



ÉCOLE POLYTECHNIQUE



CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

COMPETITION FOR ATTENTION IN THE INFORMATION
(OVERLOAD) AGE

Simon P. ANDERSON
André DE PALMA

March 2006, revised August 2010

Cahier n° 2010-23

DEPARTEMENT D'ECONOMIE

Route de Saclay

91128 PALAISEAU CEDEX

(33) 1 69333033

<http://www.economie.polytechnique.edu/>
<mailto:chantal.poujouly@polytechnique.edu>

Competition for attention in the information (overload) age

Simon P. Anderson and André de Palma*

March 2006, revised August 2010.

Abstract

Limited consumer attention limits product market competition: prices are stochastically lower the more attention is paid. Ads compete to be the lowest price in a sector but compete for attention with ads from other sectors: equilibrium ad shares follow a CES form. When a sector gets more profitable, its advertising expands: others lose ad market share. The “information hump” shows highest ad levels for intermediate attention levels. The Information Age takes off when the number of viable sectors grows, but total ad volume reaches an upper limit. Overall, advertising is excessive, though the allocation across sectors is optimal.

JEL CLASSIFICATION: D11, D60, L13, M37.

KEYWORDS: economics of attention, information age, price dispersion, advertising distribution, consumer attention, information filtering, size distribution of firms, CES, information congestion.

*Simon P. Anderson: Department of Economics, University of Virginia, PO Box 400182, Charlottesville VA 22904-4128, USA, sa9w@virginia.edu. André de Palma: Département Economie et Gestion, Ecole Normale Supérieure, 61 Ave du Président Wilson, 91235 FRANCE, and Ecole Polytechnique. andre.depalma@ens-cachan.fr. The first author gratefully acknowledges research funding from the NSF under grant SES-0752923 and from the Bankard Fund at the University of Virginia. The second author thanks Institut Universitaire de France. We thank the Autoridade da Concorrença in Lisbon for its gracious hospitality, and the Portuguese-American Foundation for support. Comments from conference participants at EARIE 2007 (Valencia), Intertic-Milan (2008), and the Economics of Advertising Conference in Bad Homberg (2008), the CITE conference on Information and Innovation in Melbourne (2009) and seminar participants at the Sauder School (UBC), Stern School (NYU), National University of Singapore, James Madison University, Catholic University of Louvain (KUL), the Universities of Oklahoma, New South Wales, Copenhagen, and North Carolina (Chapel Hill) are gratefully acknowledged.

1 Introduction

According to a Wiki cite, perhaps the first academic to articulate the concept of attention economics was Herbert Simon when he wrote

...in an information-rich world, the wealth of information means a dearth of something else: a scarcity of whatever it is that information consumes. What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it. (Simon 1971, p. 40-41).

This is echoed in Lanham (2006), in the idea that we are drowning in information, but short of the attention to make sense of that information.¹ Our interest in this paper is in turning around Simon's point and looking how restricted attention affects the market information provided. In particular, we look at competition between firms providing information in the form of ads for their products. Facing limited attention, a firm might try and get away with a high price in the hope that its competitors' ads about their lower prices has been crowded out from the information receiver's attention span. This leads us to consider price dispersion in the face of endogenous congestion where information from each sector competes within the sector and with each other sector (even though sectors do not directly compete except for attention). We track several overall dimensions of the economics of information overload, advertising volume and clutter, and price dispersion (some results are in the Abstract).

The Information Age has seen the means of reaching people increase enormously, with a corresponding drop in the costs of doing so. In response, the amount of advertising has gone up drastically. Shenk (1997) states that the average American encountered 560 daily advertising messages in 1971, and over 3,000 per day by 1997.² At the same time though, consumer retention rates for ads remains low, perhaps

¹See Eppler and Mengis (2004) for a multi-disciplinary review of Information Overload.

²Current exposure levels are a matter of considerable debate and estimates range from 245 through 5000 exposures a day: see e.g. <http://www.amic.com/guru/results.asp?words=media+exposure&submit=Search&op=AND> Skeptics point out that one would have to see an ad every 10 seconds for 8 hours to reach 3000 ads per day (although, in riposte, people are now spending *nine* hours a day with media!). However, web-sites can easily include 10 ads, and by one estimate we view 2000 web-sites a day. Likewise, commuting in a built-up area gives multiple bill-boards and signs per minute, not to mention the number of ads per page in newspapers.

one ad a day (the reader can ask himself or herself how many ads s/he remembers from yesterday!) Perhaps not surprisingly, the costs of reaching prospective consumers is low, and varies from \$6 CPM (cost per thousand impressions) for Internet, through \$18 for network TV and \$26 for newspapers (for 2006, http://www.wikinvest.com/concept/Impact_of_Internet_Advertising),³ and, of course, spam email is very low cost. This picture of many messages relative to retention underscores our modeling approach.

The Information Age comes from several sources, primarily lower costs of sending information as more (and cheaper) channels now reach potential consumers. Traditional billboards and newspaper ads have been supplemented by Internet pop-ups, telemarketing, and product placements within TV programs (and on football players' jerseys). Information costs have not been lowered uniformly across the board, though, and some sectors' messages are more appropriately delivered by the new media. However, cheaper access to attention also means that rivals can access attention more cheaply too, intensifying in-sector competition. This effect renders competition more acute, lowering prices and benefiting consumers. Scarcity of attention brings spillovers into other sectors, raising their prices and making it more likely interesting offers are missed.

New product categories cause pricing churn for other advertised goods. A new product class depresses existing classes' ads relatively (as a fraction of the total volume of messages), and it drives down weaker ones absolutely. Surprisingly, it may cause stronger sectors to increase in size because price competition is relaxed (prices are stochastically higher). Thus there are information complementarities across product classes.

The third effect we track is the attention span of consumers. Since both work and leisure time are spent increasingly on information-carrying activities,⁴ it is likely that consumer attention spans (the amount of ads that can be absorbed) have risen. This may induce more or less information transmission. When consumer attention is sparse, little information will be sent because there is not much chance of getting a look-in. Prices will be near monopoly levels because there is little chance of running across a rival. With a lot of

³Here are CPMs for 2002 (http://bpsoutdoor.com/article.php?article=how_effective):

Billboard (30-Sheet Poster), \$2.05; Radio Ad (During Prime Drive Time), \$8.61; Magazine (One page with 4 colors), \$9.35; Television Commercial (30 Seconds on a Prime-Time Network), \$17.78; Newspaper (1/3 of a page in black and white), \$22.95.

⁴Indeed: "While television is still by far the dominant medium in terms of the time average Americans spend daily with media at 240.9 minutes, the computer has emerged as the second most significant media device at about 120 minutes." from http://www.marketingvox.com/study_most_waking_hours_spent_with_media-020005/ who also say: "The average person spends about nine hours a day using some type of media, which is arguably in excess of anything we would have envisioned 10 years ago."

attention, not much information is sent because there is a good chance the consumer will get a better offer from the same sector. Prices will be low, so the benefit from sending a message is low. The middle ground - the “information hump” - is the fertile ground for messages, yielding a fair shot at making a sale at a reasonably high price, both by being seen but no rival from the same sector being found.

We also track the distribution of messages across sectors. With low levels of attention, highly profitable sectors will be most prominently represented. Increasing consumer attention brings firms into more competition with each other, which drives down sector profitability and serves to equalize opportunities across sectors while generally lowering mark-ups. Improved communication costs in specific sectors lower prices there, though the extra crowding can relax competition (and raise prices) in other sectors.

