partial W}{\partial q_i} = 0 \\ \frac{\partial W}{\partial \hat{x}_i} = 0 \end{array} \right.
$$

$$
\Leftrightarrow \left\{ \begin{array}{c} \lambda \left(v \left(\left(1 - \hat{x}_i + q_i\right) \frac{\partial \hat{y}_i}{\partial q_i} + \hat{y}_i \right) - T_S \left(\hat{y}_i, i\right) \frac{\partial \hat{y}_i}{\partial q_i} \right) + \\ + \left(\phi \lambda \frac{\partial \hat{y}_i}{\partial q_i} + \lambda \left(\hat{y}_i + q_i \frac{\partial \hat{y}_i}{\partial q_i} \right) - \beta q_x \right) \hat{x}_i + \\ + \sum\limits_{j \neq i} \left[\lambda \left(v \left(1 - \hat{x}_j + q_j\right) \frac{\partial \hat{y}_j}{\partial q_i} - T_S \left(\hat{y}_j, j\right) \frac{\partial \hat{y}_j}{\partial q_i} \right) + \lambda \frac{\partial \hat{y}_j}{\partial q_i} \left(\phi + q_j\right) \hat{x}_j \right] = 0 \\ \lambda \left(v \left(\frac{\partial \hat{y}_i}{\partial \hat{x}_i} \left(1 - \hat{x}_i + q_i\right) - \hat{y}_i \right) - T_S \left(\hat{y}_i, i\right) \frac{\partial \hat{y}_i}{\partial \hat{x}_i} \right) + \\ + \left(\phi \lambda \frac{\partial \hat{y}_i}{\partial \hat{x}_i} + q_x \lambda \frac{\partial \hat{y}_i}{\partial \hat{x}_i} \right) \hat{x}_i + \phi D_i + q_x \lambda \hat{y}_i - \beta q_x^2/2 - T\left(\hat{x}_i, i\right) + \\ + \sum\limits_{j \neq i} \left[\lambda \left(v \left(1 - \hat{x}_j + q_j\right) \frac{\partial \hat{y}_j}{\partial \hat{x}_i} - T_S \left(\hat{y}_j, j\right) \frac{\partial \hat{y}_j}{\partial \hat{x}_i} \right) + \lambda \frac{\partial \hat{y}_j}{\partial \hat{x}_i} \left(\phi + q_j\right) \hat{x}_j \right] = 0 \end{array} \right.
$$

Note that

$$\begin{aligned}
\frac{\partial\left(\int_0^{\hat{y}_i} T_S\left(y,i\right)dy\right)}{\partial \hat{y}_i} &= T_S\left(\hat{y}_i,i\right), \ \frac{\partial\left(\int_0^{\hat{x}_i} T\left(x,i\right)dx\right)}{\partial \hat{x}_i} = T\left(\hat{x}_i,i\right), \\
\sum_{i=1}^{N}\frac{\partial \hat{y}_i}{\partial q_j} &= 0, \text{ and } \sum_{i=1}^{N}\frac{\partial \hat{y}_i}{\partial \hat{x}_j} = 0 \text{ since } \sum_{i=1}^{N}\hat{y}_i = 1 \text{ by P.3}',
\end{aligned}$$

and in a symmetric equilibrium,

$$\begin{aligned}
\hat{y}_i &= \hat{z}_i = 1/N, \text{ for all } i, \text{ and} \\
\hat{x}_i &= \hat{x}_j,\ q_i = q_j \text{ for any } i, j,\ i \neq j.
\end{aligned}$$

Therefore, the f.o.c. system can be simplified to

$$\left\{\begin{array}{c}
\frac{\lambda v}{N} + \left(\frac{\lambda}{N} - \beta q_i\right)\hat{x}_i = 0 \\
\frac{\phi-\lambda v}{N} + \frac{q_i\lambda}{N} - \beta\frac{q_i^2}{2} - T\left(\hat{x}_i,i\right) = 0
\end{array}\right. \Leftrightarrow \left\{\begin{array}{c}
q_i^o = \frac{\lambda}{\beta N}\left(1+\frac{v}{\hat{x}_i}\right) > \frac{\lambda}{\beta N} = q_x^* \\
\frac{\phi-\lambda v}{N} + \frac{q_i\lambda}{N} - \beta\frac{q_i^2}{2} - T\left(\hat{x}_i,i\right) = 0.
\end{array}\right.$$

The s.o.c. equals

$$\begin{aligned}
H &= \begin{bmatrix} -\beta\hat{x}_i & \frac{\lambda}{N}-\beta q_i \\ \frac{\lambda}{N}-\beta q_i & -\frac{\partial T(\hat{x}_i,i)}{\partial \hat{x}_i} \end{bmatrix} \\
|H_1| &= -\beta\hat{x}_i < 0 \\
|H_2| &= \beta\hat{x}_i\frac{\partial T\left(\hat{x}_i,i\right)}{\partial \hat{x}_i} - \left(\frac{\lambda}{N}-\beta q_i\right)^2 = \\
&= \beta\hat{x}_i\frac{\partial T\left(\hat{x}_i,i\right)}{\partial \hat{x}_i} - \left(\frac{\lambda v}{N\hat{x}_i}\right)^2 > 0,
\end{aligned}$$

since

$$\beta > 0 \text{ and } \frac{\partial T\left(\hat{x}_i,i\right)}{\partial \hat{x}_i} > \frac{\left(\lambda v\right)^2}{N^2\beta\hat{x}_i^3} \text{ by assumption.}$$

7.2.2 The Second-Best: Ad-Quality Regulation

$$\begin{gathered}
\operatorname*{Max}_{q_i} W \\
\text{s.t.} \quad \phi D_i + q_x\lambda\hat{y}_i - \beta q_x^2/2 - p_i^* - T\left(\hat{x}_i,i\right) = 0
\end{gathered}$$

The f.o.c. equals

$$\frac{\partial W}{\partial q_i}+\frac{\partial W}{\partial \hat{x}_i}\frac{d\hat{x}_i}{dq_i}=0.$$

In a symmetric equilibrium,

$$\frac{\lambda v}{N}+\left(\frac{\lambda}{N}-\beta q_i\right)\hat{x}_i^*\left(q_i^s\right)+\left(\frac{\phi-\lambda v}{N}+\frac{q_i^s\lambda}{N}-\beta\frac{\left(q_i^s\right)^2}{2}-T\left(\hat{x}_i^*\left(q_i^s\right),i\right)\right)\frac{d\hat{x}_i^*\left(q_i^s\right)}{dq_i}=0.$$

The ad-quality solution from the f.o.c. is a maximizer of W if the s.o.c.

$$\frac{\partial^2 W}{\partial q_i^2}+\frac{\partial^2 W}{\partial \hat{x}_i^2}\frac{d\hat{x}_i^*}{dq_i}+\frac{\partial W}{\partial \hat{x}_i}\frac{d^2\hat{x}_i^*}{dq_i^2}<0$$

holds.

