Information Goods and Advertising:

An Economic Model of the Internet

December, 2003

(Revision: May, 2004)

Mark Stegeman

VPI

* This paper substantially revises a 1999 manuscript, "Financing Public Goods Through Advertising." I am grateful to Simon Anderson, Randy Ellis, Hans Haller, Stephen Coate, Kevin Lang, Michael Schwarz, Claudio Mezzetti, Jim Snyder, Muhamet Yildiz, and especially John Conlon, Paul Rhode, and Mark Walker for useful discussions and comments on earlier drafts and predecessor papers. An early version of the model was presented in the UNC/Duke Theory workshop in 1991; I am grateful to the participants of that seminar and several other seminar audiences for helpful comments. Email: stegeman@vt.edu

ABSTRACT

Small firms produce information goods, which have properties of both private and nonrival goods, under conditions of constant returns to scale and free entry. The firms embed messages in their goods, selling access to the good to small consumers and message content to small advertisers. Information goods are excludable if positive access fees are feasible and includable if negative access fees are feasible; we study various cases. In equilibrium, firms could generally increase total surplus by increasing the quality of the good, supplying less advertising, and reducing access fees. They could similarly increase surplus by supplying less advertising and making a profit-compensating adjustment in access fees. Firms may over- or underproduce information goods and circumstances producing each outcome are identified. The welfare results are mostly robust to the presence of small to moderate negative externalities from advertising.

1. Introduction

The internet supports myriad firms who produce web content and myriad consumers who access it. Many of these firms are relatively small and produce close substitutes, entry occurs continually, and competition seems vigorous. Yet two features put this market outside the standard theory of competitive and monopolistically competitive markets. First, producing a small quantity of web content acts as a kind of capacity constraint, but not the standard one: even one paragraph of news can be sold to an unlimited number of persons. Second, most web sites are free, which seems consistent with intense competition, but then how do the producers survive? The conspicuous answer (aside from the self-financed vanity press) is advertising. Advertisers subsidize content to draw attention to the advertising messages embedded within it.

To study the markets for web content and related activities such as broadcasting and publishing, this paper introduces the concept of *information goods*. Information goods resemble nonrival goods in that the quantity produced (e.g., the number of news stories posted) bounds the benefit available to any given consumer but not the number of consumers who can benefit (e.g. the number who access those stories).[1] Like a private good, however, giving access to the marginal consumer (e.g., additional server capacity) may be costly. We assume that firms produce information goods under conditions of constant returns to scale and free entry. Firms nonetheless enjoy increasing returns in consumption, because production costs per consumer decline as access increases. The firms sell embedded messages to advertisers, while selling access to consumers for a fee that may be positive, negative, or zero. The combination of constant returns in production, increasing returns in consumption, the advertising subsidy, and the forced joint consumption of advertising and other content raises issues rarely addressed in market theory.[2] Yet the huge trade in information goods, which include many entertainment products, makes it important to understand these markets.

A sparse literature, beginning with Spence and Owen (1977), studies the link between the markets for advertising and the goods that it subsidizes, usually focusing on broadcasters or print media. We review this literature briefly in Section 7. These models typically adopt Steiner's (1952) discrete model of program

---

[1] For a traditional nonrival good such as a fireworks show, this distinction is the one between the length of the show and the number of persons who watch it.

[2] The web has many features that are unmodeled here. Aside from the vanity (or public-spirited) press, the internet offers certain services, such as Ebay and Google, which are subject to substantial increasing returns in production and do not fit comfortably within the present model. We also abstract from the hierarchical structure of the internet, through which portals of various sizes link down to content. Our model assumes a "flat" hierarchy, through which consumers access content directly.

choice, which assumes that each consumer uses the product of at most one broadcaster, who selects his content from a finite set of alternatives. This paper departs from that standard framework by assuming that firms, advertisers, and consumers are so numerous that strategic interactions can be ignored. From a technical standpoint, this allows us to study a continuous model in which all agents are small and demand and supply are differentiable. Studying a continuous model has three major advantages.

First, eliminating strategic interactions allows us to focus on the pure consequences of linking the markets for advertising and information goods. The case of small agents thus takes its usual role as a benchmark. For example, are the linked markets efficient? Or does the expression "The Worldwide Waste of Time" capture a genuine problem? By assuming that all agents are small, we give efficiency its best chance to appear.

Second, the continuous model is tractable. Differentiability makes it easy to adopt general functional forms and – unlike most previous studies – to treat all prices and quantities in the two markets as endogenous. Differentiability also produces sharp and intuitive conclusions. Discrete models of the advertising subsidy have often exhibited a high ratio of assumptions to results.

The third reason to study a continuous model is that it is increasingly realistic. Steiner (1952) assumes the presence of three broadcasters, but consumers now have scores of choices in many television markets and on the internet. The agents in these markets are not infinitesimal, but they are numerous enough to suggest that a continuous model can provide useful insights into the real economic landscape.[3]

An important modeling issue is: which access fees are technically feasible? Is most web content free because it is expensive to administer non-zero prices or because the equilibrium price is naturally close to zero? Are negative prices an interesting possibility? We remain agnostic on these issues by considering all four permutations, that information goods may or may not be *excludable*, meaning that positive fees are feasible, and may or may not be *includable*, meaning that negative fees are feasible. The results for the various cases overlap considerably. The case of *unpriced* goods – goods that are neither excludable nor includable – is the hardest to assess, because a missing price provides no information about marginal rates of substitution. Our results for unpriced goods depend largely upon whether the zero price constraint binds from above or below, and we suggest empirical strategies for answering that question.

---

[3] In television, an important blemish on this increasingly competitive landscape (in the United States) is that the failure to regulate cable and satellite providers as common carriers has given these intermediaries substantial market power. In radio, a long sequence of mergers has produced enough concentration to make strategic interaction important. In the context of the internet, however, production and consumption are fragmented and the market power of intermediaries such as AOL and Yahoo is at present quite limited.

The model produces positive and normative results. This paper describes the equilibrium, interprets it, and states mostly normative results. We show that, in equilibrium, firms *set access fees too high* at the margin, relative to what would maximize total surplus. This is true even if the equilibrium fee is negative. Firms also put *too little quality* and embed *too much advertising* into information goods. In this sense, the "Worldwide Waste of Time" proves to be correct. The main exception is that if nonincludability binds, then firms may produce information goods containing too much quality or too little advertising, as they try to compensate for their inability to pay consumers to access their products.

By collectively reducing the supply of advertising, firms can often increase their own profits as well as total surplus (if they can also block new entry). This is an instance of collusion improving efficiency. Such collusion against advertisers is not always profitable, but we show that firms can generally increase profits and surplus by reducing the supply of advertising and simultaneously raising access fees, if such increases are technically feasible. This shows, in a strong sense, that firms rely too much on advertising as a source of revenue.

Unlike the sharp results concerning fees, advertising, and quality, the advertising subsidy can cause firms to over- or underproduce information goods. We describe circumstances implying each. In general, overproduction is more likely if: consumers' aggregate demand for information goods is relatively inelastic, they would pay relatively little to specialize their consumption further into their favorite kinds of goods, and nonincludability is a binding constraint. Section 5 states these welfare results and provides intuition.

Previous studies often refer to "externalities" that advertisements impose on the consumers who see them. Because we assume that consumers account for the effects of advertising when making their access decisions, we treat such effects as internal to the consumer's decision and not externalities.[4] Advertising may, however, create non-pecuniary spillovers for agents other than the advertiser demanding it, the firm supplying it, and the consumer accessing it. For example, seeing liquor ads may promote behavior that hurts persons who never saw the ads. Another example is that ads may take sales and profits away from the advertiser's rivals; this effect appears in many models of advertising and underpins many informal discussions of "wasteful" or mutually cancelling advertising. We show that such negative externalities have little impact on our results, unless they reach magnitudes that seem economically implausible. (A partial exception is that negative externalities increase the likelihood that information goods are overproduced, especially when firms derive most of their revenue from advertising.)

[4] Endogenizing the decision to consume advertising provides an answer to Fisher's and McGowan's (1979) criticism of using demand shifts to measure the apparently unmeasurable changes in utility that may be induced by persuasive advertising. When the decision to consume advertising is endogenous, any such benefits of consuming advertising are internalized.

Alternatively, advertising may generate positive spillovers, as when it provides useful information to friends of the consumer who saw the ad. Net positive externalities could overturn our conclusion that firms advertise too much but would reinforce our findings of excessively low quality and high access fees.

Sections 2 through 5 comprise the heart of the paper. Section 2 describes the model, Section 3 derives equilibrium conditions, Section 4 defines the welfare measure, and Section 5 states all of the welfare results. Section 6 adds detail by describing a particular model of the demand for advertising. Finally, Section 7 discusses related literature and Section 8 offers concluding remarks. An appendix collects the proofs.

2. The model.

We study *information goods*, meaning that each unit of output can be consumed by any number of consumers but only once by each consumer. The idea is that giving a fixed bit of information to a second consumer creates new surplus, but giving it to the same consumer a second time does not (presuming that she stored or remembered it the first time).[5] Many entertainment goods have essentially the same quality.

We measure information goods in units of *images*. An image could be a page of text, a photograph, a musical recording, or a website. We take the consumption decision to be binary: a consumer chooses only whether to get an image, not how much she uses it.[6] Because the model is static, it implicitly describes the production and consumption of images during some fixed period of time. If that period is one year, then an image could represent the entire content of one website during that year.

The model has three kinds of agents: firms, consumers, and advertisers. The firms produce images containing embedded advertising messages. Embedding means that a consumer cannot disentangle the messages from the image: she must consume them together. If consumers filter out some messages, then our

---

[5] Because information, once obtained, may give the consumer a flow of services, our use of the word "consume" may conceal what is really an investment decision.

[6] This model adopts a simple view of complicated "images" such as the New York Times website. By taking the decision to consume this website to be binary, we suppress the considerable question of which bits of it the consumer actually chooses to use. One interpretation is that the decision to consume is a decision to consume some fixed fraction, say 20%, of the website. Alternatively, we can disaggregate the website and interpret each article or page of text to be a different image, but that ignores the service provided by linking these images together in a useful way. The disaggregated interpretation also disregards the empirical fact that the images may be bundled and sold as a package (e.g. the annual fee required to access the Wall Street Journal website). In short, the market for information goods is complicated and we focus only on the key feature that such goods may be "sold" to many consumers but not twice to one consumer, while abstracting from many (probably important) details.

model is sensible if filtering is incomplete and all agents account for it. (If consumers filter out all messages then advertising should have disappeared by now.)

In one market the firms sell images to consumers, for a fee f. A consumer must pay f to *access*, meaning consume, one image. In a separate market the firms sell advertising. Specifically, the firms sell the right to choose message content to advertisers, for a price p per *impression*: one message accessed by one consumer. This reflects pricing in real advertising markets, where, abstracting from differences among consumers, the advertiser pays roughly in proportion to the number of consumers who see the message. Our assumptions will ensure that $p>0$ in equilibrium. Because impressions are a private good, we are modeling the joint production of private and information goods.[7]

We consider several versions of the model, according to which access fees f are technically feasible. A good is *excludable* if all $f>0$ are feasible or *nonexcludable* if no $f>0$ is feasible. A good is *includable* if all $f<0$ are feasible or *nonincludable* if no $f<0$ is feasible. We assume that $f=0$ is always feasible. A magazine is typically excludable but nonincludable, because it is impractical to monitor whether a consumer reads it. An open-air display may be includable but nonexcludable, if it is practical to monitor whether a consumer looks at it. In reality, what is excludable and includable is not sharply defined, but it seems reasonable to think of internet goods as mostly excludable – and perhaps includable. For other information goods, such as radio broadcasts, the set of feasible fees may differ. By studying various cases, we allow the reader to use her own judgment concerning which fees are feasible in various contexts. (In some cases non-zero access fees may merely await improvements in the technology for implementing such fees.) Goods that are nonexcludable and nonincludable are *unpriced*; in that case f is constrained to $f=0$.[8]

Having introduced the two prices of the model (f and p), we summarize all seven endogenous variables in the following anticipatory table. The rest of this section describes firms, advertisers, consumers, the remaining five variables, and the sequence of actions.

---

[7] Because of the issue raised in the first paragraph of the paper, it seems misleading to describe information goods as ordinary private goods. We discuss the relationship between private goods, public goods, and information goods in Section 2F.

[8] We say "unpriced" rather than "free" because in market theory "free" customarily refers to an an endogenous outcome rather than a constraint.

Variable	Range	Name	Interpretation	Stage
p	$p>0$	price of advertising	price of one impression	2
q	$q \geq 0$	consumption	quantity of images accessed by one consumer	2
s	$s \in [0,1]$	access rate	fraction of consumers who access a given image	2
c	$c \geq \underline{c} > 0$	quality	cost of producing one image	1
f	$f \in \mathbb{R}$	access fee	what one consumer pays to access one image	1
m	$m \geq 0$	message density	number of advertising messages in each image	1
x	$x \geq 0$	output	total images produced	1

A. Firms.

To gain an initial understanding of the role of the firms, assume for now that each firm produces exactly one image. It is characteristic of information goods that they are differentiated: that is why a consumer can get more utility from consuming two different images than from consuming the same image twice (see also fn. 10). We represent this differentiation by assuming that the firms are arranged symmetrically around the unit circle. A firm's location on the circle identifies the horizontal attribute of its image. (For simplicity, we take each firm's horizontal attribute to be exogenous.) If images are information, then different points on the circle might represent news, biography, recipes, etc. If images are entertainment, then different points might represent movies, songs, photographs, etc.[9] Firms are otherwise identical ex ante. The points on the circle are labeled in the obvious fashion from $t=-½$ to $t=½$, where $t=-½$ and $t=½$ represent the same point, that which is furthest from the point $t=0$. To avoid asymmetries in notation the label $t=0$ is not permanently fixed to any point or firm. It is assigned in context and all other points are labeled relative to that point.