The framework we use to model firms’ actions is adapted from Butters’ (1977) seminal work on informative advertising, which is remarkable for delivering a tractable and intuitive description of equilibrium price dispersion. Butters derives a density of advertised prices and sales prices; he proposes a monopolistic competition framework distinct from that of Chamberlin (1933). In both the Butters and Chamberlinian formulations of monopolistic competition, the competitive part comes from a free-entry zero profit condition that closes the model. The monopolist part in Chamberlin’s work comes from heterogeneity of the products sold by firms; in Butters it comes from the market power that firms have due to imperfect information that consumers do not know all firms’ prices.

We meld Butters’ approach with the advertising clutter approach formalized in Van Zandt (2005) and Anderson and de Palma (2009). Reception of messages is passive: the consumer does not search. This contrasts with Baye and Morgan (2001, 2002) who consider active consumers choosing whether to visit a price comparison site, and a “gatekeeper” firm charging access fees (with a single product sector which generates price dispersion from a mixed strategy equilibrium). Our context corresponds to passively getting messages from bulk mail, from the television, from billboards, etc. We focus on the interaction of multiple industries competing for individuals’ attention. While Butters generates price dispersion because each individual gets only a subset of the price messages (likewise Baye and Morgan, 2001, insofar as firms play with positive probability the option of not posting a price on the comparison site), we suppose that the individual misses

some of the messages sent. This reflects advertising clutter because an individual is bombarded by too many messages (in “junk” mail, billboards, television, and internet pop-ups) to pay full attention to all.

Several authors model both the consumer’s choice of how much attention to supply and the actions of firms vying for that attention by sending messages advertising their wares. In Falkinger (2008), consumers can choose how tight to make message filters but have a limited attention capacity. In Anderson and de Palma (2009), consumers have congested attention spans because they choose how many messages to examine. In Johnson (2010), consumers choose whether to examine all messages or to block them all.

A common feature of these models is that the consumer’s attention is a common property resource insofar as a message sender ignores the effects of its own message on other senders. This means there is a congestion externality, and a tax on messages can improve the allocation of resources.⁵ One concern with this conclusion is that direct business-stealing effects are closed down: message senders do not compete directly in the marketplace, they just compete for attention. A tax might *a priori* reduce price competition by reducing message volume, and so harm consumer welfare. We investigate this question by specifically modeling competition within each of several sectors vying for consumer attention. The focus on firm competition necessitates simplifying the consumer side of the model: it is assumed here the consumer’s attention span is fixed outside the model.

We first characterize an equilibrium model with interaction both within and across sectors. Competition within a sector means that a lower price is more likely to be the lowest sector offer in the set of messages the consumer has screened. Nonetheless, higher-price senders can remain in equilibrium: there is a trade-off between sales probability and mark-up, so all can earn zero profits despite price dispersion. Competition among sectors (industries) comes from overall competition for consumer attention, and price dispersion in each sector depends on all other active sectors.

Surprisingly, the model endogenously generates an inverse IIA property for sector message fractions, and a CES form for the total number of messages sent. This bears an intriguing parallel to the CES utility functional form so often used to parameterize Chamberlinian models. Information congestion gives a new

⁵However, if the consumer’s attention is not congested, a tax may worsen the allocation insofar as message senders do not internalize the consumer surplus from contacting prospective clients.

rationale for the CES specification, but it is now coupled with price dispersion within multiple sectors.

The model also generates a different welfare prescription from Butters (1977). While Butters' model has the optimal and equilibrium level of information equal, we find that the market allocation can be improved by taxing messages.⁶ This reflects the property that advertising is excessive, in contrast to most of the theoretical economics literature on the subject (see Bagwell, 2007, for a survey). Indeed, the standard result in the economics of informative advertising is that there is not enough advertising because firms do not capture the consumer surplus. This is the monopoly result (see Shapiro, 1981, for example). Under oligopoly, this is somewhat offset by business stealing: overadvertising arises in the Grossman and Shapiro (1983) model of informative advertising when the business stealing effect outweighs the consumer surplus one.⁷ Along similar lines, Stegeman (1991) shows that the market advertising is insufficient when the Butters model is amended to allow demand to have some elasticity: firms then tend to overprice without sufficient regard to the consumer surplus lost. In our context, over-advertising is quite natural as it dissipates rents.

The next Section describes the model and solution technique. Section 3 derives the CES form for total advertising and characterizes message volume by sector. Section 4 finds the advertising and sales price distributions by sector, and ties them into the earlier comparative static results. Section 5 sets out the normative properties, the neutrality result that no real changes ensue from transmission cost changes, the optimal allocation property, and the tax prescription to deal with over-advertising. Section 6 allows for distractions, which break the neutrality result, but retain the basic CES form. Section 7 concludes. The Appendix gives a quick reminder of the Butters (1977) model.

2 Message reception and transmission

2.1 Assumptions

There are $\bar{\Theta}$ potential commercial sectors, indexed by $\theta = 1, \dots, \bar{\Theta}$. Each *active* sector θ comprises a continuum of firms of mass n_θ .⁸ Each active firm sends just one message per consumer at a cost γ_θ (which can represent

⁶This finding reinforces the conclusion of Anderson and de Palma (2009) of the desirability of a tax on transmission in the presence of congestion.

⁷Excessive advertising is also found in the controversial Dixit and Norman (1978) paper on persuasive advertising.

⁸In Section 3.4 we discuss how these sectors are endogenously determined.

the cost of a letter, or the cost of a billboard divided by the number of consumers reached).⁹ A message is an (ex-ante anonymous) advertisement containing the price at which a consumer can buy the product from the sending firm. Firms within each sector produce homogenous goods, and each sector therefore transmits n_θ messages, for a total number of $N = \sum_{\theta=1}^{\bar{\theta}} n_\theta$ messages (per consumer). The cost of producing the good advertised in the message is c_θ (which is only incurred if the good is bought – think of a pizza for example): if the good must be produced beforehand regardless of whether the consumer buys, it suffices to set $c_\theta = 0$ and fold the production cost into the transmission cost, γ_θ .

Consumers are assumed to be identical. Messages could be sent to them by bulk mail, by email, or they could be posted on billboards, or on TV programs. However, *reaching* a consumer does not mean the message is *registered*. Each consumer has the same probability of registering a message (which means retaining the price offer). Since we assume constant returns to scale in production (constant marginal costs), we can treat the consumer as the unit of analysis and so we henceforth refer to a single consumer.¹⁰

The consumer registers a fixed number of messages, $\phi \geq 1$, which are drawn at random from the N messages sent. This reflects limited information processing capability. In what follows we will assume that not all messages sent are registered ($\phi < N$) in order to capture advertising clutter / information congestion: this condition will hold in equilibrium under a mild assumption that ensures there are always some inactive potential sectors. After registering the ϕ messages, the consumer makes her purchase decisions. She chooses the lowest priced offer received from each sector (we argue below that the probability of ties is zero) and buys q_θ units if that price is no larger than her reservation price for the sector, b_θ .

The model can also be interpreted as competition with traditional physical stores as follows. A consumer can buy a product in a store, or else she can receive an ad enabling her to buy it cheaper. For advertisers, her reservation price, b_θ , is the full price paid at the store. This full price will include her transportation costs, etc. The consumer may receive unsolicited ads from other sellers; to be entertained they must be priced below her reservation price. Assume that traditional stores are competitive so that the store price is

⁹Indeed, in equilibrium no sender would want to send a second message: to do so would give a negative profit given the original message just made a zero profit under the free-entry assumption below.

¹⁰Allowing for multiple consumer types would be useful for extending the model to analyze consumer targeting.

at the sum of marginal production plus distribution costs. Other sellers may have lower distribution costs (think bricks vs. clicks), and they might deliver the product more cheaply. Assume that consumers do not search out sellers, but they do know about the store option.