7.2.3 The Second-Best: Ad-Airtime Regulation

$$\begin{gathered}\operatorname*{Max}_{\hat{x}_i} W\\ \text{s.t.}\quad q_i^*=\frac{\lambda}{N\beta}.\end{gathered}$$

The f.o.c. equals

$$\frac{\partial W}{\partial \hat{x}_i}=0\Leftrightarrow\frac{\phi+\left(q_i^*-v\right)\lambda}{N}-\beta\frac{\left(q_i^*\right)^2}{2}-T\left(\hat{x}_i^s,i\right)=0.$$

The s.o.c. equals

$$-\frac{\partial T\left(\hat{x}_i^s,i\right)}{\partial \hat{x}_i}<0\Leftrightarrow\frac{\partial T\left(\hat{x}_i^s,i\right)}{\partial \hat{x}_i}>0,\text{ which is true by assumption in }(15).$$

7.2.4 The free-to-air broadcasting case: a simple illustration with explicit solution

This example is based on Gantman and Shy (2007). Consider the free-to-air broadcasting market with two media platforms, $N=2$, that may only collect revenues from the advertising side of the market, that is $w_i=0$. As a matter of computational simplification, ad-quality decisions are disregarded in this simple illustration.

Suppose the utility of an ad-sensitive consumer, indexed by $0\leq y\leq 1$, is defined by

$$U_y\equiv\begin{cases}v\left(1-\hat{x}_0\right)-T_S.y & \text{if subscribes platform 0}\\ v\left(1-\hat{x}_1\right)-T_S.\left(1-y\right) & \text{if subscribes platform 1,}\end{cases}\tag{21}$$

whilst the utility of an ad-indifferent consumer, indexed by $0\leq z\leq 1$, is given by

$$U_z \equiv \begin{cases} v - T_I.z & \text{if subscribes platform 0} \\ v - T_I.(1-z) & \text{if subscribes platform 1.} \end{cases} \tag{22}$$

Advertising firms are indexed by $0 \leq x \leq 1$. Let p_0 and p_1 denote, respectively, advertising prices in platform 0 and 1. Let firm x's profits of advertising be defined by

$$\Pi_x \equiv \begin{cases} \phi D_0 - p_0 - T.x & \text{if subscribes platform 0} \\ \phi D_1 - p_1 - T.(1-x) & \text{if subscribes platform 1} \\ 0 & \text{if does not advertise.} \end{cases} \tag{23}$$

Assume that both types of subscribers and advertising firms are uniformly distributed on $[0,1]$. Since platforms may only collect payments from the advertising side of the market, the profit function of platform $i = 0, 1$ is defined as

$$\Pi_i \equiv p_i \hat{x}_i.$$

Subscribers are assumed to be fully-served. Then, from (21) we have

$$\begin{aligned} v(1-\hat{x}_0) - T_S.y &= v(1-\hat{x}_1) - T_S.(1-y) \Leftrightarrow \\ \hat{y}_i &= \frac{1}{2} + \frac{v(\hat{x}_j - \hat{x}_i)}{2T_S}, \end{aligned}$$

for $i, j = 0, 1$ and $i \neq j$. Similarly, from (22) we have

$$\begin{aligned} v - T_I.z &= v - T_I.(1-z) \Leftrightarrow \\ \hat{z}_0 &= \hat{z}_1 = \frac{1}{2}. \end{aligned}$$

Hence,

$$D_i \equiv \lambda \hat{y}_i + (1-\lambda)\hat{z}_i = \lambda \left(\frac{1}{2} + \frac{v(\hat{x}_j - \hat{x}_i)}{2T_S} \right) + (1-\lambda)\frac{1}{2}, \tag{24}$$

for $i, j = 0, 1$ and $i \neq j$.

Suppose that advertising firms are partially served. From (23), the firms' demand function for advertising airtime in platform i is

$$\begin{aligned} \phi D_0 - p_0 - T.\hat{x}_i = 0 \Leftrightarrow \\ \hat{x}_i = \frac{\phi D_i - p_i}{T}, \end{aligned} \tag{25}$$

for $i = 0, 1$.

The SPNE. To compute the SPNE we plug (24) into (25) and solve the system of simultaneous equations to obtain

$$\begin{aligned}
&\begin{cases} \hat{x}_0 = \frac{\phi\left(\lambda\left(\frac{1}{2}+\frac{v(\hat{x}_1-\hat{x}_0)}{2T_S}\right)+(1-\lambda)\frac{1}{2}\right)-p_0}{T} \\ \hat{x}_1 = \frac{\phi\left(\lambda\left(\frac{1}{2}+\frac{v(\hat{x}_0-\hat{x}_1)}{2T_S}\right)+(1-\lambda)\frac{1}{2}\right)-p_1}{T} \end{cases} \\
\Leftrightarrow\ &\begin{cases} \hat{x}_0 = \frac{v\lambda\phi^2+\phi(TT_S-v\lambda(p_0+p_1))-2TT_Sp_0}{2T(TT_S+v\lambda\phi)} \\ \hat{x}_1 = \frac{v\lambda\phi^2+\phi(TT_S-v\lambda(p_0+p_1))-2TT_Sp_1}{2T(TT_S+v\lambda\phi)}. \end{cases}
\end{aligned} \tag{26}$$

Despite the assumption that the market for commercials is partially served, the market shares $\hat{x}_0\,(p_0, p_1)$ and $\hat{x}_1\,(p_1, p_0)$ depend on the advertising price of both platforms.