Each firm makes three decisions, choosing its *message density* $\hat{m}$, its *quality* $\hat{c} \geq \underline{c}$, where $\underline{c}>0$ is an exogenous lower bound, and its *access fee* $\hat{f} \in \mathbb{R}$ (which may be subject to constraints, as already described). The message density $\hat{m}$ represents the number of advertising messages embedded in the single image that the firm produces. A carat always indicates a choice or outcome for an individual firm, where this might

[9] In the context of broadcasting, the time of day is an important horizontal attribute, partly because it may be difficult to consume different images which are broadcast at the same time. This illustrates the general point that in reality information goods are differentiated along more than one dimension. Among previous studies of the advertising subsidy, Steiner (1952) appears to be unique in accounting for several dimensions of horizontal differentiation.

otherwise be confused with an outcome that is common across firms. The quality $, measured in money units, represents the firm's investment in its image. We interpret $ to include the cost of producing the messages m$. Advertisers merely choose the content of those messages. We will assume that, all else equal, consumers prefer images of higher quality.

A firm's decisions, given other firms' decisions and consumers' preferences, determine its *access ratio*, $0[0,1], the fraction of the large population of consumers that accesses its image. The total measure of consumers is normalized to one. Therefore, the total consumption of a firm's image is 1×$=$. Each firm seeks to maximize its profit:

(1) B($,m$,f$,$;p) / m$$p + (f$! P)$! $,

where P$0 denotes the exogenous marginal cost of providing access to one consumer. For websites, P>0 may represent the small cost of running a server; for a print newspaper, P>0 may represent the relatively large cost of printing one issue of the paper; for a radio broadcast, it may be reasonable to assume that P=0.

The firm's profit (1) has three components. The term ! $ is the firm's cost of producing its (single) image. The term m$$p is advertising revenue: the firm sells m$ messages, yielding m$$ impressions, at a market price of p per impression. We assume that p is set in a competitive market for impressions and thus exogenous to the firm. The term $[f$! P] is fee revenue, net of access costs: the firm sells access to its image to $ consumers, for a fee of f$ against a marginal cost of P.

B. Continuous differentiation.

We assume, in general, that firms can produce fractional images under conditions of constant returns to scale. A firm that produces $0ú+ images earns profits equal to $B($,m$,f$,$;p), where m$ now represents messages per image, $ represents production cost per image, and f$ represents the access fee per image. In reality, constant returns to scale rarely obtain exactly, but constant returns may be a reasonably good representation of the costs of producing many kinds of information goods. It is plausible that doubling the size of an online newspaper approximately doubles the cost of production; moreover, it is possible to produce small (i.e., fractional) quantities of news.

Assuming constant returns to scale allows us to assume that firms are numerous. Specifically, we assume that each firm produces such a small quantity of images (equivalently, one image represents such a small quantity of output) that removing the firm would have negligible effects on everyone else in the model. In particular, no one firm has any appreciable impact on any other firm's profit, any consumer's utility, or the price of advertising impressions. Technically, this means that the firms are packed so densely around the circle that we can treat them as continuous: we suppose that every point on the circle is occupied by one small firm, which produces a negligible quantity of images (equivalently, one image of negligible measure), and there are no horizontal gaps in the kinds of images produced. Continuity makes the model tractable and allows us to focus on the pure consequences of linking the markets for information goods and advertising, where we have made those markets as competitive as they can possibly be.[10]

Assigning unit measure to the full circle of firms, and pinning the label t=0 to an arbitrary point, total output is

$$x \equiv \int_{-1/2}^{1/2} x(t)dt,$$

where x(t) denotes the density of output at point t (i.e. the number of images per unit interval in the neighborhood of t).

Continuous differentiation implies that each firm competes against firms offering arbitrarily close substitutes, but this does not imply that access fees are set competitively, because firms producing close substitutes also produce few total images. This inevitably creates market power in the market for access, even if firms are infinitesimal. A firm that lowers its access fee can attract a few more consumers, but only a few, because (as we shall see) most of the consumers using the products of nearby firms are already also using its own product. If firms were not small, then market power would arise from more conventional origins; the point is that, in the market for information goods, market power does not vanish in the limit as total output is divided among smaller and smaller firms that produce increasingly subtle variations of an essentially fixed universe of products (e.g., the unit circle).

Finally, constant returns to scale allow us to assume that free entry drives profits to zero exactly. If discrete firms are already tightly packed around the circle, then new entry does not significantly increase the variety of images available, but if total output increases then it "deepens" that variety, by packing more images into any given interval. For example, doubling total output might imply that a consumer can download twice as many meatloaf recipes.

In principle, a firm should also be able to choose how many images to it produce. Our assumptions of constant returns to scale and free entry imply (as usual) that a firm gains no advantage from size, and so

---

[10] In principle, multiple firms could locate at the same point on the circle and produce identical images, for example the day's baseball scores. The formation of anything resembling a competitive market for a single image seems unlikely, though, for several reasons. First, because each consumer uses each image at most once, the market is in some sense small. Second, increasing returns to scale in consumption, evident in profit (1) increasing more than proportionately in β, implies that the market for a single image is a natural monopoly. Third, constant returns to scale in production implies that near-continuous product differentiation, which is advantageous for firms, is technically efficient.

we suppress the firm's quantity decision by assuming that firms are small as described above. The free entry condition will determine x, the total quantity of images produced in equilibrium.

C. Advertisers.

Advertisers are small and modeled only as an exogenous demand for advertising, measured in impressions. To lighten notation, (2) assumes that the message density m and the access ratio s are common across firms, so that total impressions equal xms. Then the inverse demand for advertising is:

$$(2) \qquad p = \alpha(xms)$$

where $\alpha:\mathbb{R}_+ \to \mathbb{R}_{++}$ is differentiable with $\alpha' < 0$. The suppliers in the market for advertising are the firms modeled in the last section, and the market supply curve emerges implicitly from their optimization problem.[11] The advertising market is unusual in that the quantity supplied can be positive at p<0: if consumers prefer to consume at least a small quantity of advertising messages, as we will allow, then firms might be willing to pay advertisers for message content. At p<0, however, it seems likely that advertisers would demand an unlimited quantity of impressions. Therefore, it seems reasonable and also simplifies matters to assume that $\alpha>0$, implying that p>0 in equilibrium.

This model treats advertising as a "black box": the content of advertisers' messages and the nature of their impact is unspecified. Advertising could be persuasive or informative. It could promote products, charities, or candidates, include public service announcements, or contain any other information that the advertiser wants consumers to absorb. Section 6 describes a specific model of persuasive advertising, which produces a demand curve of form $\alpha(xms) = (xms)^{-(1-\epsilon)}$, where $\epsilon \in (0,1)$ is a consumer preference parameter.

We assume, as the result of advertiser optimization, that $\alpha(xms)$ measures the value of the marginal impression to the advertiser. Integrating under the advertisers' demand curve, their total surplus is

$$(3) \qquad \int_0^{xms}[\alpha(k) - p]dk.$$

Whether the area under the demand curve is a reliable measure of advertisers' surplus is a minor issue here, because we study the model only in the neighborhood of equilibria.

D. Consumers.

The consumers are more complicated, because they are differentiated by tastes and we must put enough structure on those tastes to indicate how consumers would respond to individual firms' defections

[11] Stegeman (2002) provides an explicit graphical derivation of the supply curve.

from equilibrium behavior. Consumers are distributed uniformly around the same unit circle as are the firms and have total measure one. A consumer located at an arbitrarily chosen point most prefers (other things equal) a good having the same horizontal attribute.

To describe the preferences of a consumer j located at an arbitrary point t=0, suppose that the points on the circle are labeled from $t=-½$ to $t=½$, as described in Section 2A. Let $A_j \subseteq [-½,½]$ denote the (measurable) set of images that consumer j elects to access, and let

(4) $$q_j \equiv \int_{A_j} x(t)dt$$

denote her total consumption, where x(t) denotes firm t's output density. Incorporating the possibility that firms at different points on the circle choose different values of (c,m,f), we posit consumer j's utility to be:

(5) $$U_j = \int_{A_j} x(t)[\beta(c(t),m(t)) - \gamma(|t|) - f(t)]dt + \mu(q_j,\bar{m}_j) - \varphi(q_j) ,$$

where c(t), m(t), and f(t) denote firm t's choices of (c,m,f), $\bar{m}_j \equiv \int_{A_j} x(t)m(t)dt/q_j$ denotes the average density of messages embedded in the images in A_j, and β, γ, μ, and φ are exogenous twice-differentiable functions.[12] We now discuss those functions.

The expression $\beta(c(t),m(t)) - \gamma(|t|)$, where $\beta:\mathbb{R}_+\times\mathbb{R}_+\rightarrow\mathbb{R}$ and $\gamma:[0,½]\rightarrow\mathbb{R}_+$, represents the benefit that consumer j derives, per image, from accessing the goods produced by a firm at distance t, before deducting the access fee f(t). Note that $\beta<0$ is possible, and it will become clear that $\beta<0$ can occur in equilibrium if $f<0$. We assume that $\beta_1>0$, meaning that consumers prefer higher quality images. (Otherwise, a firm producing high quality would have wasted its money.) Because consumers may plausibly get positive or negative benefits from the consumption of advertising, we assume nothing about the sign of β_2.[13] The function γ describes the utility penalty, $\gamma|t|$, associated with consuming a good at "distance" $|t|$ from the consumer's point on the circle. Assume that $\gamma(0)=0$ and $\gamma'>0$. Let $\delta:\mathbb{R}\rightarrow[0,½]$ denote the inverse of γ, with the convention that $\delta=0$ if its argument is negative and $\delta=½$ if its argument exceeds $\gamma(½)$.

---

[12] As a check on the internal consistency of the model, note that (3) implies that the advertisers' total payments equal xmsp, which in a symmetric equilibrium equals firms' total advertising revenues, as computed by integrating unit profits (1) over the set of firms. Similarly, integrating (5) over the unit measure of consumers shows that, in a symmetric equilibrium, total access fees equal fxs; integrating (1) over the set of firms shows that total fees received also equal fxs.

[13] It is possible that $\beta_2<0$ even if consumers derive positive "intrinsic" benefits from advertising messages, because higher message density may crowd out more desirable qualities of the good, holding c constant. For example, a consumer might value a two-inch ad in an online newspaper article (i.e., she would not prefer to see white space), but she might get more value if equal resources had instead been invested in producing two inches of additional news content. In that case, $\beta_2<0$.

The differentiable function :: ú+6ú+ accounts for whatever benefits consumer j derives from consuming advertising, which are independent of her consumption of images and hence not captured by $; its argument $q_j m)_j$ equals the total quantity of messages that she accesses. The benefits described by : may include the later consequences of seeing advertisements, such as getting higher utility from advertised products or learning of such products. (Recall that the advertised products are outside the present model.) Most extant models of advertising implicitly require :…0 (i.e., they require : to be non-linear, because linear : could be incorporated into $). Assume that :(0)=0 and :N$0, implying that the consumer experiences any negative impacts of advertising through $, and :O#0. For the specific model of advertising studied in Section 6, : takes the form :(q_jm) = (q_jm)^0, where 00(0,1) is exogenous.

Finally, the differentiable function 6: ú+6ú+ describes the opportunity cost of the time spent accessing firms' images. In some contexts, such as broadcasting, q_j (and its twin quantity x) might actually be measured in units of time. For other information goods, a sensible interpretation of 6 requires only that the time spent consuming images is strictly increasing in the number of images consumed. Assume 6(0)=0, 6N>0, and 6O>0.

The quasi-linear utility function (5) assumes considerable separability, which may not be realistic. For instance, interaction between $ and (is plausible: a consumer might care either more or less about vertical quality, as horizontal quality increases. We impose separability mainly to increase the tractability of the model. The separable case also provides a useful benchmark, especially as it is unclear which signs one would expect on various cross-partial derivatives if separability were not forcing those derivatives to zero.

E. Timing.

This cast of agents performs in (on) two stages. In stage one, each firm chooses its image attributes (m̂,$̂) and the fee f̂ that it charges for access, if f̂…0 is feasible. The free entry (i.e., zero profit) condition determines total output x. In stage two, each consumer observes the firms' decisions and chooses which set of images to access (the set A_j in (5)), and the market for advertising clears at price p. No game is played in stage two, because what is optimal for one consumer is independent of other consumers' actions. In stage one, we study a standard Nash equilibrium among the firms, assuming perfect anticipation of the stage two consequences of firms' play. The Nash equilibrium takes a simple form, however, because there is no strategic interaction. Finally, note that nothing in the model is random.

F. The connections between information goods, private goods, and public goods.

This section, which the reader may skip without loss of continuity, discusses how information goods are related to standard notions of private and public goods. Consider the example of books. Clearly, books are private goods and not public goods. Yet books are unlike most private goods in having two distinct dimensions of output: the number of books published (here, x) is distinct from the number of copies sold (here, q). If only one book is published, then any number of copies can be produced, but the quantity of books available to any one person is effectively only one. If the unique book is arbitrarily short, then the constraint on one person's consumption is arbitrarily tight. Nonrival goods similarly have two dimensions of output: the number of paintings in a museum is distinct from the number of persons who see them; the miles of hiking trails blazed are distinct from the number of people who use them.

The two-dimensional aspect of output plays a crucial role because it supports the dichotomy that information goods (in our model) exhibit constant returns to scale in production but increasing returns in consumption. Increasing returns are evident from the observation (from (1)) that B/s is increasing in s.

If P=0, then information goods are nonrival: once the production cost c>0 is sunk, it costs nothing to give every consumer access. Because P . 0 in many contexts (but not for books) many information goods are approximately nonrival.[14] The distinction between P=0 and P . 0 plays a minor role in the results. Alternatively, if c=0, which we rule out, then information goods would be standard private goods produced at constant unit cost P.