Clearly, both types of goods can coexist – some products are available both in stores and through advertised offers; others are not available in stores. The model allows for this by judicious interpretation of the reservation price. In both cases, if an advertised offer is accepted at price p , the consumer surplus ascribed to the advertising sector is $b_\theta - p$.

Finally, we assume that the number of firms in each sector (and the density of messages in a sector at any advertised price) is determined by a zero-profit (free-entry) condition.

2.2 Solution technique

A firm’s expected demand (at any price it may charge) is the probability that its message is registered and it is the lowest price received from its sector. Its expected demand also must satisfy the zero profit condition for the price charged. We equate the probability of making a sale at a particular price from these two different angles to find the relation between the price and the advertised price distribution.¹¹

The highest price set by any firm, b_θ , plays a key role because the only way the sender can avoid a loss at such a price is if it is the only message drawn from that sector. This ties down the number of messages n_θ sent from sector θ as a fraction of the total number of messages sent, N . Summing over sectors yields the total number sent, N , from which we can back out the number in each sector (the n_θ ’s). Armed with that statistic, we can recover the equilibrium price distribution in each sector and its support. This technique also enables us to determine endogenously the equilibrium number of active sectors.

More formally, an equilibrium to the model maps the primitives $\left(\phi, \left\{b_\theta, c_\theta, q_\theta, \gamma_\theta; \theta = 1, \dots, \bar{\Theta}\right\}\right)$ into a set of non-negative sector message numbers $\left(n_1, \dots, n_{\bar{\Theta}}\right)$, which sum to the total message volume N . A sector is active if and only if $n_\theta > 0$. For each active sector, the equilibrium specifies sector purchase probabilities,

¹¹An alternative interpretation is to assume a finite number of firms per sector (see Baye and Morgan, 2001, for a related analysis with a single sector and a gatekeeper). At a symmetric mixed strategy equilibrium conditional on the number of firms entering, all prices played must yield the same profit level. Appending a prior entry game implies that the equilibrium profit level is zero. We have preferred to follow the Butters description with a continuum of firms, not least because the choice probabilities we use are exact (see (1) and the discussion below).

\mathbb{P}_θ , for the consumer, a price distribution within each sector, and corresponding choice probabilities for each product $\mathbb{P}(p, \theta)$. We show that equilibrium is unique, with an endogenous cut-off between active and inactive sectors. We proceed in Lemma 2 by determining message volume by sector as a function of the total message volume, N (to be determined later). We then sum over active sectors in Proposition 5, to find the N consistent with a given number of active sectors. Then, in Proposition 6 we identify the active sectors. Intermediate results describe properties of the solutions.

2.3 Message selection probability

We first seek the probability that a particular message is the only one registered from a sector. Assume that $n_\theta < N$, so at least two sectors are active.¹² Given there is a total mass of N messages, and a finite number of draws, the probability of drawing any given message on a given draw is $\frac{1}{N}$. It is independent of which draw we consider, since the mass of messages left does not noticeably fall with the number of messages already drawn, so the probability of drawing the message is $\frac{\phi}{N}$. By the same token, the mass of firms in sector θ is $n_\theta < N$, so that the probability of drawing a message from a competitor in the same sector is $\frac{n_\theta}{N}$. Hence the probability of avoiding all competitors in the same sector is the probability that no message from the sector is drawn on the other $\phi - 1$ draws, namely $(1 - \frac{n_\theta}{N})^{\phi-1}$. Then

$$\mathbb{P}_\theta = \frac{\phi}{N} \left(1 - \frac{n_\theta}{N}\right)^{\phi-1} \quad (1)$$

is the probability that one (specific) message from sector θ is registered,¹³ and no other message is registered from that sector.¹⁴

¹²As will be seen later, this will be true under mild conditions.

¹³We will later use the notation $\mathbb{P}(p, \theta)$ to denote the probability of a sale at price p in sector θ : hence $\mathbb{P}_\theta = \mathbb{P}(b_\theta, \theta)$, since we shall show that a sale at the top price sent, b_θ , only happens when the message is the only one drawn from the sector.

¹⁴This formula can also be derived as an approximation for a finite number of firms. Suppose that search is with replacement. Then the probability of getting the message on the first draw, and missing the rest of the sector on all subsequent draws is $\frac{1}{N} \left(1 - \frac{n_\theta - 1}{N}\right)^{\phi-1}$. The probability of missing the whole sector on the first $k - 1$ draws, drawing the message on the k^{th} draw and missing the rest of the sector subsequently is $(1 - \frac{n_\theta}{N})^{k-1} \frac{1}{N} \left(1 - \frac{n_\theta - 1}{N}\right)^{\phi-k}$. Thus the chance of getting the message alone is the sum of these events, namely $\sum_{k=1}^{\phi} (1 - \frac{n_\theta}{N})^{k-1} \frac{1}{N} \left(1 - \frac{n_\theta - 1}{N}\right)^{\phi-k}$. This sum simplifies to $\left(1 - \frac{n_\theta - 1}{N}\right)^\phi - (1 - \frac{n_\theta}{N})^\phi$. The first-order Taylor approximation to the first term, $\left(1 - \frac{n_\theta - 1}{N}\right)^\phi$, is $(1 - \frac{n_\theta}{N})^\phi + \frac{1}{N} \phi (1 - \frac{n_\theta}{N})^{\phi-1}$, and so, to the first-order, \mathbb{P}_θ is $\frac{1}{N} \phi (1 - \frac{n_\theta}{N})^{\phi-1}$, which is the expression given in the text.

2.4 Price distribution

Recall that each firm chooses a single price, and there is a continuum of firms in each sector. There can be no equilibrium with all firms choosing the same price (and hence sharing the market): a common price above $c_\theta + \gamma_\theta/q_\theta$ could be profitably undercut; any price $c_\theta + \gamma_\theta/q_\theta$ or below would give negative profits.

We first argue that the support of the equilibrium advertised price distribution (for any firm in active sector θ) is a compact interval $[\underline{p}_\theta, b_\theta]$ with no atoms nor gaps, where the lower bound, \underline{p}_θ , is to be determined below. There are *no atoms* in the price distribution because if there were, any sender choosing the same price as a mass of other senders would raise profits by infinitesimally cutting its price. This would leave its mark-up essentially unchanged but raise sales discretely because it then beats all others at the purported mass point whenever two lowest price messages were the same. The interval has *no gaps* on the support because if there were, the lower price at a gap can be raised leaving the sales probability unchanged but increasing the mark-up. This same argument implies the support must go up to b_θ : if it stopped short, the highest price firm could raise its price with no penalty on sales probability and increase its mark-up. Finally, the lower bound of the support must exceed $c_\theta + \gamma_\theta/q_\theta$ because at any lower price the transmission cost cannot be recouped. It must strictly exceed this bound because there is a positive probability that the message is not read (contrast Butters).

Lemma 1 *Prices in industry θ are distributed on a compact support $[\underline{p}_\theta, b_\theta]$ where $\underline{p}_\theta > c_\theta + \gamma_\theta/q_\theta$, and there are no atoms.*

Let $F(p, \theta)$ denote the *fraction* of messages in sector θ sent at price p or below. (Then $F(\underline{p}_\theta, \theta) = 0$ and $F(b_\theta, \theta) = 1$). A message at price p is successful as long as the price is the lowest one received: using the same logic as used to derive (1), the sales probability is

$$\mathbb{P}(p, \theta) = \frac{\phi}{N} \left(1 - \frac{n_\theta F(p, \theta)}{N} \right)^{\phi-1}, \quad (2)$$

where we simply note that $n_\theta F(p, \theta)$ is the number of messages sent from the sector with a price below p .