Finally, each platform $i = 0, 1$ solves

$$\begin{aligned} &\operatorname*{Max}_{p_i} \Pi_i = p_i \hat{x}_i \\ \text{s. t. } \hat{x}_i &= \frac{v\lambda\phi^2 + \phi\left(TT_S - v\lambda\left(p_i + p_j\right)\right) - 2TT_Sp_i}{2T\left(TT_S + v\lambda\phi\right)}. \end{aligned}$$

The f.o.c.[27] is as follows:

$$\frac{\partial \Pi_i}{\partial p_i} = 0 \Leftrightarrow \frac{v\lambda\phi^2 + \phi\left(TT_S - v\lambda\left(p_i + p_j\right)\right) - 2TT_Sp_i}{2T\left(TT_S + v\lambda\phi\right)} - p_i\frac{\phi v\lambda + 2TT_S}{2T\left(TT_S + v\lambda\phi\right)} = 0.$$

In a symmetric equilibrium, $p_0^* = p_1^*$,

$$p_0^* = p_1^* = \phi\frac{v\lambda\phi + TT_S}{3\phi v\lambda + 4TT_S}.$$

Plugging the equilibrium prices into (26), we get

$$x_0^* = x_1^* = \phi\frac{2TT_S + v\lambda\phi}{2T\left(4TT_S + 3v\lambda\phi\right)}. \tag{27}$$

Here we compare the advertising airtime in the SPNE (27) to the socially-optimal advertising airtime, i.e., the advertising airtime that maximizes social surplus W. The social surplus corresponds to the sum of aggregate subscriber surplus, advertisers' surplus, and

[27] The second derivative equals

$$\frac{\partial^2 \Pi_i}{\partial p_i^2} = 0 \Leftrightarrow -\frac{\phi v\lambda + 2TT_S}{T\left(TT_S + v\lambda\phi\right)} < 0,$$

ensuring the solution from the f.o.c. to be profit-maximizing.

platforms' profit. Formally,

$$\begin{aligned}
\underset{\hat{x}_0,\hat{x}_1}{\text{Max}}\, W &= \lambda\left[\int_0^{0.5}\left[v\left(1-\hat{x}_0\right)-T_S.y\right]dy+\int_{0.5}^{1}\left[v\left(1-\hat{x}_1\right)-T_S.\left(1-y\right)\right]dy\right] \\
&\quad+\left(1-\lambda\right)\left[\int_0^{0.5}\left[v-T_S.z\right]dz+\int_{0.5}^{1}\left[v-T_S.\left(1-z\right)\right]dz\right] \\
&\quad+\Pi_0+\Pi_1+\int_0^{\hat{x}_0}\Pi_x dx+\int_{1-\hat{x}_1}^{1}\Pi_x dx.
\end{aligned}$$

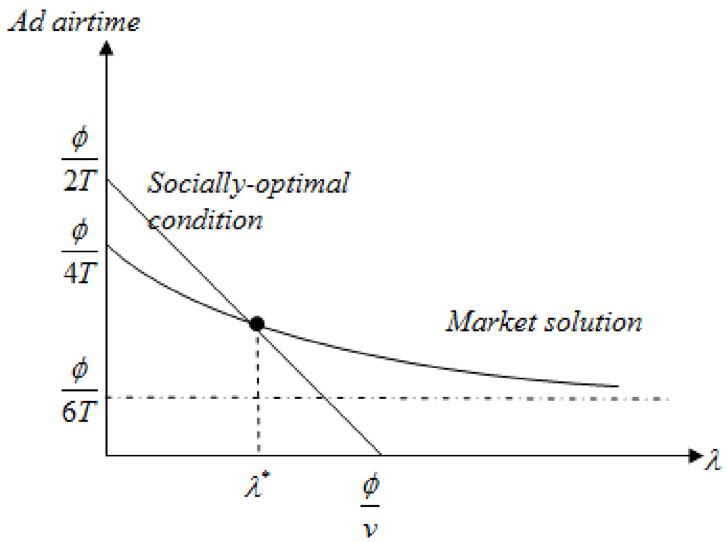
Simplifying

$$W=\frac{4v-2v\lambda\left(\hat{x}_0+\hat{x}_1\right)-T_S}{4}+\phi\left(\frac{\hat{x}_0+\hat{x}_1}{2}-\frac{v\lambda\left(\hat{x}_0-\hat{x}_1\right)^2}{2T_S}\right)-T\frac{\hat{x}_1^2+\hat{x}_0^2}{2}.$$

Computing the f.o.c. we have[28]

$$\left\{\begin{array}{l}\frac{\partial W}{\partial\hat{x}_0}=0\\ \frac{\partial W}{\partial\hat{x}_1}=0\end{array}\right.\Leftrightarrow\left\{\begin{array}{l}-\frac{v\lambda}{2}+\phi\left(\frac{1}{2}-\frac{v\lambda(\hat{x}_0-\hat{x}_1)}{T_S}\right)-T\hat{x}_0=0\\ \frac{-2v\lambda}{4}+\phi\left(\frac{1}{2}+\frac{v\lambda(\hat{x}_0-\hat{x}_1)}{T_S}\right)-T\hat{x}_1=0\end{array}\right.\Leftrightarrow\left\{\begin{array}{l}\hat{x}_0^o=\frac{\phi-v\lambda}{2T}\\ \hat{x}_1^o=\frac{\phi-v\lambda}{2T}\end{array}\right..\qquad(28)$$

We can now compare the advertising airtime in the market solution described by (27) with the socially optimal solution (28). Figure 2 performs this comparison.



[28]The second-order conditions are fulfilled since from

$$H=\left[\begin{array}{cc}-\phi\frac{v\lambda}{T_S}-T & \phi\frac{v\lambda}{T_S}\\ \phi\frac{v\lambda}{T_S} & -\phi\frac{v\lambda}{T_S}-T\end{array}\right],$$

we get

$$\begin{aligned}
|H_1| &= -\phi\frac{v\lambda}{T_S}-T<0 \\
|H_2| &= \left(\phi\frac{v\lambda}{T_S}+T\right)^2-\left(\phi\frac{v\lambda}{T_S}\right)^2>0.
\end{aligned}$$

Figure 2: Advertising airtime market outcome VS socially-optimal solution.

Note that there exists a λ^* such that $\hat{x}_i^* = \hat{x}_i^o$, for $i = 0, 1$. Lemma 4 formalizes the intuition from Figure 2 regarding the existence of λ^*.