The second characteristic of a classic public good is that is nonexclusive. As discussed in Section 2, we consider both possibilities, that information goods are exclusive and that they are not. If our information goods are unpriced and P=0, then they are pure public goods. We thus study, as a special case, the provision of pure public goods through advertising.

3. Equilibrium conditions.

We study symmetric outcomes, meaning that the endogenous variables of the model, (c,m,f,x,q,s,p)0ú^7, take common values across all firms and consumers.

---

[14] The marginal cost of provision is not the only difference between our information goods and certain nonrival goods. A consumer can choose to consume a fraction of the available books, hiking trails, or paintings in museums, but it is hard to consume only part of the national defense provided by the U.S. government. (Perhaps living in Canada would qualify.) For such indivisible nonrival goods, our assumption of constant returns to scale in production loses its meaning.

Definitions. An *outcome* of the model is a vector $(c,m,f,x,q,s,p) \in \mathbb{R}^7$, which satisfies $m,x,q,p \geq 0$, $c \geq 0$, and $s \in [0,1]$. An *interior* outcome satisfies these inequalities strictly.

Note that these definitions ignore any constraints that may apply to f. The constraint $f \geq 0$ or $f \leq 0$ binds in many of our "interior" equilibria.

We study two kinds of equilibria. Informally, a "consumer equilibrium" is an outcome such that (q,s,p) clears the advertising market and consumers respond optimally to firms' symmetric actions (c,m,f,x). A "full equilibrium" also requires that (c,m,f,x) reflect equilibrium behavior by firms in stage one. We will derive and discuss seven conditions, comprising six first-order conditions and one identity, which must be satisfied in a full equilibrium. At the end of this section is a simple example of an equilibrium.

A. Consumer demand.

The first step toward describing the symmetric equilibria of the model is to derive the demand for the output of a single defector from an otherwise symmetric profile of firms' actions. To derive that demand, we must retreat slightly from the continuous limit at which firms are infinitesimal, because consumers are indifferent to the consumption of an infinitesimal firm's output.

Suppose that all firms but one choose the same (c,m,f). Then for any consumer j at t=0 and any firm t, which chooses $(\hat{c},\hat{m},\hat{f})$, let $V(j,t) \equiv \beta(\hat{c},\hat{m}) - \gamma(t) - \hat{f} + \hat{m}\mu\varphi(q_j m) - \lambda\varphi(q_j)$; this is the rate at which accessing small but discrete quantities of firm t's output would change consumer j's utility.[15] Consumer j should access the images having the highest V(j,t) values, until her total consumption q_j becomes so high that the remaining images have negative value. If total consumption is the same, $q_j=q$, for every consumer j, as will be true in equilibrium, then this decision rule implies that the consumers who access the defector's good comprise an interval with the defector at the center, and the fraction of such consumers equals:

(6) $\hat{s}(\hat{c},\hat{m},\hat{f};q,m) \equiv 2\theta[\beta(\hat{c},\hat{m}) - \hat{f} + \hat{m}\mu\varphi(qm) - \lambda\varphi(q)] \in [0,1]$

(Recall from Section 2D that $\theta = \gamma^{-1}$ with an extended domain.) Equation (6) is the demand curve facing an individual firm in a symmetric equilibrium. Note that, in the continuous limit, $\hat{s}(\hat{c},\hat{m},\hat{f};q,m)$ depends on other firms' actions only through their impact on consumers' aggregate consumption of images (q) and impressions (qm), which are unaffected by the actions of any other single firm. That is analogous to the arms-length relationship among the firms in a standard competitive market, except that here the firm's optimization problem is framed by the aggregates q and qm rather than by the market price.

[15] The appendix provides a more careful derivation.

B. The consumer equilibrium (q,s,p).

Recall that by "consumer equilibrium," we mean that consumers act optimally and the advertising market clears, given firms' symmetric actions (c,m,f,x) in stage one. (One may imagine that the advertisers choose message content in stage one, but they pay only after the anticipated impressions occur in stage two.) To make this definition precise, we first establish a relationship between s and q, which are alternative measures of consumption: s is the fraction of consumers who access any given firm's image, and q is the total quantity of images accessed by any given consumer. Assuming that each consumer t=0 accesses the interval of firms $A=(-s/2,s/2)$, equation (4) implies that $q=\int_{-s/2}^{s/2} x\,dt$, implying:

$$q = sx. \tag{7}$$

Given symmetry in firms' and consumers' behavior, this is an identity rather than a behavioral relationship. It says that each firm's access ratio is the ratio of total consumption to total production.

Definition: An outcome (c,m,f,x,q,s,p) is a *consumer equilibrium* if, given (c,m,f,x): (q,s,p) satisfies the identity (7), clears the advertising market (2), and lies on the consumers' demand curve:

$$s = \hat{s}(c,m,f;q,m) \tag{8}$$

Lemma 0 shows that any choice of (c,m,f,x) in period 1 leads to a unique equilibrium in period 2. The proof is in the appendix.

Lemma 0. Given any symmetric firm actions (c,m,f,x), the consumer equilibrium is unique: equations (2), (7), and (8) have exactly one solution (q,s,p).

C. The period one equilibrium (c,m,f,x).

A full equilibrium must be, in addition to a consumer equilibrium, a Nash equilibrium among the firms in period one and also satisfy the free entry condition $\pi=0$:

$$msp + (f - P)s - c = 0 \tag{9}$$

The problem of the individual firm is to choose $(\hat{c},\hat{m},\hat{f})$ to maximize its (unit) profit $\pi(\hat{c},\hat{m},\hat{f},\hat{s};p)$, given: other firm's equilibrium actions (c,m,f), total output x, and the anticipated price of impressions p. (Posing the profit maximization problem as one of maximizing unit profit is standard for firms operating under conditions of free entry and constant returns to scale.) In equilibrium, profit-maximization for the individual firm requires, from (1) and (6):

(10) B̂(ĉ,m̂,f̂;p,q,m) is maximized at (ĉ,m̂,f̂)=(c,m,f), where

(11) B̂(ĉ,m̂,f̂;p,q,m) / B(ĉ,m̂,f̂,ŝ(ĉ,m̂,f̂;q,m);p)

Definition: An outcome (c,m,f,x,q,s,p) is a *full equilibrium* if it is a consumer equilibrium, satisfies (9) (free entry), and (10) (individual firm optimization).

We can now provide a consolidated statement of necessary conditions for an equilibrium. The appendix collects precise statements of all propositions and their proofs.

Proposition 1. If the interior outcome (c,m,f,s,q,x,p)0ú^7 is a full equilibrium, then:

(12a)	s = 2.($+m:N!f!6N(q)) .	[consumer optimization]
(12b)	p = "(qm) .	[advertiser optimization]
(12c)	q = sx .	[identity]
(12d)	(pm + f!P)s = c .	[free entry]
(12e)	$\$_2$+:N + ps$\$_1$ = 0.	[firm FOCs for c and m]
(12f)	($\$_2$+:N)(pm + f!P)2.N(((s/2)) + ps = 0 .	[firm FOC for m]
(12g)	$\$_2$ + :N + p = 0 if f...0 ;	[firm FOCs for f and m]
	$\$_2$ + :N + p $ 0 if f=0 and images are excludable ;	
	$\$_2$ + :N + p # 0 if f=0 and images are includable .	

where the arguments of $(c,m), $\$_1$(c,m), $\$_2$(c,m), and :N(qm) are suppressed and understood to take equilibrium values.

For the case of unpriced goods, f/0 is fixed and condition (12f) disappears. For a consumer equilibrium in stage two, recall that only conditions (12a)-(12c) are required.

The necessary conditions (12) are not sufficient for a full equilibrium, because the firm's profit function (11) is typically not globally concave in (ĉ,m̂,f̂).[16] That is unimportant for the results in this paper, because we use (12) only to derive properties of interior equilibria. Section 3E provides an example of an interior equilibrium.

[16] The necessary (local) second-order conditions for the firm's optimization problem are more complicated that the first-order conditions. If $\$_{12}$=0 (as is true for the example below), then the second-order conditions reduce to: $\$_{11}$ < 0; $\$_{22}$ # 0; !s^3$\$_{11}$[(N+s(0/4] $ 1.

D. Interpretation of the firms' first-order conditions.

We have already discussed equilibrium conditions (12a)-(12d), which have straightforward interpretations. The less obvious features of the equilibrium arise mainly from the firms' optimization problem. To aid intuition, we now discuss the first-order conditions for that problem, (12e)-(12g), in detail. To simplify condition (12g), we assume f…0 throughout.

The recurring expression \$₂+:N represents the marginal impact of an impression on consumers' utility. Consumers may gain by the presence of advertising messages, but p>0 (from (12b)) and mp+f!P>0 (from (12d)) imply, from (12f), that advertising must occur to the point that \$₂+:N<0. The intuition is that if consumers preferred to get more impressions, then any firm could increase its revenue by increasing message density, without any offsetting cost. These properties of equilibrium will be important throughout the analysis:

(13) mp+f!P>0; \$₂+:N<0.

We now study the firm's decision by breaking it into three parts.

i. The message density decision.

The firm's choice of message density is unlike any decision that appears in standard firm or market theory, and so it is worthwhile to study (12f), the associated first-order condition. Using carats to denote variables affected by the firm's choice of m\$, we rewrite it slightly:

(12fN) 2.N((\$/2))[\$₂(c,m\$)+:N(qm)] = ! p\$/(m\$p+f\$!P)

The left-hand side of (12fN) is the rate at which the marginal message reduces access to the firm's good: M\$/Mm\$<0, as derived from the firm's demand curve (6). If .N>0 is high, indicating that the marginal consumer sacrifices relatively little by switching to the closest available substitutes, and the consumer's marginal utility from a message, \$₂+:N<0, is very negative, then the firm's demand is more sensitive to its message density. The right-hand side of (12fN) is the firm's marginal rate of substitution (on the revenue side) between messages and access: !(MB/Mm\$)/(MB/M\$), as derived from the profit function (1). For simplicity, suppose that .N is fixed. Then, as the firm increases m\$, the left-hand side of (13eN) becomes more negative as consumers become increasingly intolerant of messages (from \$₂₂<0) and the right-hand becomes less negative, because M\$/Mm\$<0. In other words, as the firm increases message density, the marginal message has an increasingly negative impact on access, and increasing access becomes relatively more important for profits than increasing messages, because fewer consumers are watching more messages. Equilibrium occurs where message density is so high that these combined effects cause the return to the marginal message to fall to zero.

ii. Optimizing the revenue mix.

To explain the second condition for firm optimization, (12g), a picture is useful. For Figure 1, let (m*,f*,x*,q*,c*,p*,s*) be an equilibrium, and we take all of these values to be fixed, except for a *single* firm's choice of (m,f). Think of this firm as producing a single image. In Figure 1, the firm's linear iso-profit lines are derived from (1) and fixed. Lower isoprofit lines indicate more ads and higher access fees and consequently higher profits (holding the firm's access ratio fixed at s*). The curved line, from the demand curve (8), shows which (m,f) combinations cause consumers to access the firm at the equilibrium rate s*. Points above the curve would produce higher access rates. It is useful to interpret this curve as an indifference curve for the firm's marginal consumer; its slope is the rate at which that consumer is willing to exchange a higher access fee for less advertising, holding her total consumption constant. (The curve has a positive slope at small m to reflect the possibility that the consumer gets positive marginal utility from the first few messages, but p>0 implies that equilibrium must occur where the curve has a negative slope.) Profit-maximization requires the firm to choose whichever (m,f) combination on the access constraint reaches the highest iso-profit line. This obviously occurs at a tangency, and that is the content of condition (12g). The slope of the isoprofit line is $-p$, differentiating (8) shows that the slope of the access constraint is $\beta_2(c^*,m)+\mu\phi(q^*m^*)$, and (12g) states that these slopes must be equal in equilibrium.

Condition (12g) thus means that the firm should balance its two sources of revenue, advertising and access fees, to generate access as efficiently as possible. (That result does not depend on free entry, but the free entry condition requires that x* be high enough to eliminate profits and bring the tangency to the iso-profit line that represents zero profits.) Condition (12g) is analogous to the usual first-order conditions for an optimal mix of inputs, except that in this case the firm is optimizing its mix of revenue sources. Note that if the firm chooses f<0 (points near the top of the figure), then advertisers contribute more than 100% of its revenue.

It is interesting to understand what determines the firm's revenue mix, the ratio mp/f, but a thorough analysis is beyond the scope of this paper. Figure 1 indicates that, as one would expect, the firm favors advertising revenue when the price of impressions (p) is high and consumers become less averse to advertising at the margin (i.e. $|\beta_2(c^*,m)+\mu\phi(q^*m^*)|$ is small).