We proceed in Section 3 by determining aggregate numbers of messages per sector and total messages, and in Section 4 we derive the price distribution for each sector.

3 Advertising levels

3.1 Advertising shares by sector

We have just argued that some firms will send out an ad at a price equal to the reservation price b_θ , and the probability of finding a second ad at the same price is zero. So consider an advertisement which is sent out with price b_θ . Since \mathbb{P}_θ (as given by (1)) is the probability this is the only ad found from sector θ , the equilibrium zero profit condition (which will tie down n_θ) reads:

$$(b_\theta - c_\theta) q_\theta \mathbb{P}_\theta = \gamma_\theta, \quad (3)$$

where we recall that q_θ is the quantity of good θ demanded. Define π_θ by:

$$\pi_\theta = \frac{(b_\theta - c_\theta) q_\theta}{\gamma_\theta},$$

which measures the economic performance (social surplus per \$ transmission cost) of sector θ . It is necessary (but not sufficient) for an active sector that $\pi_\theta > 1$ because $(b_\theta - c_\theta) q_\theta$ must exceed γ_θ in order for the sender to incur the cost of a message, given that messages are not read with certainty. Indeed, if $\pi_\theta \leq \frac{N}{\phi}$, then even a single message sent from sector θ at the highest price would not be expected to cover its costs: i.e.,

$$(b_\theta - c_\theta) q_\theta \frac{\phi}{N} \leq \gamma_\theta, \quad (4)$$

where $\frac{\phi}{N}$ is the probability the message is registered.¹⁵ The zero profit condition (3) for the equilibrium probability the highest-priced sender in active sector θ makes a sale is then

$$\mathbb{P}_\theta = \frac{1}{\pi_\theta}. \quad (5)$$

This probability depends only on the intrinsic economic performance index, π_θ , of the sector.

Let $\bar{\Theta} \in (1, \bar{\Theta}]$ be the number of sectors for which $\pi_\theta > 1$: this is the *maximum* number of active sectors.¹⁶ We rank these sectors such that π_θ is decreasing in the index θ , i.e. from highest to lowest

¹⁵As we shall see below, this is also the condition for the lowest price in the price support to be below b_θ . (For the lowest price, γ_θ equals the mark-up times the probability of being drawn. The latter is ϕ/N since a sale is guaranteed for the lowest price in the sector, conditional on being drawn. Since the critical value of the low price is b_θ , the condition follows immediately.)

¹⁶The model is degenerate if there is a single sector. From (9), there are zero messages. This could rather be interpreted as a single firm setting the monopoly price in one message: any other entrant rationally anticipates negative profit if it enters (messages from rivals are necessarily read, so price would go to marginal cost if there is more than one firm, so losses ensue). Section 6, by introducing an "outside" option, allows for a non-degenerate outcome with one active sector.

economic performance. For simplicity (except when we do the symmetric analysis) we will assume that all the π_θ 's are different across sectors. In the sequel, we will find the endogenous number of active sectors.¹⁷

Lemma 2 *Let $N > \phi$. All sectors θ such that $\pi_\theta > \frac{N}{\phi}$ are active sectors, and the rest are inactive. The relative sector sizes are*

$$\frac{n_\theta}{N} = \max \left\{ 1 - \left(\frac{N-1}{\phi \pi_\theta} \right)^{\frac{1}{\phi-1}}, 0 \right\}, \quad \theta = 1, \dots, \bar{\Theta}. \quad (6)$$

Proof. Equating the probability derived from the zero-profit condition, (5), with the probability that she gets no other message from the sector, (1), implies $\mathbb{P}_\theta = \frac{\phi}{N} \left(1 - \frac{n_\theta}{N} \right)^{\phi-1} = \frac{1}{\pi_\theta}$, and so determines the ad market shares by rewriting this as (6). Hence, sector θ sends a positive number of messages if and only if $\pi_\theta > \frac{N}{\phi}$. ■

We defer considering the overall comparative static properties of equilibrium because N is still to be determined in (6). However, we can use the expression to compare across sectors of different economic characteristics within an equilibrium. Sectors with larger economic performance send more messages because they are more attractive to senders. That is, $n_\theta > n_{\theta'}$ if and only if $\pi_\theta > \pi_{\theta'}$. We proceed by further characterizing the relation that sector sizes must satisfy at any equilibrium.

3.2 The inverse IIA property

Sector message sizes exhibit a type of IIA property (Independence of Irrelevant Alternatives) in the sense that the ratio of ad market shares of two sectors depends only on their profitabilities *for a given N* . However, contrary to the usual IIA property (first pointed out by Debreu (1960) in his critique of Luce's (1959) Choice Axiom), which stipulates that the ratio of market shares does not change with the number and type of other options, this ratio *does* change here since N changes with the profitability of a third sector (see also (9) below). Thus, the standard IIA property does *not* hold for this model. However, a related IIA property holds, with respect to the market shares of all competing sectors. We call this the *inverse* IIA property,

¹⁷As we show below, there will be at least 2 sectors under the mild condition that $\pi_2 > 1$.

which pertains to the ratios $m_{-\theta} \equiv \frac{N-n_\theta}{N}$. From (6), the inverse IIA property is:¹⁸

$$\frac{m_{-\theta}}{m_{-\theta'}} = \left(\frac{\pi_{\theta'}}{\pi_\theta} \right)^{\frac{1}{\phi-1}} \quad \text{for all } \theta = 1, \dots, \hat{\Theta}, \quad (7)$$

where $\hat{\Theta}$ is the number of *active* sectors. This is a property of invariance of the ratio of *all rivals'* advertising levels as the appeal of any rival (outside the pair) changes. Analogously to the way the IIA property implies the Logit model (see Luce, 1959), the inverse IIA property implies an *inverse* Logit formulation:

Proposition 3 *At any equilibrium with $\hat{\Theta}$ active sectors, the non- θ shares have a logit form:*

$$\frac{m_{-\theta}}{(\hat{\Theta}-1)} = \frac{\pi_\theta^{-\frac{1}{\phi-1}}}{\sum_{\theta'=1}^{\hat{\Theta}} \pi_{\theta'}^{-\frac{1}{\phi-1}}} \equiv \Psi_\theta, \quad \theta = 1, \dots, \hat{\Theta}, \quad (8)$$

where the LHS is the non-share of sector θ over the total non-share of all sectors.

Proof. Inverting (7),

$$\frac{m_{-\theta'}}{m_{-\theta}} = \left(\frac{\pi_\theta}{\pi_{\theta'}} \right)^{\frac{1}{\phi-1}}.$$

Summing over θ' gives $\frac{(\hat{\Theta}-1)}{m_{-\theta}} = (\pi_\theta)^{\frac{1}{\phi-1}} \sum_{\theta'=1}^{\hat{\Theta}} \pi_{\theta'}^{-\frac{1}{\phi-1}}$, and the result follows directly by inversion. ■

The value of $\hat{\Theta}$ is endogenous here (and is determined below). Only the active sectors are counted: inactive sectors π_θ are excluded from the summation. The same caveat applies below.

As π_θ increases, the RHS of (8) falls: as the profitability of a sector rises, it produces proportionately more ads while the others produce relatively less. Even a mature sector may enjoy a higher profitability if γ_θ falls, perhaps because of the advent of a new medium which might complement advertising its goods and get larger ad market shares which come at the expense of the others.¹⁹ Indeed, as shown in sections 3.4 and 3.5, weak sectors might be pushed out of the market entirely. The effects of raising ϕ on the distribution of messages by sector are fundamentally those of the logit formulation (see for example Anderson, de Palma, and Thisse (1992)), though the derivation of that form above differs from the usual roots.