Lemma 4 *Assume that $\phi \leq v$. There exists a unique $0 \leq \lambda^* \leq 1$ such that the market equilibrium advertising level equals the socially-optimal level.*

Proposition 4 *(i) If the proportion of ad-sensitive subscribers is above (below) λ^*, then the market provision of advertising airtime will be above (below) the socially-desirable level.*
(ii) If the proportion of ad-sensitive subscribers equals or exceeds ϕ/v, then the absence of advertising will be socially-optimal.
(iii) The advertising airtime market equilibrium is strictly positive for any proportion of ad-sensitive subscribers and converges to $\phi/6T > 0$ from above.

Note that ϕ/v may be interpreted as firms' advertising surplus, ϕ, discounted by subscribers' valuation of an ad-free platform, v. On the one hand, if the ratio decreases, this implies that the range of values of λ for which the market solution is excessive relatively to the socially-optimal advertising level will expand. On the other hand, if ϕ/v is high, firms' surplus generated by commercials exceeds subscribers' valuation of ad-free programs, implying that the advertising market provision will be too low *vis-à-vis* the socially-optimal level over a wider interval of values of λ.

It is important to stress that in the case considered in part (ii) of Proposition 4 the socially optimal advertising airtime is nil. However, this result ignores the fact that without commercials there may be no resources to support the free-to-air programs, namely in the absence of government subsidies.

Contrary to the socially-optimal advertising level, the market provision of advertising is strictly positive for any λ as stated in part (iii) of Proposition 4. In fact, for any proportion of ad-sensitive subscribers the market provision of advertisements is above $\phi/6T$.

7.3 Proofs

Proof of Lemma 0 See stage II of the SPNE in subsection 2.7.1. $\square$

Proof of Lemma 1 Using $\hat{z}_i = \hat{z}_i(w_i; \mathbf{w}_{-i})$ and

$$\hat{\mathbf{z}}_{-i} = \left(\hat{z}_1(w_1, \mathbf{w}_{-1}), ..., \hat{z}_{i-1}\left(w_{i-1}, \mathbf{w}_{-(i-1)}\right), \hat{z}_{i+1}\left(w_{i+1}, \mathbf{w}_{-(i+1)}\right), ..., \hat{z}_N(w_N, \mathbf{w}_{-N})\right),$$

substitute both in $\hat{x}_i$ to get

$$\hat{x}_i = \hat{x}_i \left(p_i, \hat{y}_i, \hat{z}_i \left(w_i, \mathbf{w}_{-i} \right), \mathbf{p}_{-i}, \hat{\mathbf{y}}_{-i}, \left(\begin{array}{c} \hat{z}_1 \left(w_1, \mathbf{w}_{-1} \right), ..., \hat{z}_{i-1} \left(w_{i-1}, \mathbf{w}_{-(i-1)} \right), \\ \hat{z}_{i+1} \left(w_{i+1}, \mathbf{w}_{-(i+1)} \right), ..., \hat{z}_N \left(w_N, \mathbf{w}_{-N} \right) \end{array} \right) \right).$$

The system of demand functions from the advertising firms is now

$$\begin{cases} \hat{x}_1 = \hat{x}_1 \left(p_1, \hat{y}_1, \hat{z}_1 \left(w_1, \mathbf{w}_{-1} \right), \mathbf{p}_{-1}, \hat{\mathbf{y}}_{-1}, \hat{\mathbf{z}}_{-1} \left(w_1, \mathbf{w}_{-1} \right) \right) \\ \hat{x}_2 = \hat{x}_2 \left(p_2, \hat{y}_2, \hat{z}_2 \left(w_2, \mathbf{w}_{-2} \right), \mathbf{p}_{-2}, \hat{\mathbf{y}}_{-2}, \hat{\mathbf{z}}_{-2} \left(w_2, \mathbf{w}_{-2} \right) \right) \\ \vdots \\ \hat{x}_i = \hat{x}_i \left(p_i, \hat{y}_i, \hat{z}_i \left(w_i, \mathbf{w}_{-i} \right), \mathbf{p}_{-i}, \hat{\mathbf{y}}_{-i}, \hat{\mathbf{z}}_{-i} \left(w_i, \mathbf{w}_{-i} \right) \right) \\ \vdots \\ \hat{x}_{N-1} = \hat{x}_{N-1} \left(p_{N-1}, \hat{y}_{N-1}, \hat{z}_{N-1} \left(w_{N-1}, \mathbf{w}_{-(N-1)} \right), \mathbf{p}_{-(N-1)}, \hat{\mathbf{y}}_{-(N-1)}, \hat{\mathbf{z}}_{-(N-1)} \left(w_{N-1}, \mathbf{w}_{-(N-1)} \right) \right) \\ \hat{x}_N = \hat{x}_N \left(p_N, \hat{y}_N, \hat{z}_N \left(w_N, \mathbf{w}_{-N} \right), \mathbf{p}_{-N}, \hat{\mathbf{y}}_{-N}, \hat{\mathbf{z}}_{-N} \left(w_N, \mathbf{w}_{-N} \right) \right). \end{cases}$$

The system of structural functions can be implicitly represented as

$$\begin{cases} \hat{x}_1 \left(\hat{\mathbf{y}}, \mathbf{p}, \mathbf{w} \right) - \hat{x}_1 = 0 \\ \vdots \\ \hat{x}_N \left(\hat{\mathbf{y}}, \mathbf{p}, \mathbf{w} \right) - \hat{x}_N = 0 \\ \hat{y}_1 \left(\hat{\mathbf{x}}, \mathbf{w} \right) - \hat{y}_1 = 0 \\ \vdots \\ \hat{y}_N \left(\hat{\mathbf{x}}, \mathbf{w} \right) - \hat{y}_N = 0 \end{cases} . \tag{29}$$