Figure 1: Optimal Revenue Allocation

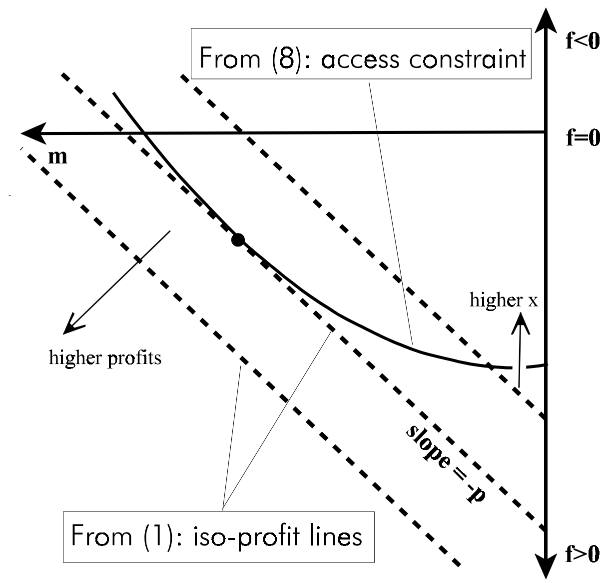
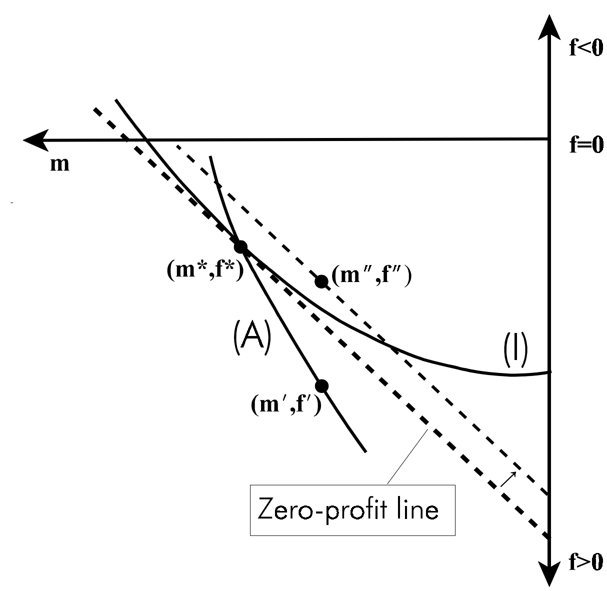


Figure 2: Surplus-Enhancing Reallocation



iii. Quality depends on the access ratio.

The third condition for firm optimization, (12e), concerns quality choice. For simplicity, continue to assume that f...0, so that (12g) holds with equality. Then (12g) simplifies (12e) to s$\$_1$=1. This condition means that quality is provided efficiently, holding other firm decisions and the access ratio fixed: the marginal dollar spent on quality gives each consumer accessing an image a marginal benefit of $\$_1$, and the fraction of consumers reaping that benefit is s. This efficiency result can be expressed as a tangency condition similar to that in Figure 2: the firm resolves the tradeoff between higher quality and lower access fees in whichever way minimizes its costs of achieving the equilibrium level of access.

The equilibrium condition s$\$_1$=1 also indicates that quality is determined substantially by the access rate. Indeed, quality may be very sensitive to the access rate, because totally differentiating s$\$_1$(c,m)=1 yields dc/ds=1/(s$\$_{11}$), and small *$\$_{11}$* is economically reasonable. Higher access rates tend to increase quality for another reason, unrelated to condition (12e): the higher unit profits implied by higher access rates must be exhausted somewhere, and higher costs per unit are one outlet for those profits.

In principle, the advertising subsidy could allow firms to increase the quality of their images, but we have just shown this can occur through only two channels: advertising must either increase consumers' willingness to pay for quality on the margin (which is possible because the sign of $\$_{12}$ is unclear), or it must increase the access ratio. This model provides no obvious indication that advertising has these effects.[17]

E. An example.

We now provide a simple example of a full equilibrium, assuming an includable good.

Example. Assume that prices are measured in dollars and a consumer's utility function (5) is characterized by \$(c,m) = 120$c^{1/3}$ + m/100 ! m^2/20,000, ((t) = 1,600t, : / 0, and 6(q) = q^2. Assume that P=\$2 and q)=\$1, and the inverse demand for impressions is "(qm) = .218 ! qm/500,000.[18] Then it is easy to confirm that the unique solution of the equilibrium conditions (12) is x=640, q=64, s=0.1, c=\$8, m=1,000, f= ! \$8, and p=\$0.09. This point also satisfies the necessary second-order conditions for the individual firm's

---

[17] Stegeman (2002) studies the comparative statics properties of the model, and if $\$_{12}$=0 (meaning that message density does not influence consumers' tastes for quality), then quality is generally insensitive to changes in consumers' tastes or advertisers' demand for impressions. In particular, the marginal profits generated by increases in consumers' or advertisers' demand are typically dissipated through channels other than higher quality. This result is consistent with the discussion above, because such demand changes also prove to have little effect on equilibrium access rates.

[18] This demand curve violates our assumption ">0, but the only role of that assumption is to ensure that p>0. We set q) for completeness; all that matters is that q)<8.

maximization problem (stated in fn. (16)), and numerical methods show that they represent a global maximum for the individual firm.

In this equilibrium, the firms collectively produce x=640 images, and each consumer accesses s=10% of that total. The firms pay the consumers a total of $-$xsf=\$512 in access fees (i.e. they pay the consumers to access the good) and produce xsm=64,000 impressions. The marginal impression reduces the consumer's utility by β_2=\$.09, matching the cost (p) to the advertiser. For each image, a firm earns smp=\$9 in advertising revenue, against which it pays c=\$8 in production costs, sP=\$0.20 in access costs, and sf=\$0.80 in access fees to consumers, for a net profit of zero.

4. Welfare measure.

To set the stage for the welfare results, the last step is to define (total) surplus. A basic measure sums the surpluses accruing to the firms, advertisers, and consumers. (Firms get no surplus in equilibrium but we will consider perturbations from equilibrium that violate the zero-profit constraint.) We generalize that basic measure slightly by accounting for the possibility of impression externalities.

A. Impression externalities.

Economists' formal models of advertising often assume (at least implicitly) that advertising creates externalities, and broader discussions of advertising often concern normative issues that turn on externalities. Therefore, it seems appropriate to generalize our welfare analysis by accounting for the possibility that consuming advertising has external effects on persons other than the advertiser, the firm, and the consumer who chose to create and access the ad. Such *impression externalities* could be positive or negative. All of our welfare results apply to the benchmark case that there are no impression externalities, and the results prove to be less sensitive to such externalities than one might expect.

Let Φ(xms), where $\Phi:\mathbb{R}_+ \to \mathbb{R}$ is exogenous, denote the external impact of the (xms)th impression. If Φ(xms)>0, then the xmsth impression has a positive external effect; if Φ(xms)<0, then the effect is negative. Our general model is agnostic about the source of such effects, but the specific model of advertising in Section 6 fills that black box by assuming that advertising affects only advertisers' market shares without increasing their total output, and this ultimately implies that Φ(xms)=$-$p in equilibrium. That may be an overly pessimistic view of advertising, but our welfare conclusions can survive impression externalities of that magnitude.[19]

B. Welfare measure.

Most of the welfare results consider the impact of identical perturbations of firms' actions away from equilibrium values, assuming that consumers and the advertising market reach the consumer equilibrium implied by those changes. Global analysis is useful, but isolating the various margins in the model and studying them individually seems helpful for gaining a basic understanding of whatever distortions exist. Also, while this paper is not about policy, part of the rationale for incremental analysis is that most policies cause incremental changes.

Because all of our welfare comparisons are between consumer equilibria, we need a welfare measure that applies to all such equilibria. In a consumer equilibrium, $q=sx$, $A=[-s/2,s/2]$, and (5) imply that each consumer accrues utility:

$$(14)\qquad u^* = q[\beta(c,m) - f] - 2x\int_0^{s/2} \gamma(t)dt + \mu(qm) - \kappa(q)\,,$$

Firms' total surplus (profit) equals $x\pi(c,m,f,s;p)$. Summing the surpluses accruing to firms, advertisers, and consumers, using (1), (3), and (14), accounting for impression externalities if $\sigma\neq 0$, and canceling terms representing payments between groups, the (total) surplus in a consumer equilibrium is:

$$(15)\qquad W \equiv \int_0^{mq}[\alpha(t)+\sigma(t)]dt - cx + [\beta(c,m) - P]q - 2x\int_0^{s/2} \gamma(t)dt + \mu(qm) - \kappa(q)\,.$$

If images are unpriced ($f\equiv 0$), then an awkward issue arises: it is unclear how consumer utility should be scaled relative to other sources of surplus. Multiplying β, γ, μ, and κ by a common factor leaves consumer behavior unchanged but alters the welfare measure (15). It is hard to obtain welfare results that are insensitive to this scaling, but we obtain unambiguous results based upon whether the $f\equiv 0$ constraint binds from above or below. This works because the sign of $\partial\pi/\partial f|_{f=0}$ contains implicit information about consumers' marginal rate of substitution between images and money. We discuss an example of this in Section 5A.

---

[19] Many persons might agree that the liquor industry presents one of the most plausible cases (if there are any) of substantial negative externalities to advertising. In that context, $\sigma > -p$ means roughly that the positive impact of the marginal liquor ad on the combined profits of all liquor companies and related businesses, ignoring the cost of placing the ad, exceeds the negative external impacts on consumers because *other* consumers see the ad: impacts such as being the victim of a drunk driver. That claim is not conspicuously absurd, but it is an empirical question that the author is unqualified to answer.

5. Welfare analysis of firms' equilibrium behavior.

In this section, we study how small deviations in firms' actions, away from a full equilibrium, affect surplus, after accounting for consumers' and advertisers' responses. Specifically, we study how a small symmetric change in each of (c,f,m,x), leaving the other three actions fixed at original equilibrium values, affects surplus as defined by (15), assuming that (p,q,s) adjust as required by conditions (12a)-(12c). For changes in c, f, and m, we obtain surprisingly sharp results. After presenting those results, we study the ambiguous impact of changes in output.

Definitions. If images are nonexcludable, then it is possible that in equilibrium f=0 but $\partial B/\partial f>0$; in this case we say that *f=0+*. If images are nonincludable, then it is similarly possible that in equilibrium f=0 but $\partial B/\partial f<0$; in this case we say that *f=0 −*. For all other equilibria, we say that *f≠0±*.

Our benchmark case is *f≠0±* (i.e., f is an unconstrained optimum for each firm). All of our results apply to that case and to at least one of the other two cases.

A. Advertising and fees are too high; quality is too low.

Our first results support popular impressions: firms put too little quality and too much advertising into information goods, and they set access fees too high. (Indeed, the efficient access fee is generally negative.)

Proposition 2.

A. Assume that the model starts in a full interior equilibrium, with f≠0 −. Suppose that F=0 (no impression externalities).

			Then:
			Surplus (W)
If the firms:	Increase quality (c)	|	Rises
	Reduce access fees (f)	|	Rises[20]
	Reduce message density (m)	|	Rises

If F<0 (a negative externality), but $F>-p-(f-P)/m$, then the same results obtain. If F>0 (a positive externality), then the first two results obtain, but the result for "m" is no longer assured.

---

[20] If f=0 −, then reducing the access fee would again increase surplus, but the constraint f≥0 prevents firms from doing so. The statement of Proposition 2 in the appendix covers this case.

B. Consider any fixed (c,m,x). Then any surplus-maximizing choice of f, if it generates non-negative surplus (W>0) and incomplete access (s<1), must satisfy f = P ! m[p+F(mq)].

The intuition for Proposition 2 is closely related to the familiar underutilization of nonrival goods, due to increasing returns in consumption. The standard analysis shows that the owner of a nonrival good could maximize total surplus by setting the access fee equal to the zero marginal cost of access, but a zero access fee requires a subsidy. Advertising is a potential source of this subsidy, but Proposition 2A shows that a similar inefficiency remains: reducing the access fee (f) below its equilibrium value increases surplus by reducing the deadweight loss caused by too little access. This is true even if the equilibrium access fee is f=0, because the access fee is only part of the "price" that consumers pay for images: they must also consume advertising messages to the point of getting negative utility from the marginal message. Proposition 2B shows that the surplus-maximizing access fee is, in the absence of negative impression externalities, smaller than P, the marginal cost of access. In particular, to achieve efficiency the access fee should be reduced by the market value of all impressions that the accessing consumer generates.

In the example of Section 3E, the equilibrium access fee is ! $8, but Proposition 2B shows that, all else equal, the efficient fee would be f = P ! mp = ! $88, to account for the considerable surplus that the decision to access a good creates for the advertisers who have bought the embedded messages.

It is unsurprising that firms set fees too high, given the increasing returns to consumption implied by (1) (and the inherent market power discussed in Section 2B), but what is the intuition for excess advertising? It is the same. Message density (m) is the implicit second component of the fee that firms extract from consumers, and firms overprice on both margins. Raising quality (c) would also increase surplus, because that is yet another way to transfer surplus to the marginal consumer.

The example of Section 3E illustrates Proposition 2. Starting from the equilibrium described, we can use (12a)-(12c) to study the impact of incremental changes in m on q and p, keeping c, f, and x fixed at original equilibrium values. Routine calculations show that dq/dm . ! 0.0277, d(qm)/dm . 36.3, and dp/dm . ! 0.0000726. Expressing total profits as xB = qmp+(f-P)q-cx (from (1) and (7)), it follows that d(Bx)/dm = p×d(qm)/dm + qm×dp/dm + (f-P)dq/dm . ! 1.1 < 0. Reducing advertising messages would thus *increase* the firms' profits: the higher price of impressions would more than offset the smaller number of impressions sold and the larger payments to consumers as their consumption increases. Proposition 2A shows that such a collective advertising reduction would also increase surplus.

We conclude our discussion of Proposition 2 with extended remarks about the circumstances under which the claims of Proposition 2A need not hold. The most important exception is equilibria with f=0 !, meaning that nonincludability binds. If f=0!, then the firms' inability to set a negative access price causes them to encourage access in other ways – by increasing quality and reducing message density – and these adjustments may exhaust any welfare gains obtainable by increasing c or reducing m. Given such an equilibrium, the one adjustment that would clearly increase surplus would be a reduction in the access fee, but that is precisely what nonincludability does not allow.

Large access costs P tend to increase equilibrium access fees and reduce the likelihood that nonincludability binds. In other words, if P is large then (all else equal) it is more likely that f>0, which ensures that firms advertise too much and the conclusions of Proposition 2A obtain.