¹⁸Therefore $\frac{n_\theta}{n_{\theta'}} = \frac{1-(\hat{\Theta}-1)\Psi_\theta}{1-(\hat{\Theta}-1)\Psi_{\theta'}}$, which indicates that IIA does not hold, where Ψ_θ is defined below in (8).

¹⁹We see in Section 3.5 that the number of ads from sector θ' may actually rise if that sector is sufficiently attractive.

Proposition 4 For $\hat{\Theta} > 1$ constant, as ϕ rises, the ad market share of the most profitable sector decreases with ϕ , and the share of the least profitable sector increases. As ϕ falls to 1, almost all messages are sent by the most profitable sector.

Proof. To show the first point, fix $\hat{\Theta} > 1$. The relation in (8) gives the fraction of messages in sector θ as $\frac{n_\theta}{N} = 1 - (\hat{\Theta} - 1) \Psi_\theta$. Note that $\frac{d\pi_\theta^{\frac{-1}{\phi-1}}}{d\phi} = -\frac{\pi_\theta^{\frac{-1}{\phi-1}}}{(\phi-1)^2} \ln \frac{1}{\pi_\theta}$, so that

$$\frac{d\Psi_\theta}{d\phi} \stackrel{s}{=} -\ln \frac{1}{\pi_\theta} + \sum_{\theta'=1}^{\hat{\Theta}} \Psi_{\theta'} \ln \frac{1}{\pi_{\theta'}},$$

(where the symbol $\stackrel{s}{=}$ denotes that the derivative has the sign of the expression), or (since $\sum_{\theta'=1}^{\hat{\Theta}} \Psi_{\theta'} = 1$),

$$\frac{d\Psi_\theta}{d\phi} \stackrel{s}{=} \sum_{\theta'=1}^{\hat{\Theta}} \Psi_{\theta'} \left(\ln \frac{1}{\pi_{\theta'}} - \ln \frac{1}{\pi_\theta} \right).$$

Hence, the share decreases with ϕ for the most profitable sector (1), and increases for the least profitable one ($\hat{\Theta}$). Finally, note from (8) that

$$\Psi_1 = \frac{\pi_1^{\frac{-1}{\phi-1}}}{\sum_{\theta=1}^{\hat{\Theta}} \pi_\theta^{\frac{-1}{\phi-1}}} = \frac{1}{1 + \sum_{\theta=2}^{\hat{\Theta}} \left(\frac{\pi_1}{\pi_\theta} \right)^{\frac{1}{\phi-1}}}.$$

Hence, $\lim_{\phi \downarrow 1} \Psi_1 = 0$: almost all messages are sent from sector 1. ■

If the attention span is very limited (ϕ close to 1), virtually all messages are from the highest profit sector, 1, because this yields the greatest profit conditional on making “the” hit. The messages sent tend to quote the monopoly price because there is almost no chance of being undercut by another message. Monopoly prices are most attractive for the sector with the highest monopoly profit. The number of messages sent from this sector tends to π_1 .²⁰ This corresponds to pure dissipation of the monopoly profit in sector 1. It is possible that there is a huge number of such messages if π_1 is very high: even if π_2 is high too (but strictly below π_1), it attracts virtually no messages. This case arises if the transmission cost for one sector tends to zero while the other sectors retain positive costs: the sector crowds out all other sectors. This is clearly wasteful because all other sectors are closed out, while the affected sector just dissipates all the rents in excessive message transmission.²¹

²⁰This can be seen as follows. If N messages are sent, all from sector 1, and one is drawn, then monopoly pricing implies the profit from a message is $\frac{b_1 - c_1}{N} q_1 - \gamma_1$. The zero profit condition implies the number of messages is π_1 .

²¹As we shall see below, if all sector transmission costs fall proportionately, the range of prices stays the same in each sector: the density of messages sent at any price simple rises proportionately (to the cost decrease) for all sectors.

At the other extreme, when the attention span is extensive, any price above the lowest in the sector will almost certainly be beaten. All sectors are very competitive, so sectors become equally (un)attractive: a lot of price competition means very few messages per sector.

When $\hat{\Theta} > 2$, the advertising shares of the intermediate sectors are not necessarily monotonic in the level of consumer attention, ϕ . To see this, consider 3 sectors. Sectors 1 and 2 have very high profits, with 2 slightly less than 1, while sector 3 has very low profit. When the attention span is slightly above one message, sector 1 is active while 2 is virtually silent. For middling values of ϕ , both 1 and 2 have almost half the market each. For ϕ large, all have around one third shares. Sector 2's share is not monotonic here.

Expression (8) in turn gives rise to a familiar functional form.

3.3 Aggregate advertising

The next step is to determine the equilibrium message volume, N . Expressions (6) and (8) give two different expressions for m_{-i} . Equating them yields:²²

Proposition 5 *The equilibrium total message size given $\hat{\Theta} > 1$ active sectors and $N > \phi$ takes a CES form:*

$$N = \phi \left(\hat{\Theta} - 1 \right)^{\phi-1} \left(\sum_{\theta=1}^{\hat{\Theta}} \pi_{\theta}^{\frac{-1}{\phi-1}} \right)^{-(\phi-1)}. \quad (9)$$

Thus N is increasing in each profitability, π_{θ} , and homogenous of degree one in the sector profitabilities.

Adding another viable sector raises N .

Proof. The properties are straightforward except for the last one. Consider introducing a “barely viable”

sector s with $n_s = 0$: by (6), the corresponding performance of such a new sector s is $\pi_s = N/\phi$. We now

verify that introducing this barely viable sector s leaves (9) unchanged:

$$\frac{N}{\phi} = \frac{\left(\hat{\Theta} - 1 \right)^{\phi-1}}{\left(\sum_{\theta'=1}^{\hat{\Theta}} \pi_{\theta'}^{\frac{-1}{\phi-1}} \right)^{\phi-1}} = \frac{\left(\hat{\Theta} \right)^{\phi-1}}{\left(\sum_{\theta'=1}^{\hat{\Theta}} \pi_{\theta'}^{\frac{-1}{\phi-1}} + \left(\frac{N}{\phi} \right)^{\frac{-1}{\phi-1}} \right)^{\phi-1}}.$$

Thence, by continuity, introducing a strictly viable sector, with $\pi_s > N/\phi$, will cause N to increase even if some sectors exit. ■

²² N can also be derived from summing up the expressions for market shares in (6).

The CES form has well-known properties.²³ Raising the profitability of any sector causes the total volume of messages to rise because the extra clamor causes a larger total without a fully compensating backlash from the other sectors.