The Jacobian of the demand system (29) is

$$\begin{bmatrix} -1 & 0 & ... & 0 & \frac{\partial \hat{x}_1}{\partial \hat{y}_1} & ... & ... & \frac{\partial \hat{x}_1}{\partial \hat{y}_1} \\ 0 & \ddots & & \vdots & \vdots & & & \vdots \\ \vdots & & \ddots & 0 & \vdots & & & \vdots \\ 0 & ... & 0 & -1 & \frac{\partial \hat{x}_N}{\partial \hat{y}_1} & ... & ... & \frac{\partial \hat{x}_N}{\partial \hat{y}_N} \\ \frac{\partial \hat{y}_1}{\partial \hat{x}_1} & ... & ... & \frac{\partial \hat{y}_1}{\partial \hat{x}_1} & -1 & 0 & ... & 0 \\ \vdots & & & \vdots & 0 & \ddots & & \vdots \\ \vdots & & & \vdots & \vdots & & \ddots & 0 \\ \frac{\partial \hat{y}_N}{\partial \hat{x}_1} & ... & ... & \frac{\partial \hat{y}_N}{\partial \hat{x}_N} & 0 & ... & 0 & -1 \end{bmatrix} = \begin{bmatrix} -I_N & A \\ B & -I_N \end{bmatrix},$$

where I_N denotes the identity matrix of dimension N. Thus, the Jacobian determinant of system (29) is different from 0, since $|-I|\,|-I| - |A|\,|B| = (-1)^N \times (-1)^N - |A|\,|B| = 1 - |A|\,|B|$ and by assumption $|A|\,|B| \neq 1 \Leftrightarrow 1 - |A|\,|B| \neq 0$, and consequently the matrix is invertible.

Since $\hat{x}_i$ and $\hat{y}_i$, $i = 1, 2, .., N$ are twice differentiable by assumption, and thus are also C^1 functions, and the Jacobian of the system (29) is invertible, then, by the implicit function theorem, we can write

$$\begin{cases} \hat{x}_i = \hat{x}_i\left(p_i, w_i, \mathbf{p}_{-i}, \mathbf{w}_{-i}\right) \\ \hat{y}_i = \hat{y}_i\left(w_i, p_i, \mathbf{w}_{-i}, \mathbf{p}_{-i}\right) \quad \text{for } i \in \{1, 2, ..., N\}. \ \square \\ \hat{z}_i = \hat{z}_i\left(w_i; \mathbf{w}_{-i}\right) \end{cases}$$

Proof of Lemma 2 Assumptions,

$$\begin{aligned} \frac{\partial \hat{y}_i\left(w_i, \mathbf{w}_{-i}, p_i, \mathbf{p}_{-i}\right)}{\partial w_i} &< 0, \ \frac{\partial \hat{y}_i\left(w_i, \mathbf{w}_{-i}, p_i, \mathbf{p}_{-i}\right)}{\partial w_j} > 0, \forall i, j \text{ and } i \neq j, \\ \frac{\partial \hat{x}_i\left(p_i, \mathbf{p}_{-i}, w_i, \mathbf{w}_{-i}\right)}{\partial p_i} &< 0, \ \frac{\partial \hat{x}_i\left(p_i, \mathbf{p}_{-i}, w_i, \mathbf{w}_{-i}\right)}{\partial p_j} > 0, \forall i, j \text{ and } i \neq j. \end{aligned}$$

Hence,

$$\frac{\partial \hat{x}_i(\cdot)}{\partial w_i} = \underbrace{\frac{\partial \hat{x}_i(\cdot)}{\partial \hat{y}_i}}_{>0} \underbrace{\frac{\partial \hat{y}_i(\cdot)}{\partial w_i}}_{<0} + \underbrace{\frac{\partial \hat{x}_i(\cdot)}{\partial \hat{z}_i}}_{>0} \underbrace{\frac{\partial \hat{z}_i(\cdot)}{\partial w_i}}_{<0} + \sum_{j \neq i} \underbrace{\frac{\partial \hat{x}_i(\cdot)}{\partial \hat{y}_j}}_{<0} \underbrace{\frac{\partial \hat{y}_j(\cdot)}{\partial w_i}}_{>0} + \sum_{j \neq i} \underbrace{\frac{\partial \hat{x}_i(\cdot)}{\partial \hat{z}_j}}_{<0} \underbrace{\frac{\partial \hat{z}_j}{\partial w_i}}_{>0} < 0.$$

$$\sum_{j \neq i} \frac{\partial \hat{x}_j(\cdot)}{\partial w_i} + \frac{\partial \hat{x}_i(\cdot)}{\partial w_i} = 0 \Leftrightarrow \frac{\partial \hat{x}_j(\cdot)}{\partial w_i} = -\frac{1}{N-1} \frac{\partial \hat{x}_i(\cdot)}{\partial w_i} > 0, \text{ by P.3}''.$$

$$\frac{\partial \hat{y}_i(\cdot)}{\partial p_i} = \underbrace{\frac{\partial \hat{y}_i(\cdot)}{\partial \hat{x}_i}}_{<0} \underbrace{\frac{\partial \hat{x}_i(\cdot)}{\partial p_i}}_{<0} + \sum_{j \neq i} \underbrace{\frac{\partial \hat{y}_i(\cdot)}{\partial \hat{x}_j}}_{>0} \underbrace{\frac{\partial \hat{x}_j(\cdot)}{\partial p_i}}_{>0} > 0.$$

$$\sum_{j \neq i} \frac{\partial \hat{y}_j(\cdot)}{\partial p_i} + \frac{\partial \hat{y}_i(\cdot)}{\partial p_i} = 0 \Leftrightarrow \frac{\partial \hat{y}_j(\cdot)}{\partial p_i} = -\frac{1}{N-1} \frac{\partial \hat{y}_i(\cdot)}{\partial p_i} > 0, \text{ by P.3. } \square$$

Proof of Lemma 3 See subsection 2.7.1 “The SPNE”, stage I. □

Proof of Proposition 1

$$\frac{\partial p_i^*}{\partial \lambda} = \frac{\frac{D_i}{-\frac{\partial D_i}{\partial w_i}} \frac{\partial \hat{y}_i}{\partial p_i}}{-\frac{\partial \hat{x}_i}{\partial p_i} + \frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial D_i}{\partial w_i}}} > 0, \text{ and } \frac{\partial w_i^*}{\partial \lambda} = -\frac{D_i - \hat{x}_i \cdot \frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}}{\left(\frac{\partial D_i}{\partial w_i} - \lambda \frac{\partial \hat{y}_i}{\partial p_i} \frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}\right)^2} \cdot \frac{\partial \hat{y}_i}{\partial p_i} \frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}} < 0, \text{ by Assumption 1. } \square$$