A useful way to interpret the distinction between f=0+ and f=0! equilibria, given f=0, is that it contains implicit information about consumers' valuation of unpriced goods. For example, consider the claim that reducing advertising increases surplus. Reducing advertising helps consumers but reduces the collective surplus of advertisers and firms.[21] Given this tension, any claim about the net impact on surplus must rest on an implicit assumption about consumers' valuation of access, but if the access fee is constrained to zero then no market reveals that value. How then, can we claim net benefits from reducing advertising in unpriced goods? The answer is that if consumers' valuation of access was small enough to overturn the claim, then it would be profitable for firms to pay consumers to access their images, implying f=0!. By assuming f…0!, Proposition 2A rules this out.

These observations underscore the point that, if images are unpriced, then determining the welfare properties of an equilibrium depend partly on the ability to distinguish f=0! from f=0+ equilibria. Clues may exist. For example, inventions allowing firms to charge access fees for radio or television signals have often led to substantial fees; this suggests that the equilibrium for free radio and television was of the f=0+ kind, which in our model implies excessive advertising. On the other hand, some internet sites have attempted to pay consumers to access their sites, suggesting that equilibria in this market may have f=0!, indicating that advertising may not be excessive. (This is slightly counterintuitive: it is precisely when firms embed so much advertising in their images that they would be willing to pay consumers to access them that more ads might increase welfare.) It may also be possible, through experiments or surveys, to estimate $\$_2+:N$, consumers' marginal disutility from advertising messages, and this may be easier than measuring directly the value that they get from an unpriced good. Comparing $\$_2+:N$ to the market price of a impression, p, indicates (in principle) whether the equilibrium is of kind f=0! or f=0+ (cf. Lemma A1 in the appendix).

---

[21] If reducing m across all firms were to increase impressions per image qm, then a single firm could do the same, and this would always be profitable when $f\$0$. Therefore, reducing m must decrease qm in the consumer equilibrium. When f=0, the combined surplus of advertisers and firms varies with the area under the demand curve for advertising (holding cx fixed), which falls as qm falls.

We conclude our discussion of Proposition 2 by considering how impression externalities may change the results. The results are fairly robust: externalities ($\Phi \neq 0$) do not affect the main claims of Proposition 2A, with two exceptions. First, if advertising impressions create sufficiently large positive spillovers, then (as one would expect) it is possible that firms could increase surplus by supplying more advertising.

Second, and less obvious, very negative externalities can overturn all of the results. For intuition, suppose that persons who see web advertisements become serial murderers. This consideration may overwhelm all other normative issues, implying that the best policy is not to encourage access by increasing product quality, etc., but instead to persuade people to turn off their computers. Externalities must be very negative, however, before this effect can dominate. For example, if access fees cover the marginal cost of access ($f \geq \Pi$), then Proposition 2A shows that only $\Phi < -p$ can overturn the results, meaning that the negative externality from an advertising impression must exceed the gross benefit of the impression to the advertiser. That seems unlikely in most circumstances (cf. fn. 19). If $f<\Pi$, so that advertisers are effectively subsidizing consumption of their messages, then less extreme negative externalities can potentially overturn the results.

B. Over-reliance on advertising revenue.

In many cases, unlike in the example of Section 3E, the welfare-enhancing adjustments described by Proposition 2 would cause firms to make losses. The next proposition describes a welfare-enhancing adjustment that respects the zero-profit condition. In many equilibria, the firms could increase surplus by supplying less advertising and making a profit-compensating adjustment in the access fee.

Proposition 3. Assume that the model starts in a full interior equilibrium, with $f \neq 0$. Assume that consuming advertising creates no positive externalities ($\Phi \leq 0$) and any negative externality is bounded by $\Phi(qm) \geq -p-(f-\Pi)/m$. If the firms decrease message density (m) and make a profit-compensating adjustment in the access fee (f), then surplus increases.[22]

A related result appears in Stegeman (2002). There it is shown that a budget-balanced change in the corporate profits tax, which taxes revenue from advertising at a higher rate than revenue from access fees, increases consumer surplus, total surplus, and the total output of information goods.

[22] It is easy to show that the result $dW/dm<0$, in Proposition 2A, extends to the case $\Phi(qm)= -p-(f-\Pi)/m$. Separately, note that if f=0 in equilibrium and images are nonexcludable, then the profit-compensating adjustment in the access fee may not be feasible.

The rest of this section explains the source of this (somewhat subtle) welfare gain. It comes through two channels. The more obvious channel is that when the firms supply less advertising, they effectively collude against advertisers, reducing the quantity of impressions supplied and driving up their price (p). Pairing this policy with a profit-compensating adjustment in the access fee uses the surplus extracted from the advertisers to subsidize consumers' access. This partially corrects consumers' underconsumption of images, the fundamental inefficiency behind the results in Proposition 2. A possibly counterintuitive aspect of this policy is that it often mitigates the underutilization problem by *raising* the access fee, but the higher access fee is accompanied (from the consumers' viewpoint) by a more-than-compensating reduction in advertising.

The second channel of gain is similar but less obvious. By collectively reducing message density, the firms also "collude" against consumers by reducing their consumption of messages and increasing the implicit price that consumers are willing to pay for a message. Because that implicit price is negative, however, reducing message density also, paradoxically, returns the surplus to the consumers. The firms could recover the surplus by raising the access fee, but the profit-neutral policy proposed by Proposition 3 requires them to invest it (like the surplus extracted from the advertisers) in an access fee low enough to increase consumers' access rate. The higher access rate creates new surplus that exceeds the surplus destroyed through "collusion." Figure 2 helps to explain how extracting surplus from consumers, only to return it, produces gains.

Figure 2 reproduces Figure 1, but we now interpret (m,f) as the symmetric choice of *all* firms. We assume throughout that (c*,m*,f*,x*,q*,s*,p*) is a full equilibrium and that (x,c,p) remains fixed at its equilibrium value (x*,c*,p*). (By fixing p*, we set aside any surplus extracted from the advertisers.) Therefore, the isoprofit line shown in Figure 2 is the zero-profit line. Figure 2 omits the indifference curve shown in Figure 1, which represented the preferences of a single consumer for the product of a single firm, and instead depicts two curved lines reflecting the implications of aggregate firm behavior. The indifference curve (I) comprises the (m,f) choices which, if adopted by *all* firms, would leave consumers just as well off as they are in equilibrium, assuming that they reoptimize their consumption q=sx*. Obviously (I) must pass through (m*,f*). Totally differentiating the utility function (5) after the substitution $A_j=(-q_j/(2x),q_j/(2x))$ and invoking the envelope theorem to abstract from the consumers' adjustments of q shows that the slope of (I) is $\$_2(m)+:\mathbb{N}(qm)$. Therefore, much as in Figure 1, condition (12g) implies that (I) and the zero-profit line must be tangent at (m*,f*). The tangencies in Figures 1 and 2 show that in equilibrium firms choose the (m,f) mixture that generates access most efficiently, in the aggregate as well as on the individual level.

Curve (A) in Figure 2 depicts an aggregate iso-access locus, comprising the values of (m,f) that cause the consumers to choose an access rate equal to the equilibrium rate s*. Curve (A) comes from the consumers' equilibrium condition (12a); totally differentiating (12a) shows that its slope is $\beta_2(m)+\mu'(qm)+qm\mu''(qm)$. If $\mu''(q^*m^*)<0$, then (A) is steeper than (I) at (m*,f*). Points to the right produce access rates s>s*. Intuitively, $\mu''<0$ is important because that is what makes a consumer's marginal rate of substitution between money and any one firm's messages sensitive to the aggregate quantity of messages and so creates the possibility of profitable "collusion" against the consumers by collectively curtailing advertising.

Returning to Figure 2, suppose that all firms deviate from equilibrium play to a nearby point such as (m',f'). Then the access rate is unchanged (implying that the zero-profit line does not move), and the firms earn higher profits. The consumers are accessing fewer messages, and the higher implicit price that the firms receive for those messages produces the increased profits. When the firms return those profits to the consumers by reducing the access fee from f', as described by Proposition 3, the access rate increases, and this produces a first-order surplus gain that exceeds the second-order surplus lost by the firms' incremental "collusion" in the implicit market for impressions. After this adjustment, the zero-profit line rises (because of the higher access rate) and the firms choose a point such as (m'',f''), which leaves them with zero profits but consumers strictly better off than they were in the original equilibrium.

In the proof of Proposition 3, it is crucial that $\alpha'+\mu''<0$. Our original assumption that $\alpha'<0$ allows the firms to collude profitably against the advertisers, and if $\mu''<0$ then the firms can likewise collude against the consumers. Using the surplus extracted through both channels to subsidize access produces the efficiency gain described by Proposition 3. The social advantages of substituting away from advertising revenue are robust. They reflect the advantage of transferring surplus from an "efficient" margin, here the efficient provision of a given level of access, to an inefficient margin, the overpricing of access evident in Proposition 2.

C. Over- and underproduction.

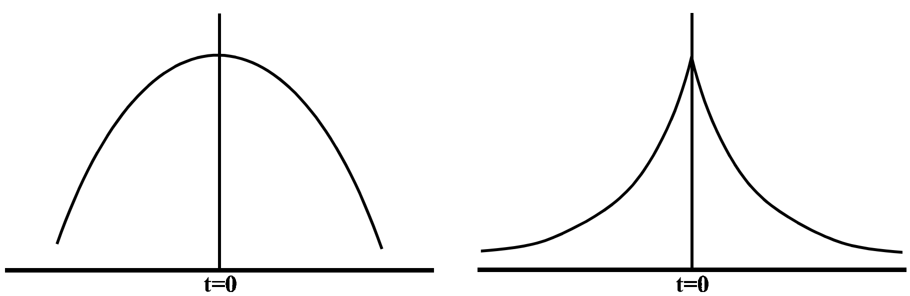
The most ambiguous dimension of the welfare analysis is the sign of dW/dx, the question of whether firms produce too many or too few images. Spence (1976) points out that high own-price and cross-price elasticities can cause socially excessive entry by producers of private goods, and similar effects appear here.[23]

---

[23] It is hard to improve on Spence's (1976) explanation: "Under monopolistic competition, there are basically two forces at work... First, because revenues do not capture the consumer surplus, revenues may not cover costs even when the social value of the product [at the margin, given existing products] is positive.

Proposition 4A states four conditions that jointly imply overproduction of information goods, while Proposition 4B shows that reversing the four conditions implies underproduction. Proposition 4 is not a very strong result, but it shows which sorts of circumstances tend to cause over- and underproduction. The shape of the horizontal penalty function (plays an important role. If (is convex, then we say that consumers have *concave horizontal utility* because $! ((t) is concave in t, for fixed $. This means that if we ordered the images produced in equilibrium, according to the preferences of an arbitrary consumer, then she would have weaker preferences between #1 and #2 than between #n and #n+1. Similarly, if (is concave, then we say that consumers have *convex horizontal utility*, meaning that the consumer's strongest preferences would be between #1 and #2.[24] Figure 3 illustrates these concepts.

Figure 3



Concave horizontal utility (γ convex)   Convex horizontal utility (γ concave)

Proposition 4. Assume that the model starts in a full interior equilibrium.

A. If horizontal utility is concave, consuming advertising creates no positive externalities (F#0), f…0+ (equivalently, $p\#*\$_2+:N*$), and $x > (N/[2(60!m^2:0)]$, then firms overprovide at the margin (i.e., decreasing output would increase surplus).

This is a force tending to eliminate products that should be produced. Second, when a product is introduced, it affects other firm's profits as well as increasing consumer surplus. If the products are substitutes, the effect on other firms' profits is adverse. Since the entering firm does not take these effects into account, it may enter when it is not generating a net social benefit. This is a force tending to generate too many products in the case of substitutes."

[24] Though the terminology does not make this distinction, for Proposition 2 it is sufficient that this ordering be constructed only through n such that the consumer actually consumes image #n in equilibrium.

B. If horizontal utility is convex, consuming advertising creates no negative externalities (F$0), f...0! (equivalently, p$*$$_2$+:N*), and x < (N/[2(60!m^2:0)], then firms underprovide at the margin.

For the benchmark case of linear (, no externalities, and f...0, Proposition 4 provides a sharp test: whether the firms overprovide or underprovide images depends solely upon whether x exceeds (N/[2(60!m^2:0)]. This question has no presumptive answer, because neither the necessary first-order conditions nor the necessary second-order conditions for equilibrium place any restrictions on 60 and :0, beyond our original assumptions that 60>0 and :0#0.[25] The example of Section 3E satisfies the benchmark assumptions, and in that case x=640 and (N/[2(60!m^2:0)]=400, implying overproduction.[26]

To gain intuition for the "x-test" of Proposition 4, suppose that 60 or !:0 is very large. Then consumers' aggregate demand for images (q) is very inelastic in the neighborhood of equilibrium, and the only welfare gain from higher output (x) is that consumers shift their consumption toward images that better satisfy their specialized tastes (e.g. they can spend all of their online time reading news about Uzbekistan). Because x > (N/[2(60!m^2:0)] must hold for 60 and !:0 sufficiently large, the x-test shows that gains from such demand-shifting alone *never* justify the marginal cost of the additional output (in the benchmark case). In other words, the external benefit to consumers from allowing them to specialize their consumption never overcomes the external harm to firms of capturing their customers. Underproduction is possible only if higher output would cause sufficiently higher aggregate consumption.

Beyond the benchmark case, the shape of (is important. If horizontal utility is convex, then consumers get increasing marginal benefits from confining their consumption to the images they like best, implying that they would benefit greatly from greater availability of images catering to their specialized tastes. These preferences remain unexpressed to the market, however, because they are inframarginal. Consequently, if utility "spikes" at t=0, then the potential welfare gains from catering to specialized tastes are unbounded and underproduction is an inescapable possibility. On the other hand, if horizontal utility is concave, then the potential gains available from catering to specialized tastes are bounded and this makes it possible to obtain results such as Proposition 4A.