3.4 Sector viability

When sectors are asymmetric, some may be precluded by the strength of those in the market. We determine the equilibrium set of active sectors. Recall that $\bar{\Theta}$ denotes the number of sectors for which $\pi_\theta > 1$ (any sector with $\pi_\theta \leq 1$ is not viable, and so can be eliminated from the discussion). Furthermore, define $\hat{\Theta}$ by

$$\pi_{\hat{\Theta}+1} \leq \left(\hat{\Theta} - 1\right)^{\phi-1} \left(\sum_{\theta'=1}^{\hat{\Theta}} \pi_{\theta'}^{\frac{-1}{\phi-1}}\right)^{-(\phi-1)} < \pi_{\hat{\Theta}} \quad (10)$$

and assume that $\hat{\Theta} \in (1, \bar{\Theta})$. As we show, this means that there will be some sectors (all those with index above $\hat{\Theta}$) which do not advertise, and the existence of such sectors implies that the congestion condition $N > \phi$ necessarily holds.²⁴

Proposition 6 *Assume that $\hat{\Theta} \in (1, \bar{\Theta})$ as defined by (10). Then there exists a unique equilibrium: sectors $1, \dots, \hat{\Theta}$ are active, and the total volume of messages is given by (9).*

Proof. From Lemma 2, a sector is active in equilibrium if $\pi_\theta > \frac{N_{\hat{\Theta}}}{\phi}$, where $N_{\hat{\Theta}}$ denotes the number of messages for $\hat{\Theta}$ active sectors as given by (9). Next, we show there is a unique cut-off between active and inactive sectors. The condition for a sector to be active is $\pi_\theta > \frac{N_{\hat{\Theta}}}{\phi}$. Given the ranking of sectors, the LHS decreases in the marginal sector, $\hat{\Theta}$, while we showed in Proposition 5 that the RHS increases as sectors are added. Thus there is a unique solution for $\hat{\Theta}$, and it is given by (10), where the term in the middle is $\frac{N_{\hat{\Theta}}}{\phi}$ (see (9)). Notice that necessarily the congestion condition holds: $N_{\hat{\Theta}} > \phi$ since $\frac{N_{\hat{\Theta}}}{\phi} > \pi_{\hat{\Theta}+1} > 1$ by (10).

It remains to show that the equilibrium follows the ranking: there cannot be an equilibrium with some sector θ excluded while some sector $\theta' > \theta$ is included. If there were, then the profit from sending a single message from sector θ (at its monopoly price, b_θ) is $\pi_\theta \frac{\phi}{N}$. However, messages sent from sector θ' return a

²³For example, it is maximal at symmetry (under the constraint that the sum of the inverse π_θ 's is constant).

²⁴If all potential sectors are active, we get into a corner solution where the condition $N > \phi$ does not necessarily hold. If the model returns a solution with $N < \phi$, it contradicts the congested formula used in setting up the choice probabilities. The existence of some latent sectors is enough to avoid that.

profit of at most $\pi_{\theta'} \frac{\phi}{N}$. Hence, since $\pi_{\theta} > \pi_{\theta'}$, a message from sector θ would supplant one from sector θ' , so the starting point cannot be an equilibrium.²⁵ ■

Viability constraints imply that equilibrium congestion across sectors may close down a sector when another sector becomes more attractive. Similarly, a newly entering or improved sector raises the congestion on the incumbents. This we illustrate next.

3.5 Raising a sector's profitability

Proposition 5 shows that an increase in a sector's profitability will increase the total number of messages sent (even if this causes exit of other sectors). Since the other sectors all send smaller shares of this larger total, the affected sector must send more messages. We now determine what happens to the other sectors. Recall $\frac{n_{\theta}}{N} = 1 - \left(\frac{N-1}{\phi \pi_{\theta}}\right)^{\frac{1}{\phi-1}}$ from (6). Hence for an unaffected sector (where π_{θ} has not changed) it is clear that the sector *share* goes down. However, it is possible the *number* of messages it transmits goes up, as we now show (that is, we show that $\frac{dn_{\theta}}{d\pi_{\theta'}}$ can be positive). Indeed, $\frac{dn_{\theta}}{d\pi_{\theta'}} = \frac{dn_{\theta}}{dN} \frac{dN}{d\pi_{\theta'}}$ has the sign of $\frac{dn_{\theta}}{dN}$ since $\frac{dN}{d\pi_{\theta'}} > 0$. From (6), we have the derivative²⁶

$$\frac{dn_{\theta}}{dN} = 1 - \frac{\phi}{\phi-1} \left(\frac{N-1}{\phi \pi_{\theta}}\right)^{\frac{1}{\phi-1}}.$$

Substituting $\frac{N}{\phi}$ from (9) and defining $\chi_{\theta} = \pi_{\theta}^{\frac{-1}{\phi-1}}$ and $\bar{\chi}$ as the average value of χ_{θ} , gives

$$\frac{dn_{\theta}}{dN} = 1 - \frac{\phi}{\phi-1} \frac{(\hat{\Theta}-1)}{\hat{\Theta}} \frac{\chi_{\theta}}{\bar{\chi}}, \quad (11)$$

From a symmetric starting point (where $\chi_{\theta} = \bar{\chi}$ for all $\theta \leq \hat{\Theta}$), $\frac{dn_{\theta}}{dN}$ has the sign of $1 - \frac{\phi}{\phi-1} \frac{(\hat{\Theta}-1)}{\hat{\Theta}}$, which is negative if and only if $\phi > \hat{\Theta}$. If though $\phi < \hat{\Theta}$, a marginally higher attractivity in one sector causes message numbers to *rise* in all sectors. This result is broadly consistent with the rising part of the information hump (low ϕ) and for the early "take-off" part of the Information Age evolution depicted in Figure 1 below. There is a relatively large increase in the number of messages sent as long as the amount of competition is small.

In the asymmetric case, (11) indicates that there is a cut-off value of χ_{θ} for which $\frac{dn_{\theta}}{dN}$ is negative for higher χ_{θ} and positive for lower χ_{θ} . Since π_{θ} is inversely related to χ_{θ} , this means that *larger sectors are*

²⁵If there are several sectors with the same profitability, then they are either all active or all inactive.

²⁶From which we see that higher $\pi_{\theta'}$ increases the likelihood that the expression is positive.

more likely to see an increase in the number of messages sent. A summary Proposition:

Proposition 7 *The equilibrium total message volume increases as any sector becomes more profitable. The improved sector sends more messages both relatively and absolutely. All other sectors diminish in relative importance, but sufficiently profitable sectors may increase in absolute size.*

It may seem surprising that some sectors could increase in size despite more competition and even though sectors are linked only through the negative effects of congestion (there are no demand complementarities, for example). The favored sector increases in size. This has two contradictory effects on other sectors. First, any given message is less likely to be found. However, any *rival's* message is also less likely to be found. The first effect impacts all industries equally. The second favors the larger industries because each firm has more competition, so these industries will attract new entry.

3.6 The Information Age

The key driver of the information age is lower communication costs. The homogeneity property of the CES function for N in Proposition 5 implies that total message volume doubles if all communication costs are halved.²⁷ This is one obvious cause of a surfeit of information: spam email is an everyday manifestation of the problem. Any such cost improvement is offset by the rise in messages sent, so all improvements are completely dissipated.²⁸ As we show in the next Section, price dispersion also remains unaltered, and this leads to the neutrality result given in Proposition 14 that welfare remains unchanged.

However, even though a uniform cost reduction does not cause new sectors to enter, improved communication may help some sectors more than others, insofar as some are better suited to having their ads embedded in the new media. This leads us to now consider a larger set of viable sectors. The exercise can be thought of as cost reductions in hitherto excluded sectors (or, indeed, as new product classes, like PCs and software, coming to market).

²⁷No further sectors will enter, since doubling of the existing message volume will preclude them, even if their transmission costs halve. Indeed, as we just noted, a sector is viable if and only if $\pi_s > N/\phi$.

²⁸This is reminiscent of Zahavi's Law in transportation, according to which average travel times have remained constant over several decades, despite substantial increases in travel speed.