Proof of Proposition 2.1

$$\frac{d\Pi_i^*}{d\lambda} = \frac{1}{N}\left(\frac{dp_i^*}{d\lambda}+\frac{dw_i^*}{d\lambda}\right), \text{ where}$$

$$\frac{dp_i^*}{d\lambda}+\frac{dw_i^*}{d\lambda} = \left[\frac{\frac{\frac{1}{N}}{-\frac{\partial D_i}{\partial w_i}}\left(\frac{\partial D_i}{\partial w_i}-\lambda\frac{\partial \hat{y}_i}{\partial p_i}\frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}\right)^2-\frac{1}{N}\left(1-\frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}\right)\cdot\frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}\left(-\frac{\partial \hat{x}_i}{\partial p_i}+\frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial D_i}{\partial w_i}}\right)}{\left(-\frac{\partial \hat{x}_i}{\partial p_i}+\frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial D}{\partial w_i}}\right)\left(\frac{\partial D_i}{\partial w_i}-\lambda\frac{\partial \hat{y}_i}{\partial p_i}\frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial \hat{x}_i}{\partial p_i}}\right)^2}\right]\cdot\frac{\partial \hat{y}_i}{\partial p_i}$$

and the numerator can be rewritten as

$$\underbrace{\frac{\frac{1}{N}}{-\frac{\partial D_i}{\partial w_i}\left(\frac{\partial \hat{x}_i}{\partial p_i}\right)^2}}_{>0}\left[\underbrace{\left(\frac{\partial D_i}{\partial w_i}\frac{\partial \hat{x}_i}{\partial p_i}-\lambda\frac{\partial \hat{y}_i}{\partial p_i}\frac{\partial \hat{x}_i}{\partial w_i}\right)^2}_{>0}+\underbrace{\left(\frac{\partial \hat{x}_i}{\partial p_i}-\frac{\partial \hat{x}_i}{\partial w_i}\right)}_{\substack{\leq 0\\ \text{by Assumption 1}}}\underbrace{\left(-\frac{\partial \hat{x}_i}{\partial p_i}\frac{\partial D_i}{\partial w_i}+\frac{\partial \hat{x}_i}{\partial w_i}\right)}_{<0}\underbrace{\frac{\partial \hat{x}_i}{\partial w_i}}_{<0}\right].$$

Hence,

$$\frac{d\Pi_i^*}{d\lambda}>0 \text{ if } \frac{\partial \hat{x}_i}{\partial w_i} \text{ sufficiently close to } \frac{\partial \hat{x}_i}{\partial p_i}, \text{ for example if } \frac{\partial \hat{x}_i}{\partial p_i}=\frac{\partial \hat{x}_i}{\partial w_i}. \ \square$$

Proof of Proposition 2.2

$$\frac{dU_y^*}{d\lambda}=v\frac{\partial q_x^*}{\partial \lambda}-\frac{\partial w_i^*}{\partial \lambda}>0 \text{ and } \frac{dU_z^*}{d\lambda}=-\frac{\partial w_i^*}{\partial \lambda}>0. \ \square$$

Proof of Proposition 2.3

$$\frac{d\Pi_x^*}{d\lambda}=\frac{\partial \Pi_x^*}{\partial \lambda}+\frac{\partial \Pi_x}{\partial q_x}\frac{\partial q_x^*}{\partial \lambda}+\frac{\partial \Pi_x}{\partial p_i}\frac{\partial p_i^*}{\partial \lambda}=\frac{q_x^*}{N}-\frac{\partial p_i^*}{\partial \lambda}.$$

Therefore, $\frac{d\Pi_x^*}{d\lambda}>0$ if and only if $\frac{q_x^*}{N}-\frac{\partial p_i^*}{\partial \lambda}>0 \Leftrightarrow q_x^*>\frac{\partial p_i^*}{\partial \lambda}N$. $\square$

Proof of Proposition 3.1 Let F denote the f.o.c. system in (16). Thus, by the implicit function theorem

$$
\begin{aligned}
J_{\lambda}^{q_i^o,\hat{x}_i^o}\left(q_i^o,\hat{x}_i^o,\lambda\right) &= -\left[J_{q_i,\hat{x}_i}^F\left(q_i^o,\hat{x}_i^o,\lambda\right)\right]^{-1} J_{\lambda}^F\left(q_i^o,\hat{x}_i^o,\lambda\right) \\
&= -\begin{bmatrix} -\beta\hat{x}_i^o & \frac{\lambda}{N}-\beta q_i^o \\ \frac{\lambda}{N}-\beta q_i^o & -\frac{\partial T\left(\hat{x}_i^o,i\right)}{\partial\hat{x}_i} \end{bmatrix}^{-1} \times \begin{bmatrix} \frac{v+\hat{x}_i^o}{N} \\ \frac{q_i^o-v}{N} \end{bmatrix} \\
&= \frac{1}{\beta\hat{x}_i\frac{\partial T\left(\hat{x}_i^o,i\right)}{\partial\hat{x}_i}-\left(\frac{\lambda}{N}-\beta q_i^o\right)^2}\begin{bmatrix} \frac{\partial T\left(\hat{x}_i^o,i\right)}{\partial\hat{x}_i} & \frac{\lambda}{N}-\beta q_i^o \\ \frac{\lambda}{N}-\beta q_i^o & \beta\hat{x}_i^o \end{bmatrix} \times \begin{bmatrix} \frac{v+\hat{x}_i^o}{N} \\ \frac{q_i^o-v}{N} \end{bmatrix} \\
&= \frac{1}{\beta\hat{x}_i^o\frac{\partial T\left(\hat{x}_i^o,i\right)}{\partial\hat{x}_i}-\left(\frac{\lambda}{N}-\beta q_i^o\right)^2}\begin{bmatrix} \frac{\partial T\left(\hat{x}_i^o,i\right)}{\partial\hat{x}_i}\frac{v+\hat{x}_i^o}{N}+\left(\frac{\lambda}{N}-\beta q_i^o\right)\frac{q_i^o-v}{N} \\ \left(\frac{\lambda}{N}-\beta q_i^o\right)\frac{v+\hat{x}_i^o}{N}+\beta\hat{x}_i^o\frac{q_i^o-v}{N} \end{bmatrix}.
\end{aligned}
$$