---

[25] The appendix shows that the second-order condition for consumers' consumption decision is always satisfied. The firms' second-order conditions (cf. fn. 16) are unrelated to 60 and :0 for essentially the same reason that optimization by a competitive firm places no restriction on the elasticity of aggregate demand, in the standard supply and demand framework.

[26] This calculation does not imply that x=400 would maximize welfare, because the derivation of the test inequality invokes the equilibrium conditions of Proposition 1, which need not hold away from equilibrium.

The third condition invoked by Proposition 4 is the possibility that the access fee is constrained to zero. If f=0+, meaning that nonexcludability binds, then Proposition 4A suggests that underproduction is more likely. The reason is that firms' output is constrained by their inability to raise revenue directly from consumers. The alternative source of revenue is advertisers, and supplying more advertising drives up consumers' disutility from the marginal impression, $!(\$_2+:\mathbb{N})$. If this disutility far exceeds p, then it is a sign that nonexcludability binds tightly, which increases the likelihood of underproduction.[27] Conversely, if f=0!, then the firms cannot compete on price as aggressively as they would like to, and the profits generated by this involuntary collusion are dissipated by entry, which makes overproduction more likely.

The intuition behind the fourth condition, the presence of impression externalities, is obvious. If consuming advertising creates negative externalities, then the images conveying the advertising are more likely to be overprovided.

The appendix states a lemma (Lemma A3) which can be used to construct tighter tests for overproduction and underproduction than those provided by Proposition 4, in the case that (is linear. In that case, whether firms overprovide or underprovide images depends on the sign of a single (somewhat complicated) inequality.

The next proposition states economically plausible and sufficient conditions for overproduction which differ somewhat from those of Proposition 4A. Instead of the x-test, Proposition 5 uses information about the firms' revenue mix. There is no comparable result for underproduction.

Proposition 5. Assume that the model starts in a full interior equilibrium, with f...0+. If horizontal utility is concave, F<0, and $[!F/p]\times[pm/(pm+f!P)]\ \$\ ½$, then firms overprovide at the margin.

In Proposition 5, the term !F/p compares the advertisers' gross benefit from the marginal impression to the negative external effect of that impression. If a negative external effect (e.g, on the profits of the advertisers' rivals) offsets most of the advertisers' benefit, then !F/p is large. In the particular model of Section 6, !F/p=1. The term pm/(pm+f!P) indicates what fraction of the firm's unit profit from access (i.e., ignoring production costs) comes from advertisers. If f=0 and P>0, then this exceeds one.

Proposition 5 shows that if advertisers provide most of the firms' revenue, and negative externalities offset most of the advertisers' benefits from impressions, then images are overprovided. The intuition is that

[27] Much of the extensive literature on whether broadcasters should be forced to set f=0 (cf. Section 7) turns, implicitly, on the underproduction that can occur in f=0+ equilibria.

advertisers have socially excessive incentives to subsidize the consumption of their messages.[28] It is worth emphasizing that the assumptions of Proposition 5 are not necessary for overproduction. The inequality is constructed from the assumption that 6 and : are approximately linear. Any reasonable amount of curvature in 6 and :, which reduces the consumption response to higher output, makes it much easier to find overproduction in equilibrium.

Summarizing, Propositions 4 and 5 suggest that overproduction is more likely when:

• Horizontal utility is concave (or less convex).

• The marginal opportunity cost of consumption is locally highly convex.

• Images are nonincludable but firms would pay consumers for access if they could.

• Consuming advertising generates negative externalities (i.e., beyond its direct impact on the producing firm, the advertiser, and the accessing consumer). In the presence of such negative externalities, excess production is more likely when firms get most of their revenue from advertising.

Symmetrically, underproduction is more likely when horizontal utility is convex, the marginal opportunity cost of consumption is locally close to linear, nonexcludability binds, and consuming advertising generates positive externalities.

D. The underlying sources of inefficiency.

On many margins, the equilibria of the model are efficient, as illustrated by Figures 1 and 2. This is unsurprising, given that all agents are small and do not interact strategically. The inefficiencies that do arise can be traced back to the distinctive properties of information goods that *while any number of consumers can consume (i.e., access) each image, each individual consumer can consume each image only once*.

The first property leads to increasing returns to consumption, which by familiar reasoning makes marginal cost pricing incompatible with non-negative profits. A priori, one might hope that the advertising subsidy, combined with production by small firms under constant returns to scale, could somehow remedy this fundamental problem, but Proposition 2 dismisses that hope: firms "overcharge" consumers on every margin. Increasing returns to consumption also cause the systematic distortion in the revenue mix described by Theorem 3. Ironically, it is the efficient production of access that causes the revenue distortion, because adjusting the revenue mix away from advertising revenue exchanges a second-order technical loss for a first-

---

[28] The inequality is related to the inequality in Proposition 1 which requires that impression externalities not be too negative. The violation of the Proposition 1 inequality, and the satisfaction of the Proposition 5 inequality, each correspond, in slightly different ways, to the possibility that access may destroy surplus at the margin, if impressions create sufficiently large negative externalities.

order gain, sacrificing the production of impressions, a private good, in favor of increasing consumption of the information good. This inefficiency, which is the natural consequence of embedding advertising messages in information goods, seems to have no analog elsewhere in market theory.

The second property of information goods means that consumers generally have excess demand for the images that they most prefer (and this possibility seems realistic), but because the intensity of this demand finds no expression at the margin, underproduction may result. On the other hand, the first property of information goods implies that firms can strictly increase their profits by attracting the marginal consumer away from other firms, and this capture externality can cause overproduction, as is evident in the relatively weak conditions that imply overproduction if the aggregate demand for images is inelastic. These conflicting effects lead to the ambiguities studied by Propositions 4 and 5.

6. A specific model of advertising demand

In the original model, ", $, :, and F are "black boxes:" they describe the benefit that advertisers get from advertising ("), the benefit or harm that consumers derive from their consumption of advertising ($ and :), and the externalities that this activity imposes on others (F). This section describes a particular model of persuasive advertising, which generates functional forms for ", :, and F. (The model places no restrictions on $.) Its main features are that consuming advertised products gives consumers a utility bonus, but the total demand for goods is fixed; advertising influences only *which* goods consumers buy.

We develop this model independently and then show how to embed it into the original model. Consider a continuum of conventional private goods, on the [0,1] interval. Each point on this interval represents a distinct good, but points do not represent locations: the distance between goods is unimportant. Each private good k0[0,1] is produced at zero cost by two firms, a blue firm and a green firm, which specialize in its production. Private goods are thus provided by a continuum of symmetric duopolies. The blue firms have aggregate measure one, as do the green firms. The consumers also have measure one, and each small consumer buys exactly one unit of each private good k, from exactly one of the two duopolists selling good k.

The duopolists can buy advertising, measured in impressions. If a duopolist i buys h_i impressions; then an average consumer sees its message h_i times. Consumers see the messages by consuming images. Let q_j denote consumer j's consumption of images, and let q denote the average consumption of images, across all consumers. Until we integrate the two models, we take q_j and q to be exogenous. Assume that duopolist i's message is distributed evenly throughout the images, so that consumer j sees duopolist i's message $h_i q_j/q$

times. Advertising is "persuasive" in the sense that consuming an advertised product yields an unexplained utility bonus. If duopolist i is blue, then assume that consumer j's utility from buying its product is $\Lambda+(hq_j/q)^{\alpha}-r_i$, where $\Lambda>0$ is a fixed constant large enough to ensure that every consumer wants to buy every private good in equilibrium, $\alpha\in(0,1)$ is an exogenous parameter, and r_i is duopolist i's price. If duopolist i is green and producing good k, then consumer j's utility from buying its product is $\Lambda+(h_iq_j/q)^{\alpha}-r_i+\varepsilon_j(k)$, where $\varepsilon_j(k)$ is an idiosyncratic preference shock, distributed uniformly on the interval $[-½,+½]$ and statistically independent across j and k, which consumer j observes before purchase. The random variable $\varepsilon_j(k)$ introduces differentiation between the blue and green firms. Consumer j's total utility from purchasing private goods is thus:

$$U_j = \int_{B_j} [\Lambda+(h_b(k)q_j/q)^{\alpha}-r_b(k)]dk + \int_{G_j} [\Lambda+(h_g(k)q_j/q)^{\alpha}-r_g(k)+\varepsilon_j(k)]dk \;,$$

where: B_j and G_j denote, respectively, the sets of goods that consumer j buys from the blue and green duopolists, with $\{B_j,G_j\}$ a partition of [0,1]; and $r_b(k)$ and $h_b(k)$ are the price and advertising decisions of the blue firm producing the kth good, with symmetric notation for the green firms.

The next step is to derive the Nash equilibrium between each pair of duopolists, for any given good k; we suppress references to k. Assume that the blue duopolist chooses (h_b,r_b) and the green duopolist simultaneously chooses (h_g,r_g). The consumers decide which products to buy after observing these decisions. Assume that $q_j=q$ for almost all consumers j. Then a representative consumer sees the blue firm's message h_b times and the green firm's message h_g times, and she will buy good k from the blue firm if $\varepsilon_j < h_b^{\alpha}-h_g^{\alpha}+r_g-r_b$ or from the green firm if the inequality is reversed. Aggregating over all consumers and invoking the large of law numbers, the blue firm sells $y_b = ½ + h_b^{\alpha}-h_g^{\alpha}+r_g-r_b$ units. Assume that each duopolist must pay p for each impression, where p is exogenous. Then the blue firm's profit is $[y_br_b - ph_b] = [(½+h_b^{\alpha}-h_g^{\alpha}+r_g-r_b)r_b - ph_b]$. The green firm's profit function is similar, and a routine calculation shows that a symmetric equilibrium among the duopolists requires $r_b(k)=r_g(k)=½$ and $h_b(k)=h_g(k)=h^*(p)\equiv[\alpha/(2p)]^{1/(1-\alpha)}$, for each good $k\in[0,1]$.

Because the total measure of the duopolists is two, their aggregate demand for impressions is $2h^*(p) = 2[\alpha/(2p)]^{1/(1-\alpha)}$. Let $m\equiv 2h^*(p)/q$, the number of messages that a consumer sees per unit of information goods consumed. Then $qm = 2[\alpha/(2p)]^{1/(1-\alpha)}$ and the duopolists' inverse demand for impressions is:

(16) $\quad \rho(qm) = [2/(qm)]^{1-\alpha}\alpha/2.$

Given the Nash equilibrium among the duopolists, and consumer j's optimal response to that behavior, consumer j (who benefits from only half of the impressions that she sees) realizes utility

$$:_j = L + [h^*(p)q_j/q]^0 \; ! \; ½ + ¼$$

where "+¼" represents the extra gain from buying green goods when $,_j(k)>0$. Dropping the constant term, which is inconsequential for our purposes, we can express this utility as a function of consumer j's consumption of information goods:

(17) $\quad :(q_jm) \; / \; [q_jm/2]^0$

Call this Model 2, and the main model of the paper Model 1. To embed Model 2 into Model 1, assume that the same population of consumers inhabits each. Each consumer's total utility is the sum of the utilities accrued by buying images in Model 1 and private goods in Model 2; but (17) shows that the latter utility can be expressed as a function of the consumer's access decisions in Model 1 and incorporated into Model 1 by defining : as in (17). The second link between the models is that we take the duopolists of Model 2 to be the advertisers of Model 1, by defining the demand for impressions " as in (16). (Note that (16) and (17) satisfy all of the assumptions on : and " in Model 1.) The timing is: in stage 1, the firms of Model 1 choose (c,m,f,x); in stage 2, the duopolists choose (h,r), the consumers choose q (and hence s) and see their messages, and the market for impressions clears at price p (the quantity of impressions supplied is qm and the quantity demanded is 2h); in stage 3, the consumers buy private goods from the duopolists and realize the :(qm) component of their utility.

The third link between the two models is to calibrate the impression externalities in Model 1, F, to make the welfare function (15) an accurate representation of (total) surplus in the combined model. To do this, we recalculate surplus for the combined model. Ignoring transfers between agents, consumers' total utility in a symmetric equilibrium of the combined model is, from (14):

$$q\$(c,m) \; ! \; 2x I_0^{s/2} \, ((t)dt \; + \; :(qm) \; ! \; 6(q) \, .$$

After ignoring transfers, the only term in the firms' profit function (1) is their real costs !cx. The advertisers (i.e., duopolists) have no revenue or costs other than transfers, because they have no real costs of production. Therefore, total surplus is simply:

$$q\$(c,m) \; ! \; 2x I_0^{s/2} \, ((t)dt \; + \; :(qm) \; ! \; 6(q) \; ! \; cx$$

Calibrating our original definition of welfare, (15), to this expression requires

(18) $\quad F(t) = \, ! \, "(t) = \, ! \, p$ (in equilibrium).

In other words, the negative external impact of consuming advertising equals the entire area under the demand curve for advertising. This is striking, because in this model advertising really does increase consumers' utility from consumption; but the huge negative externality results because the duopolists advertise only to capture sales from each other: one duopolist's gain from advertising is fully offset by another's loss.

The combined model illustrates the application of our welfare results. If f$P, then Proposition 3 shows that the firms inefficiently distort their revenue mix toward advertising revenue.[29] In the economically reasonable case that firms get most of their revenue from advertisers, nonexcludability does not bind, and horizontal utility is concave, Proposition 5 shows that firms overprovide images at the margin.

7. Related literature.

The first study that formally recognizes the link between the markets for broadcasting and advertising appears to be that of Spence and Owen (1973, henceforth SO). SO assume that each monopolistically competitive broadcaster offers a unique and indivisible program, and each consumer accesses at most one of those programs. Advertising plays a simple role: it generates a fixed amount of revenue for each broadcaster, for each consumer who accesses its program. Wildman and Owen (1985) put advertising into consumers' utility functions and allow firms to choose how much advertising to place in their programs, and Masson, Mudambi, and Reynolds (1990) endogenize both price and quantity in the market for advertising, while providing a simpler treatment of the broadcast market.