We consider the symmetric case before returning in the next Section to the asymmetric one. For the symmetric analysis, we will assume that all $\bar{\Theta}$ potential sectors are active.²⁹ Then, with $\pi_\theta = \pi > 1$ for all $\theta = 1, \dots, \bar{\Theta}$, the expression (from (9)) for the total number of messages, N , reduces to³⁰

$$N = \phi \left(\frac{\bar{\Theta} - 1}{\bar{\Theta}} \right)^{\phi - 1} \pi. \quad (12)$$

Having more sectors, $\bar{\Theta}$, raises the total number of messages. The number N is a logistic function of the number of sectors: it is first convex (for $\bar{\Theta} < \phi/2$), and then concave, for $\bar{\Theta} > \phi/2$. If we were to view the number of (new) sectors as arriving at a constant rate, then this means the amount of information would accelerate at first (the take-off of the Information Age) before tapering off, reminiscent of the Bass (1969) diffusion of innovation model. Indeed, the amount of information has an asymptote of $\bar{N} = \phi\pi$, which is the bound to the amount of information the system can sustain.³¹

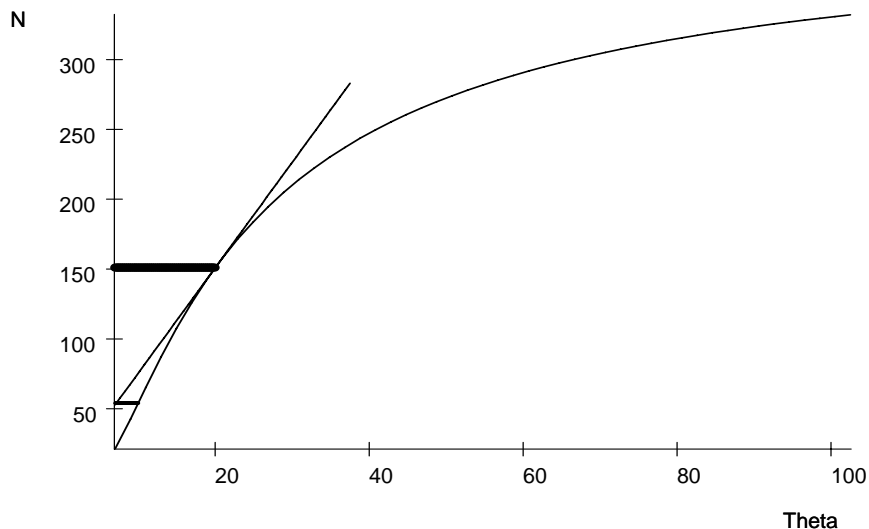


Figure 1. Total messages as a function of the number of active sectors.

The average number of messages per sector, $n_\theta = N/\bar{\Theta}$, is increasing in $\bar{\Theta}$ if and only if $\phi > \bar{\Theta}$, so it is

²⁹We then need to verify that the condition $N > \phi$ is verified with $\bar{\Theta}$ sectors: this is duly met in the Figures below.

³⁰Symmetric CES models are commonly deployed in the economics of product differentiation. Note here that the sector viability constraint, $\pi > N/\phi$, is automatically satisfied.

³¹At the limit, monopoly prices, b , are set in each sector, returning π when the message is chosen. The probability of being chosen is ϕ/\bar{N} , which therefore equals $1/\pi$ (see also (5)).

eventually decreasing (for $\bar{\Theta}$ large enough). The initial increase is explained by the idea that more sectors mean less competition, so higher prices and more incentive to send messages. The logistic function in (12) is sketched in Figure 1, for $\pi = 20$ and $\phi = 20$ (the function asymptotes to $N = 400$, the maximal value of $N/\bar{\Theta}$ is attained at $\bar{\Theta} = 20$, and the inflection point is at $\bar{\Theta} = 10$). The other comparative static property of N , with respect to ϕ , is described next.

3.7 The information overflow hump

The advent of new media means more consumer time is now spent with ad-carrying activities, like surfing the internet or sending email. This likely implies an increase in the overall consumer attention span as more hours are spent on media. The thumbnail capture in the model of this increased span is to raise ϕ .

From the symmetric analysis (see (12)), we can see that the information level, N , is decreasing in the attention span, ϕ if and only if $\phi > \hat{\phi} \equiv \frac{1}{\left(\ln \frac{\bar{\Theta}}{\bar{\Theta}-1}\right)}$, and so N is necessarily decreasing for $\phi > \bar{\Theta}$ (since $\bar{\Theta} \ln \frac{\bar{\Theta}}{\bar{\Theta}-1} > 1$: the LHS is decreasing in $\bar{\Theta}$ and goes to 1 as $\bar{\Theta}$ goes to infinity). Likewise, $\frac{N}{\phi}$ is falling in ϕ , and therefore N increases more slowly than ϕ .

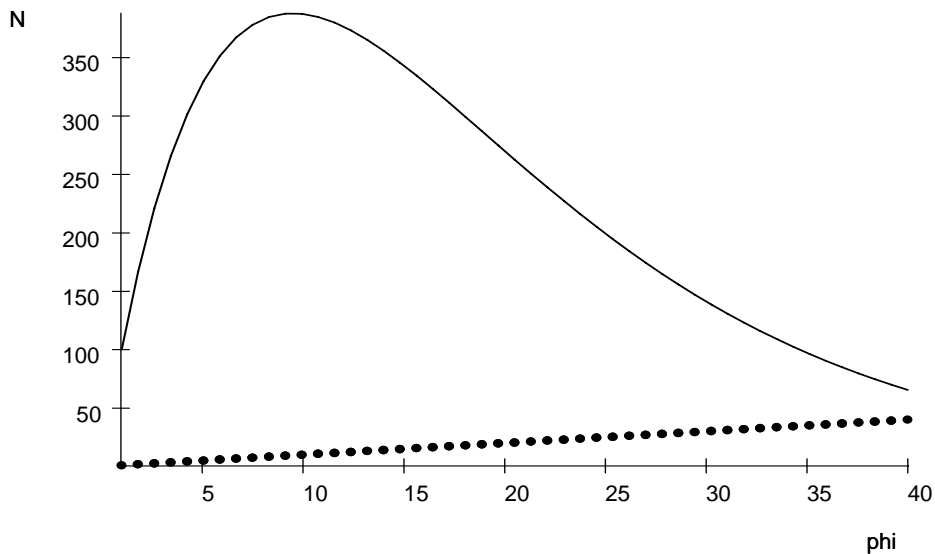


Figure 2. Total number of messages sent, N , as a function of attention span, $\phi \geq 1$.

Figure 2 plots the relation of N as a function of ϕ for $\pi = 100$ and $\bar{\Theta} = 10$ (hence $N = 100\phi \left(\frac{9}{10}\right)^{\phi-1}$ attains its maximum at $\hat{\phi} = \frac{1}{\left(\ln \frac{10}{9}\right)}$, which is slightly less than 10). The dashed line is the line $\phi = N$. Figure 2 shows the quasi-concave function, i.e., first increasing, then decreasing with the attention span, ϕ . This we term the *information overflow hump*. However, the number of messages only increases for low $\phi < \hat{\phi} (< \bar{\Theta})$. More attention has two conflicting consequences. First, it raises the probability a message from the sector is seen, which raises profitability, and hence the number of messages sent, ceteris paribus. But it also has the effect of increasing price competition (the price distribution shifts down), as it is more likely a lower price will be found in the sector. This reduces profitability and leads to a smaller number of firms (messages). For low ϕ , the price competition effect is weak in that it is quite unlikely that another message received will be from the same sector as one already received: extra messages will most likely come from unrepresented sectors. With high reception rates, the price effect dominates. In a nutshell, for low ϕ and given $\bar{\Theta}$, more examination leads to more messages sent as undiscovered sectors become more likely to be found. For higher ϕ , more examination means more hits in the same sector, which increases price competition and so decreases sector activity.

We now turn to the price distribution, whose properties underpin the economics of the results so far.