Note that $\beta\hat{x}_i^o\frac{\partial T\left(\hat{x}_i^o,i\right)}{\partial\hat{x}_i}-\left(\frac{\lambda}{N}-\beta q_i^o\right)^2>0$ by the s.o.c., $\frac{\lambda}{N}-\beta q_i^o<0$ since $q_i^o=\frac{\lambda}{\beta N}\left(1+\frac{v}{x_i^o}\right)$ by the f.o.c., and $v\geq q_i^o$ by assumption. Thus,

$$
\frac{\partial q_i^o}{\partial\lambda}=\frac{\frac{\partial T\left(\hat{x}_i^o,i\right)}{\partial\hat{x}_i}\frac{v+\hat{x}_i^o}{N}+\left(\frac{\lambda}{N}-\beta q_i^o\right)\frac{q_i^o-v}{N}}{\beta\hat{x}_i^o\frac{\partial T\left(\hat{x}_i^o,i\right)}{\partial\hat{x}_i}-\left(\frac{\lambda}{N}-\beta q_i^o\right)^2}>0,
$$

$$
\frac{\partial\hat{x}_i^o}{\partial\lambda}=\frac{\left(\frac{\lambda}{N}-\beta q_i^o\right)\frac{v+\hat{x}_i^o}{N}+\beta\hat{x}_i^o\frac{q_i^o-v}{N}}{\beta\hat{x}_i^o\frac{\partial T\left(\hat{x}_i^o,i\right)}{\partial\hat{x}_i}-\left(\frac{\lambda}{N}-\beta q_i^o\right)^2}<0.\ \square
$$

Proof of Proposition 3.2 The result comes straightforward from the LHS of equations (17) and (18), together with assumption (15) that guarantees T increasing in x_i. $\square$

Proof of Corollary to Proposition 3.2 If $\lambda=0$, then

$$
p_i^*=\frac{\frac{D_i}{-\frac{\partial D_i}{\partial w_i}}\lambda\frac{\partial\hat{y}_i}{\partial p_i}+\hat{x}_i^*}{-\frac{\partial\hat{x}_i}{\partial p_i}+\frac{\frac{\partial\hat{x}_i}{\partial w_i}}{\frac{\partial D_i}{\partial w_i}}}=\frac{\hat{x}_i^*}{-\frac{\partial\hat{x}_i}{\partial p_i}+\frac{\frac{\partial\hat{x}_i}{\partial w_i}}{\frac{\partial D_i}{\partial w_i}}}>0,\text{ for any }\hat{x}_i^*>0,
$$

and

$$
\frac{\lambda v}{N}\left(1+\frac{\lambda v}{2\beta N\left(\hat{x}_i^o\right)^2}\right)=0<p_i^*.
$$

Hence, according to the LHS of equations (17) and (18), together with assumption (15), it must be the case that $\hat{x}_i^o>\hat{x}_i^*>0$. $\square$

Proof of Proposition 3.3 In a symmetric equilibrium, the level of advertising airtime on each platform is defined by

$$
\Pi_x^*=0\Leftrightarrow\frac{\phi}{N}+q_x\frac{\lambda}{N}-\beta q_x^2/2-p_i^*-T\left(\hat{x}_i^*,i\right)=0,
$$

where q_x is the quality standard imposed by the regulator. Hence, by the implicit function theorem

$$\frac{d\hat{x}_i^*}{dq_x} = -\frac{\partial\Pi_x^*/\partial q_x}{\partial\Pi_x^*/\partial\hat{x}_i} = \frac{\frac{\lambda}{N} - \beta q_x}{\frac{\partial T\left(\hat{x}_i^*, i\right)}{\partial\hat{x}_i}} < 0,$$

since $\frac{\partial T(\hat{x}_i,i)}{\partial\hat{x}_i} > 0$ by assumption (15) and $\frac{\lambda}{N} - \beta q_x < 0$ if $q_x > \frac{\lambda}{\beta N} = q_x^*$ in (7). $\square$

Proof of Proposition 3.4 Plugging $q_i^* = \frac{\lambda}{\beta N}$ into condition (20) results in

$$\frac{\lambda^2 + 2N\beta\left(\phi - v\lambda\right)}{2N^2\beta} = T\left(\hat{x}_i^s, i\right).$$

Comparing the LHS of the previous equation to the LHS of (18), which defines the first-best advertising airtime, it is clear that $T\left(\hat{x}_i^o, i\right) - T\left(\hat{x}_i^s, i\right) = -\frac{v^2\lambda^2}{2N^2\beta\left(x_i^o\right)^2} < 0$, implying $\hat{x}_i^o < \hat{x}_i^s$ since $\partial T\left(\hat{x}_i, i\right)/\partial\hat{x}_i > 0$ by assumption in (15). $\square$

Proof of Lemma 4 Existence.

$$\begin{aligned}\hat{x}_i^o\left(\lambda\right) &= \hat{x}_i^*\left(\lambda\right) \Leftrightarrow \hat{x}_i^o\left(\lambda\right) - \hat{x}_i^*\left(\lambda\right) = 0 \\ &\Leftrightarrow \frac{\phi - v\lambda}{2T} - \phi\frac{2TT_S + v\lambda\phi}{2T\left(4TT_S + 3v\lambda\phi\right)} = 0.\end{aligned} \tag{30}$$

At $\lambda = 0$, the LHS of (30) equals

$$\frac{\phi}{2T} - \frac{\phi}{4T} = \frac{\phi}{4T} > 0,$$

while at $\lambda = 1$ the LHS of (30) equals

$$\frac{\phi - v}{2T} - \phi\frac{2TT_S + v\phi}{2T\left(4TT_S + 3v\phi\right)} < 0,$$

since $\phi \leq v$ by assumption. Hence, by the intermediate value theorem there exists at least one $\lambda^* \in (0, 1)$ such that the market equilibrium advertising level corresponds to the socially-optimal level.