The small theoretical literature on advertising and public goods (or broadcasting) has devoted the most attention to three (overlapping) issues: what market structure and regulatory environment is most likely to provide diversity in broadcast programs and cater adequately to minority tastes (e.g., Steiner (1952), Beebe (1977)); should broadcasters be allowed to set positive access fees (e.g.Minasian (1964), Buchanan (1967), Holden (1993), Doyle (1998)), Hansen and Kyhl (2001)); and what are the general impacts of market structure (e.g., Masson, Mudambi, and Reynolds (1990)). In wide-ranging studies, Spence and Owen (1977), Owen and Wildman (1992), and Anderson and Coate (1999) investigate all of these issues. Some recent manuscripts look at other issues. Gabszewicz, Laussel, and Sonnac (1999) study the impact of regulatory limits on message density. Hackner and Nyberg (2000) assume that consumers get positive utility from ads and study the asymmetric and monopolistic equilibria that can arise from network effects in a newspaper

---

[29] If f<P then, given the large negative externality attached to advertising, it is no longer true that the information good is underconsumed, so the access-promoting policies of Propositions 2 and 3 are no longer clearly desirable.

duopoly.[30] Dukes (1999), who provides a complete structural model of the demand for advertising, studies how the degree of substitutability in the media and product markets affects diversity, message density, and other quantities.[31]

The rest of this survey focuses on several welfare results, in other studies, which are related to the questions addressed in our results, especially Proposition 2. Hansen and Kyhl (2001, henceforth HK) study a monopoly broadcaster who sells access to an indivisible nonrival good, and who can sell advertising at a fixed price.[32] Following Holden (1993), HK address the policy question of whether it makes sense to force the monopolist to set f=0 (in our notation). HK show, as a side result, that if the monopolist maximizes profits at (m,f), with f>0, and (m*,f*) denotes the *global* first best, then f*=0 and m*<m if the good is nonincludable; but f*<0 and m*=m if the good is includable (using our terminology). The result that a monopolist puts more advertising in a nonincludable good than would occur at the global first best complements, at some distance, our finding that small firms choose excessively high message density at the margin, regardless of includability. Both of HK's results for m* also depend on the assumption that the demand for advertising is perfectly elastic.[33] HK's assertion that the first-best access fee is negative does not depend on a perfectly elastic demand for advertising, and their study appears to be the first to make this (relatively robust) point.

The earlier literature on advertising and public goods, at least through Spence and Owen (1973), seems to assume that optimal f=0, applying the ordinary intuition from nonrival goods. Given this

[30] Blair and Romano (1993, BR) and Rysman (2002) study related models of positive feedback between circulation and ad volume, assuming that the marginal cost of distributing the media to one more consumer is strictly positive. BR study a monopoly and Rysman studies an oligopoly. Rysman also allows one consumer to access multiple media. None of these papers adopt the assumption, common to all the other papers cited (and the present model), that advertising can be measured in impressions: in the "newspaper" models, doubling circulation and doubling advertisements may have different impacts on advertising revenue.

[31] Brown and Cave (1992) provide a useful informal survey of theoretical and empirical issues in the market for broadcasting.

[32] HK do not actually assume that price of advertising is fixed, but they assume that the revenue that the monopolist gets from advertising is proportional to the number of viewers, and a perfectly elastic demand for advertising is the only obvious interpretation of that assumption (unless, perhaps, the monopolist is advertising its own product). Of course, for a single firm, the assumption that the demand for its ads is perfectly elastic may be economically reasonable.

[33] Anderson and Coate (1999) point out that the m*<m result depends on HK's assumption that the advertisers earn zero surplus, a similar point.

assumption, it is immediate that any strictly positive access fee is socially excessive.[34] Our claim that the equilibrium access fee is too high would thus surprise none of these early writers, in the case that f>0.

Given an exogenous message density and price of advertising, Spence and Owen (1977) show informally that the entry process is biased in favor of broadcasters offering low-cost programs, relative to what would maximize social surplus. This result is consistent with our finding that the market underprovides quality at the margin, but SO obtain their result from a special demand structure: they assume that there is no subsitutability between programs, making each broadcaster effectively a monopolist, and that the demand for each program is geometric. SO provide no argument or intuition for a more general result.[35]

To the author's knowledge, the only other study that endogenizes all prices and quantities in the two markets is Anderson and Coate (1999, henceforth AC). Most of AC's results concern pure public goods, but they briefly introduce the fifth quantity, an access fee, at the end of the paper. AC study a duopoly, with discrete output and consumption in the fashion of Steiner (1953). AC address issues mostly different from the issues addressed here, but they do describe conditions sufficient for under- and overproduction of pure public goods (e.g., only one of their duopolists produces when both should, or both produce when only one should). These conditions focus on different quantities than do our Propositions 4 and 5. Roughly speaking, AC find that overproduction is more likely when the demand for advertising (similar to our α) is high, the nuisance cost of advertising (our β_2) is high, and consumer valuation and horizontal penalties (our β and $\gamma\phi$) are low. AC take most of these quantities to be exogenous rather than determined in equilibrium.

8. Summary and concluding remarks.

Markets for information goods are complicated, but we have taken an early step toward understanding them by making (at least) one major simplification. Following the traditions of market theory, we have studied the benchmark case in which agents are numerous enough to eliminate strategic interactions. Even then, however, the elementary properties of information goods make market power almost inevitable. Viewed as a model of the internet, we have made other simplifications. We have ignored: the internet's important role in organizing information through hierarchies, horizontal links, and search engines; the provision of

---

[34] Minisian (1964) argues that f>0 may be more efficient than f=0, because f=0 means that the market provides no price signals to guide output decisions.

[35] SO also study the question of the market over- or underprovides broadcasts, but their comments on this topic are too diverse to permit easy summary.

certain common (e.g., travel, auction) services, which may be hard to characterize as information goods; the bundling of goods that often occurs when access is not free.

We have focused on the role of advertising in correcting what otherwise seems to be the inevitable underprovision of information goods. The underprovision question has two dimensions, familiar from the theory of nonrival goods: is a sufficient quantity of goods produced, and are sufficiently many consumers given access to whatever quantity is produced?

We have shown that the advertising subsidy can correct underproduction but need not do so, and it is unlikely to lead to the optimal quantity of output. We have described circumstances that tend to lead to over- and underproduction.

The advertising subsidy usually cannot correct underconsumption. This leads to our robust conclusions that, on the margin, information goods contain advertising messages that are not only irritating (the model predicts that they must be irritating), but too many messages and too little investment in quality content, from the viewpoint of maximizing total surplus. The excess advertising result is unrelated to traditional arguments about "wasteful" advertising, which appear as negative impression externalities in the model. The model accommodates such externalities but they do not drive the result.

We have also shown that firms that can set non-zero access fees generally choose a socially inefficient revenue mix: firms could increase total surplus by making a budget-balanced shift away from advertising revenue (which encourages consumption) and toward fee revenue (which discourages consumption). This result stems from the joint production of an information good and the private good of advertising impressions, which creates the possibility of increasing surplus by shifting revenue production from one margin to the other. The mechanism for producing the gain is that reducing the supply of advertising amounts to collusion against both advertisers and consumers, though the negative shadow price of advertising for consumers and the budget-balanced fee adjustment imply in combination that the collusive surplus goes back to consumers and encourages consumption.

When access fees are zero, some of our conclusions may be overturned, depending upon whether fees are constrained from above or from below and how tightly they are constrained. We have shown that this can be determined empirically, if it is possible to measure what a consumer would pay to avoid seeing the marginal advertising message. Comparing this to the market price of advertising impressions reveals whether the zero-price constraint binds from above or below.

CITATIONS

S. Anderson and S. Coate, "Market Provision and Public Goods: The Case of Broadcasting," NBER WP#7513 (1999, revised 2001).

J. Beebe, "Institutional Structure and Program Choices in Television Markets," *Quarterly Journal of Economics* (1977) 91:15-37.

R. Blair and R. Romano, "Pricing Decisions of the Newspaper Monopolist," *Southern Economic Journal* (1993), 59:721-32.

A. Brown and M. Cave, "The Economics of Television Regulation: A Survey with Application to Australia," *Economic Record* (1992) 68:377-94.

J. Buchanan, "Public Goods in Theory and Practice: A Note on the Minasian/ Samuelson Discussion," *Journal of Law and Economics* (1967) 10:193-97.

C. Doyle, "Programming in a Competitive Broadcasting Market: Entry, Welfare, and Regulation," *Information Economics and Policy* (1998) 10:23-39.

A. Dukes, "The Advertising Market in a Product Oligopoly," manuscript, 2000.

F. Fisher and J. McGowan, "Advertising and Welfare: Comment," *Bell Journal of Economics* (1979) 10:726-27.

J. Gabszewicz, D. Laussel, N. Sonnac, "TV-Broadcasting Competition and Advertising," manuscript, 1999.

J. Häckner and S. Nyberg, "Price Competition, Advertising, and Media Market Concentration," manuscript, 2000.

C. T. Hansen and S. Kyhl, “Pay-per-view Broadcasting of Outstanding Events: Consequences of a Ban” *International Journal of Industrial Organization* (2001) 19:589-609.

S. Holden, “Network or pay-per-view?” *Economics Letters* (1993) 43:59-64.

R. Masson, R. Mudambi, R. Reynolds, “Oligopoly in Advertiser-Supported Media” (1990) *Quarterly Review of Economics and Business* 30(2):1-16.

J. Minasian, “Television Pricing and the Theory of Public Goods,” (1964) *Journal of Law and Economics* 7:71-80.

B. Owen and S. Wildman, *Video Economics*, Harvard University Press, 1992.

M. Rysman, “Competition Between Networks: A Study of the Market for Yellow Pages,” (2004) *Review of Economic Studies*, 71:483-512.

P. Samuelson, “The Pure Theory of Public Expenditure,” *Review of Economics and Statistics* (1964).

M. Spence, “Product Selection, Fixed Costs, and Monopolistic Competition,” *Review of Economic Studies* (1976) 43:217-35.

M. Spence and B. Owen, “Television Programming, Monopolistic Competition, and Welfare,” *Quarterly Journal of Economics* (1977) 91:103-26.

M. Stegeman, “Competitive Provision of Nonrival and Pure Public Goods Through Advertising,”, manuscript, 2002.

P. Steiner, “Program Patterns and Preferences, and the Workability of Competition in Radio Broadcasting,” *Quarterly Journal of Economics* (1952) 66:194-223.

S. Wildman and B. Owen, "Program Competition, Diversity, and Multi-Channel Bundling in the New Video Age," in *Video Media Competition: Regulation, Economics, and Technology* (E. Noam, ed.), Columbia U. Press, 1985.

APPENDIX

This appendix comprises a more complete development of the demand function (6) and statements and proofs of the main results. To derive (6), the first step is to propose a natural solution to the consumer's access decision if all firms choose common values of (c,m,f,x). Then consumer j's problem is to choose A_j to maximize

$$U_j = x\text{m}_{A_j}[\$(c,m) ! ((*t*) ! f]dt \ + \ :(q_j m) \ ! \ 6(q_j) .$$

Recall that $q_j / x \text{I}_{A_j}$ dt represents consumer j's total consumption. The natural solution to this problem is to access the products of an interval of firms: A_j=(! q_j/(2x),q_j/(2x)) for some q_j. After making this substitution, MU_j/Mq_j = $(c,m) ! f ! ((q_j/(2x))+m : N(q_j m) ! 6N(q_j) and MU_j^2/Mq_j^2 = m^2 : O(q_j m) ! (N/(2x) ! 6O(q_j). Our assumptions : #0, (N>0, and 6O>0 imply that MU_j^2/Mq_j^2 <0, implying that the unique optimal q_j satisfies

(19) $(c,m) ! ((q_j/(2x)) ! f+m : N(q_j m) ! 6N(q_j) = 0 if q_jO(0,x).

Now suppose that a single firm t deviates from the common profile (c,m,f,x). Because the deviating firm is negligible, consumer j's total consumption q_j should remain fixed. To decide which goods to consume, we propose that the consumer consider the marginal impact of consuming a small discrete quantity of each firm's good. If, in addition to her pre-deviation access plan A_j she consumes) discrete units of the good produced by firm t, where ($\hat{c}$,$\hat{m}$,$\hat{f}$) denotes firm t's choices, then her new utility is:

$$\text{m}_{A_j}[\$(c,m) ! ((*t*) ! f]xdt \ + \ [\$(\hat{c},\hat{m}) ! ((*t*) ! \hat{f}]) \ ! \ 6(q_j +)) \ + \ :(q_j m +)\hat{m}).$$

Differentiating this expression with respect to) and evaluating at)=0 measures the incremental value of accessing firm t's good: $($\hat{c}$,$\hat{m}$) ! ((*t*) ! $\hat{f}$! 6N(q_j) + $\hat{m}$: N(q_jm). (This would be the relevant test in the discrete analog of the present model, with each firm producing an arbitrarily small fixed quantity of output.) If the consumer accesses the goods of firms for which this expression is positive, then she accesses the good of firm t if and only if

$$*t* < .[\$(\hat{c},\hat{m}) ! \hat{f} + \hat{m} : N(q_j m) \ ! \ 6N(q_j)]$$

(Recall from Section that .=(!1 with an extended domain.) If all consumers j choose the same q_j=q, as in equilibrium, then this implies that the fraction of consumers who access the firm's good is (6), as posited in the text. For all firms but the negligible deviator, (6) implies that the access ratio is s = 2.[$(c,m) ! f+m : N(qm) ! 6N(q)], implying ((s/2) = $(c,m) ! f+m : N(qm) ! 6N(q), implying that each consumer is indeed consuming the quantity implied by (19), as we originally observed must be true if the consumer's firm-by-firm decision rule aggregates to optimal total consumption.