Uniqueness. Solving $\hat{x}_i^o\left(\lambda\right) = \hat{x}_i^*\left(\lambda\right)$ w.r.t. λ we obtain two roots, λ_1 and λ_2,

$$\begin{aligned}\lambda_1 &= -\frac{-\phi^2 + \sqrt{\phi^4 + 4T^2T_S^2 + 2T\phi^2T_S} + 2TT_S}{3v\phi}, \\ \lambda_2 &= \frac{\phi^2 + \sqrt{\phi^4 + 4T^2T_S^2 + 2T\phi^2T_S} - 2TT_S}{3v\phi}.\end{aligned}$$

Clearly, $-\phi^2 + \sqrt{\phi^4 + 4T^2T_S^2 + 2T\phi^2T_S} + 2TT_S > 0$ since $\sqrt{\phi^4 + 4T^2T_S^2 + 2T\phi^2T_S} > \sqrt{\phi^4} = \phi^2$. Thus, we rule out the root $\lambda_1 < 0$. Only λ_2 is admissible since $\sqrt{\phi^4 + 4T^2T_S^2 + 2T\phi^2T_S} > \sqrt{4T^2T_S^2} = 2TT_S$ guarantees $\lambda_2 > 0$ and thus $\lambda^* = \lambda_2$. $\square$

Proof of Proposition 4 (i) From Lemma 4, $\hat{x}_i^o(\lambda)$ and $\hat{x}_i^*(\lambda)$ only cross once at $\lambda^* = \lambda_2$ within the interval $\lambda \in [0,1]$, $i = 0,1$. Since $\hat{x}_i^o(0) > \hat{x}_i^*(0)$ and $\hat{x}_i^o(1) < \hat{x}_i^*(1)$, it is straightforward that $\hat{x}_i^o(\lambda) > \hat{x}_i^*(\lambda)$ for $\lambda \in (0,\lambda^*)$ and $\hat{x}_i^o(\lambda) < \hat{x}_i^*(\lambda)$ for $\lambda \in (\lambda^*,1)$. (ii) If $\lambda \geq \phi/v$, then $\hat{x}_i^o(\lambda) = \text{Max}\left\{\frac{\phi - v\lambda}{2T}, 0\right\} = 0$ since $\lambda \geq \phi/v \Leftrightarrow \frac{\phi-v\lambda}{2T} \leq 0$. Therefore, no advertising is socially-optimal. (iii) Note that the advertising airtime market equilibrium $\hat{x}_i^*(\lambda) = \phi\frac{2TT_S+v\lambda\phi}{2T(4TT_S+3v\lambda\phi)}$ is strictly decreasing in λ since

$$\frac{d}{d\lambda}\left(\phi\frac{2TT_S+v\lambda\phi}{2T\left(4TT_S+3v\lambda\phi\right)}\right) = -\frac{v\phi^2T_S}{\left(4TT_S+3v\lambda\phi\right)^2} < 0,$$

and

$$\lim_{\lambda\to\infty}\hat{x}_i^*(\lambda) = \lim_{\lambda\to\infty}\phi\frac{2TT_S+v\lambda\phi}{2T\left(4TT_S+3v\lambda\phi\right)} = \frac{\phi}{6T}.$$

Hence, the advertising airtime market equilibrium converges from above to $\frac{\phi}{6T}$. $\square$

Lemma A.1 *In the market equilibrium with partially-served advertisers, the access price for advertising firms is increasing in λ.*

Proof of Lemma A.1

$$G\left(p_i^*, \hat{x}_i^*, \lambda\right) = \begin{cases} \frac{\phi}{N} + \frac{\beta}{2}\left(\frac{\lambda}{N\beta}\right)^2 - p_i^* - T\left(\hat{x}_i^*, i\right) = 0 \\ \left(-\frac{\partial \hat{x}_i}{\partial p_i} + \frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial D_i}{\partial w_i}}\right)p_i^* + \frac{D_i}{\frac{\partial D_i}{\partial w_i}}\lambda\frac{\partial \hat{y}_i}{\partial p_i} - \hat{x}_i = 0 \end{cases},$$

$$\begin{aligned} J_\lambda^{p_i^*,\hat{x}_i^*}\left(p_i^*, \hat{x}_i^*, \lambda\right) &= -\left[J^G_{p_i^*,\hat{x}_i^*}\left(p_i^*, \hat{x}_i^*, \lambda\right)\right]^{-1} J^G_\lambda\left(p_i^*, \hat{x}_i^*, \lambda\right) \\ &= -\begin{bmatrix} -1 & -\frac{\partial T\left(\hat{x}_i^*, i\right)}{\partial \hat{x}_i} \\ -\frac{\partial \hat{x}_i}{\partial p_i} + \frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial D_i}{\partial w_i}} & -1 \end{bmatrix}^{-1} \begin{bmatrix} \frac{\lambda}{N^2\beta} \\ \frac{\frac{\partial \hat{y}_i}{\partial p_i}}{N\frac{\partial D_i}{\partial w_i}} \end{bmatrix}. \end{aligned}$$

Notationaly, let

$$\begin{aligned} t &\equiv \frac{\partial T\left(\hat{x}_i^*, i\right)}{\partial \hat{x}_i} > 0, \\ a &\equiv -\frac{\partial \hat{x}_i}{\partial p_i} + \frac{\frac{\partial \hat{x}_i}{\partial w_i}}{\frac{\partial D_i}{\partial w_i}} > 0. \end{aligned}$$

Thus,

$$J_{\lambda}^{p_i^*,\hat{x}_i^*}\left(p_i^*,\hat{x}_i^*,\lambda\right) = \begin{bmatrix} \frac{1}{at+1} & -\frac{t}{at+1} \\ \frac{a}{at+1} & \frac{1}{at+1} \end{bmatrix} \begin{bmatrix} \frac{\lambda}{N^2\beta} \\ \frac{\frac{\partial \hat{y}_i}{\partial p_i}}{N\frac{\partial D_i}{\partial w_i}} \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial p_i^*}{\partial \lambda} \\ \frac{\partial \hat{x}_i^*}{\partial \lambda} \end{bmatrix} = \begin{bmatrix} \frac{1}{at+1}\frac{\lambda}{N^2\beta} - \frac{t}{at+1}\frac{\frac{\partial \hat{y}_i}{\partial p_i}}{N\frac{\partial D_i}{\partial w_i}} > 0 \\ \\ \frac{a}{at+1}\frac{\lambda}{N^2\beta} + \frac{1}{at+1}\frac{\frac{\partial \hat{y}_i}{\partial p_i}}{N\frac{\partial D_i}{\partial w_i}} \end{bmatrix}. \square$$