Proof of Lemma 0. Suppose x>0. Definition (6), :0#0, and 60>0 imply that M[x$(c,m,f;q,m)!q]/Mq<0 (taking (c,m,f,x) to be fixed). Because $(c,m,f;q,m)0[0,1], q=0 implies x$(c,m,f;q,m)!q$0, and q=x implies x$(c,m,f;q,m)!q#0. Therefore , x$(c,m,f;q,m)!q = 0 for exactly one q*0[0,x], and let s*/q*/x. Alternatively, if x=0, then let q*=0 and s*=$(c,m,f;0,m). Let p*/"(q*m). Then (q*,s*,p*) is the unique solution to (2), (7), and (8). ☐

Except for Proposition 2B, every remaining result describes the properties of a consumer equilibrium. To lighten notation, we sometimes suppress the arguments of "(qm), F(qm), $(c,m), ((s/2), .(((s/2)), :(qm), 6(q), B($,m,f;p,q,m), W(q;c,m,x), and their derivatives, but these functions are always evaluated at the equilibrium point.

Lemma 1. Consider any interior outcome that is a full equilibrium.

A. f...0! iff MB/Mf$0 iff $$\$_2$$+:N+p#0.

B. f...0+ iff MB/Mf#0 iff $$\$_2$$+:N+p$0.

Proof. From (1), (6), (8) and (10):

(A.1) MB/Mf = $!(mp+$!P)2.N ;

(A.2) MB/Mm = $p+($$\$_2$$+:N)(mp+$!P)2.N = 0 ;

(A.3) MB/M$ = (mp+$!P)2$$\$_1$$.N ! 1 = 0 .

By definition (Section 5), f...0! iff either f...0 or MB/Mf$0, but if f...0 then individual firm optimization (condition (10)) requires MB/Mf=0. Therefore, f...0! iff MB/Mf$0. Statement (13) and MB/Mm=0 imply (mp+$!P)2.N = !$p/($$\$_2$$+:N). Therefore, using (A.1) and (13): MB/Mf$0 iff $ $ (mp+$!P)2.N iff $ $!$p/($$\$_2$$+:N) iff $$\$_2$$+:N+p#0. That establishes A. The proof of B is similar. ☐

Proposition 1. Consider any interior outcome that is a full equilibrium. Then (12) holds.

Proof of Proposition 1. Conditions (12a)-(12d) are immediate from (2), (7), (8), and (9). Condition (12e) comes from (A.2) and (A.3), and (12f) comes from (A1.2). It remains to show (12g). If f...0, then each firm's choice of f is locally unconstrained, implying MB/Mf=0. (A1.1) and (A1.2) then imply $$\$_2$$+:N+p=0. If f=0 and goods are excludable, then it must be that MB/Mf#0, because otherwise the individual firm would increase $\hat{f}$, and Lemma 1 then implies $\beta_2+\mu'+p\geq 0$. If f=0 and goods are includable, then the argument is similar. ☐

When x>0, it simplifies matters to use q=sx (12c) to eliminate s from (15), the definition of surplus. Noting that neither f nor p has any direct impact on surplus, x>0 implies:

(15′) W(q;c,m,x)

$$= \int_0^{mq}[\alpha(t)+\sigma(t)]dt - cx + [\beta(c,m)-\pi]q - 2x\int_0^{q/(2x)} \gamma(t)dt - 6(q) + \mu(qm).$$

The assumptions underlying the statics results follow.

Proposition	**Exogenous Variables**	**Endogenous Variables**	**Governing Equations**
2, 4, 5, Lemmas 2, 3	c, f, m, x	q, s, p, W	(12a)-(12c), (15′)
3	c, m, x	q, s, p, f, W	(12a)-(12d), (15′)

Lemma 2. Consider any interior outcome that is a consumer equilibrium. Then dq/dc>0, dq/df<0, and dq/dm has the sign of $\beta_2+\mu'+qm\mu O$.

Also, the derivatives of the welfare function (15′) satisfy:

$\partial W/\partial q = (p+\sigma)m - \pi + f$

$\partial W/\partial c = q\beta_1 - x$

$\partial W/\partial m = (p + \sigma + \beta_2 + \mu')q$

Proof. Let

$\rho(q;f,c,m,x) \equiv q - 2x.(\beta(c,m) + m\mu'(qm) - f - 6'(q))$

Condition (1a) implies that $\rho(q;f,c,m,x)=0$ in any consumer equilibrium. Because $\partial\rho/\partial q = 1 + (60 - m^2\mu O)2x.' > 0$, the implicit function theorem implies that, at equilibrium:

dq/dc $= -[\partial\rho/\partial c]/[\partial\rho/\partial q] = \beta_1 2x.'/[\partial\rho/\partial q] > 0$;

dq/df $= -[\partial\rho/\partial f]/[\partial\rho/\partial q] = -2x.'/[\partial\rho/\partial q] < 0$

dq/dm $= -[\partial\rho/\partial m]/[\partial\rho/\partial q] = (\beta_2+\mu'+qm\mu O)2x.'/[\partial\rho/\partial q]$

The expressions for $\partial W/\partial c$ and $\partial W/\partial m$ come from differentiating (15′) and using (12b) to eliminate α. Similarly, $\partial W/\partial q = m(p+\sigma) + \beta - \pi - \gamma(s/2) - 6' + m\mu'$, which with (12a) implies that $\partial W/\partial q = (p+\sigma)m - \pi+f$. ☐

Proposition 2.

A. Consider any interior outcome that is a full equilibrium. If F> ! p ! (f ! P)/m (which must hold if F$0), then dW/df<0. If, in addition, f…0 ! , then dW/dc>0. If, in addition, F#0, then dW/dm<0.

B. Given fixed (c,m,x), let (q(f),p(f),s(f)) describe the period 2 equilibrium implied by arbitrary f. Assume that f=f* maximizes W(q(f);c,m,x) over f0ú. If s(f*)<1 and W(q(f*);c,m,x)>0, then f* = P ! m[p(f*)+F(mq(f*))].

Proof.

Part A. If F$0, then (12d), c>0, and s>0 imply (p+F)m ! P + f > 0. Lemma 2 implies MW/Mq = (p+F)m ! P + f > 0, by the assumption on F. Equation (15N) implies:

dW/df = MW/Mq×dq/df.

dW/dc = MW/Mq×dq/dc + MW/Mc.

dW/dm = MW/Mq×dq/dm + MW/Mm

Lemma 2 implies dW/df <0. For the rest, assume f…0 ! . Then Lemma 1 implies $\$_2$+ : N+p#0. With (12c) and (12e), this implies q$\$_1$$x. Then Lemma 2 implies MW/Mc = q$\$_1$! x $ 0. Therefore, dW/dc $ MW/Mq×dq/dc, but Lemma 2 implies dq/dc >0, so dW/dc>0. Now assume F#0. Lemma 2 and $\$_2$+ : N+p#0 imply MW/Mm # Fq #0. Therefore, dW/dm # MW/Mq×dq/dm. Because $\$_2$+ : N+qm : 0< $\$_2$+ : N+p # 0, Lemma 2 implies dq/dm<0, so dW/dm < 0.

Part B. Equation (15) shows that W#0 in any period 2 equilibrium such that q=0. Therefore q(f*)>0, so (12c) implies x>0 and we can use (15N) to describe W. Given (12c), s(f*)<1 implies q(f*)0(0,x), and Lemma 2 shows that dq/df < 0 (the proof of result required only q to be "interior"). Therefore, f* a maximizer of W(q(f);c,m,x) requires dW/df = MW/Mq×dq/df = 0, which requires MW/Mq=0, which from Lemma 2 implies [p(f*)+F]m ! P+f*=0. ~

Proposition 3. Consider any interior outcome that is a full equilibrium, with f…0 ! . Assume F(qm)0[! p ! (f ! P)/m,0]. Then, taking f to be endogenous as described in the table above, and disregarding any constraint on f (i.e., f$0 or f#0): dW/dm<0.

Proof. Lemma 2 implies that MW/Mq = (p+F)m ! P+f, so F$! p ! (f ! P)/m implies MW/Mq$0. The proof of Proposition 2A shows that f…0 ! implies MW/Mm # Fq, so F#0 implies MW/Mm#0. If F<0, then MW/Mm<0; if F=0, then pm ! P+f>0 (from (13)) implies MW/Mq>0. Therefore, to show that dW/dm = MW/Mq×dq/dm + MW/Mm < 0, it is sufficient to show that dq/dm<0. Using (12c) to eliminate s from (12a), (12b), and (12d) leaves the following system:

2x.($(c,m)+m:N(qm)!f!6N(q)) !q = 0.

"(qm) ! p = 0.

(mp + f!P)q ! cx = 0

Invoking the implicit function theorem in the usual way, total differentiation of the 3×3 system shows that dq/dm = T_1(q,p,f;m,c,x)/T_2(q,p,f;m,c,x), where

T_1(q,p,f;m,c,x) / !2.Nxq[p+$\$_2$+:N+qm("N+:O)]

T_2(q,p,f;m,c,x) / 2.Nx[mp+f!P+qm^2("N+:O)!q6O] !q

Given f...0!, Lemma 1 implies p+$\$_2$+:N#0, which with (12c) and (12f) implies q $ 2.Nx(mp+f!P). After these substitutions:

T_1(q,p,f;m,c,x) $!2.Nxq^{2}m("N+:O) > 0

T_2(q,p,f;m,c,x) # 2.Nxq[m^2("N+:O)!6O] < 0

Therefore, dq/dm < 0. ~

Our analysis of dW/dx begins with a lemma, which shows that the sign of dW/dx is connected to the sign of:

N(f,c,m,x,p,q,s) / [(p+F)m+f!P]/[1+2(6O!m^2:O)x.N] ! (pm+f!P)[1+($\$_2$+:N)/(2p)]

Suppress arguments of N(f,c,m,x,p,q,s),

Proofs after here may need checking again, some changes seem lost.

Lemma 3. Consider any interior outcome that is a full equilibrium. Then:.

dW/dx # sN(f,c,m,x,p,q,s) if (is convex.

dW/dx $ sN(f,c,m,x,p,q,s) if (is concave.

Proof. To study MW/Mx, we define a linear approximation of (. Given the equilibrium value of s, define (*:[0,s/2]6ú by (*(t) / ((s/2) + (N(s/2)(t!s/2). Then (*(s/2)=((s/2) and (*N(t)=(N(s/2) for all t. Clearly: if (O(t)$0 for all t, then I$_0^{s/2}$(*(t)dt#I$_0^{s/2}$((t)dt; if (O(t)#0 for all t, then I$_0^{s/2}$(*(t)dt$I$_0^{s/2}$((t)dt. Suppose (O(t)$0 for all t. Then, from (15N):

MW/Mx	= s((s/2)!2I$_0^{s/2}$((t)dt ! c	(using (12c))
	# s((s/2)!2I$_0^{s/2}$(*(t)dt ! c	
	= (N(s/2)s^2/4 ! c	(by the definition of (*)
	= s^2/[4.N(((s/2))] ! c	(because (=.$^{!1}$)
	= !(mp+f!P)[1+($\$_2$+:N)/(2p)]s.	(by (12d) and (12f)).

If (O(t)#0 for all t, then the inequality reverses. We have thus shown:

MW/Mx # !(mp+f!P)[(1+($\$_2$+:N)/(2p)]s if (O(t)$\$$0 for all t;

MW/Mx $\$$!(mp+f!P)[(1+($\$_2$+:N)/(2p)]s if (O(t)#0 for all t.

Parallel to the proof of Proposition 2A, equation (15N) implies:

dW/dx = MW/Mq×dq/dx + MW/Mx .

Lemma 2 shows that MW/Mq = (p+F)m ! P + f . Parallel to the proof of Lemma 2, dq/dx = ![MR/Mx]/[MR/Mq] = 2./[1+2(60!m^2:O)x.N]. Putting all this together, and using (12a), shows that if (is convex then:

dW/dx # !(mp+f!P)[1+($\$_2$+:N)/(2p)]s + [(p+F)m!P+f]s/[1+2(60!m^2:O)x.N]

= sN(f,c,m,x,P,q,s)

if (is convex. The claim for (concave follows similarly. ~

Proposition 4. Consider any interior outcome that is a full equilibrium. If (is convex, F#0, f...0+ (equivalently p$\$$*$\$_2$+:N*), and x > (N/[2(60!$m^2$:O)], then dW/dx<0. If (is concave, F$\$$0, f...0! (equivalently p#*$\$_2$+:N*), and x < (N/[2(60!$m^2$:O)], then dW/dx>0.

Proof. The claims of equivalence come from Lemma 1, given (13). Suppose (is convex, F#0, f...0+, and x > (N/[2(60!m^2:O)]. Then !($\$_2$+:N)/p#1, and (12d) implies mp+f!P>0, implying that N # [mp+f!P][1/[1+2(60!m^2:O)x.N] !½], and x > (N/[2(60!m^2:O)] implies 2(60!m^2:O)x.N>1. Therefore, N < 0, and Lemma 3 then implies the result. If (is concave, F$\$$0, f...0!, and x < (N/[2(60!m^2:O)], then a symmetric argument shows that dW/dx>0. ☐

Proposition 5. Consider any interior outcome that is a full equilibrium, with f...0+. If (is convex and [!F/p]×[pm/(pm+f!P)] $\$$ ½, then dW/dx<0.

Proof. The inequality implies (p+F)m+f!P # (pm+f!P)(1!½), and f...0+ implies (from Lemma 1) !½ # ($\$_2$+:N)/(2p), so with pm+f!P>0 (from (13)) we have (p+F)m+f!P # (pm+f!P)/2 # (pm+f!P)[1+($\$_2$+:N)/(2p)]. Because the right-hand side is strictly positive, it follows that N<0, and Lemma 3 then implies the result. ☐