4 Equilibrium price dispersion

The equilibrium *sales probability* corresponding to a particular price p in sector θ can be determined independently of the other sectors. However, we need to bring in the other sectors to determine which prices are actually used in equilibrium. The equilibrium sales probability for a message announcing price p in sector θ , $\mathbb{P}(p, \theta)$, is given simply from the zero-profit condition as

$$\mathbb{P}(p, \theta) = \frac{\gamma_{\theta}}{(p - c_{\theta}) q_{\theta}} = \frac{(b_{\theta} - c_{\theta})}{(p - c_{\theta})} \frac{1}{\pi_{\theta}}, \quad (13)$$

where $\mathbb{P}(p, \theta) \in (0, 1)$ for all p in the interior of the support of the equilibrium price distribution. The above expression reduces to the zero-profit condition (5), when $p = b_{\theta}$, and using the notation $\mathbb{P}(b_{\theta}, \theta) = \mathbb{P}_{\theta}$.

The equilibrium sales probability above is decreasing and convex in p . We next want to use it to determine

the equilibrium *advertised price* distribution. This is done by equating $\mathbb{P}(p, \theta)$ in the zero profit condition (13) to the expression given in (2) for the probability of there being no lower price drawn, which gives

$$\frac{(b_\theta - c_\theta)}{(p - c_\theta)} \frac{1}{\pi_\theta} = \frac{\phi}{N} \left(1 - \frac{n_\theta F(p, \theta)}{N} \right)^{\phi-1}, \quad (14)$$

where n_θ/N is given by (6).

Proposition 8 *The equilibrium advertised price density in sector θ is decreasing and convex on $[\underline{p}_\theta, b_\theta]$, with (truncated) Pareto distribution*

$$F(p, \theta) = \frac{1 - \left(\frac{N}{\phi \pi_\theta} \right)^{\frac{1}{\phi-1}} \left(\frac{b_\theta - c_\theta}{p - c_\theta} \right)^{\frac{1}{\phi-1}}}{1 - \left(\frac{N}{\phi \pi_\theta} \right)^{\frac{1}{\phi-1}}}, \quad (15)$$

where N is given by (9) and \underline{p}_θ is given by

$$\underline{p}_\theta = c_\theta + \left(\frac{N}{\phi} \right) \frac{\gamma_\theta}{q_\theta}. \quad (16)$$

Proof. The equilibrium advertised price distribution is given from the relation (14) as

$$F(p, \theta) = \frac{N}{n_\theta} \left(1 - \left(\frac{N}{\phi \pi_\theta} \frac{b_\theta - c_\theta}{p - c_\theta} \right)^{\frac{1}{\phi-1}} \right).$$

Recalling that $\frac{n_\theta}{N} = 1 - \left(\frac{N}{\phi \pi_\theta} \right)^{\frac{1}{\phi-1}}$ from (6), we can write (15). It is readily checked that $F(b_\theta, \theta) = 1$.

Since $F(\underline{p}_\theta, \theta) = 0$, the lowest price in sector θ is determined by (14) as:

$$(\underline{p}_\theta - c_\theta) = \frac{(b_\theta - c_\theta) N}{\pi_\theta \phi}. \quad (17)$$

Then (16) follows immediately. The corresponding density, $f(p, \theta)$ is strictly positive on $[\underline{p}_\theta, b_\theta]$, where it is decreasing and convex (as shown by differentiation of (15)). ■

The distribution for sector θ depends on the other sectors through N , giving a simple general equilibrium effect. For given N , we can derive the price distributions by sector independently; since consumer surpluses by sector are additively separable and consumers are not budget constrained.

The intuition for the lowest price in the support is straightforward. A message sent at this lowest price always beats all the other messages from the sector. Hence the sales probability is just the probability that

it is read at all, which is simply $\frac{\phi}{N}$ since it has ϕ shots from a pool of N messages. Equating this probability times the mark-up to the cost of sending the message gives (16).

As in Butters (1977), lower prices are advertised more heavily. In the Butters model (with $q_\theta = 1$), the corresponding lowest price, \underline{p} , would be simply $c_\theta + \gamma_\theta$, because such a price just covers the cost of production plus sending the message. In the Butters version, the lowest price must always get a sale because there is no information congestion, and no possibility that the message remains unread. In contrast, here the lowest price in any sector does not always make a sale. Information overflow pushes up the lowest price in the support, which is needed to compensate for the likelihood that the message may not be received.

The simplest measure of price dispersion is the breadth of the support of the equilibrium prices. This is $b_\theta - \underline{p}_\theta$, where $\underline{p}_\theta = c_\theta + \left(\frac{N}{\phi}\right) \frac{\gamma_\theta}{q_\theta}$. Ceteris paribus, dispersion is smaller the greater is $\left(\frac{N}{\phi}\right) \frac{\gamma_\theta}{q_\theta}$ (recall though that N depends on all the parameters of the model, apart from the inactive sectors' profitabilities). Hence, for example, a larger γ_θ decreases N and so *increases* dispersion in all *unaffected* sectors, while decreasing dispersion in the affected sector (see (9)).

Changes within the sector affect the support as well as the aggregate message volume N . A sector can become inviable if it faces tough competition from other sectors and/or it is quite unattractive itself. Viability can be expressed as the condition that the price support does not collapse. That is $\underline{p}_\theta < b_\theta$. From (17), this means that $\frac{N}{\phi} < \pi_\theta$; this is the same condition from (6) for $n_\theta > 0$ in equilibrium.

Finally, suppose instead that sector θ is associated to a conditional downward-sloping demand and the consumer will buy $q_\theta(p)$ units at the lowest price, p , held. Assume that demand begets a quasi-concave profit function with a maximizing price \hat{p}_θ . The corresponding profit conditional on being the only message registered in the sector) is $(\hat{p}_\theta - c_\theta) q_\theta(\hat{p}_\theta)$, and so the profit per dollar transmission cost is $\hat{\pi}_\theta = \frac{(\hat{p}_\theta - c_\theta) q_\theta(\hat{p}_\theta)}{\gamma_\theta}$, which therefore plays exactly the same role as does π_θ above. In equilibrium, no firm will charge more than \hat{p}_θ because profits can be increased by charging \hat{p}_θ , and the parallel analysis to that above yields the equilibrium price distribution as

$$F(p, \theta) = \frac{1 - \left(\frac{N}{\phi \hat{\pi}_\theta}\right)^{\frac{1}{\phi-1}} \left(\frac{(\hat{p}_\theta - c_\theta) q_\theta(\hat{p}_\theta)}{(p - c_\theta) q_\theta(p)}\right)^{\frac{1}{\phi-1}}}{1 - \left(\frac{N}{\phi \hat{\pi}_\theta}\right)^{\frac{1}{\phi-1}}}$$

(cf. (15)), where N is given by (9) with $\hat{\pi}_\theta$ replacing π_θ . Now \underline{p}_θ is given implicitly by $(\underline{p}_\theta - c_\theta) q_\theta(\underline{p}_\theta) = \left(\frac{N}{\phi}\right) \gamma_\theta$ (cf. (16)), which has a unique price solution for $\underline{p}_\theta < \hat{p}_\theta$ under the assumption that profit is strictly quasi-concave.³² Compared to the distribution for rectangular demand (setting $\hat{p}_\theta = b_\theta$ and treating q_θ as invariant), the distribution is now stochastically lower (FOSD) because lower prices are relatively more attractive due to the demand expansion effect.

4.1 Advertised price dispersion and sector profitability

Greater sector profitability impacts the affected sector by increasing the volume of messages sent (Proposition 5). As we now see (Proposition 9), this increases price competition, and so stochastically lowers prices. However, this market mechanism spills over into the other sectors. Elsewhere, price competition is reduced because sector messages are crowded out in relative terms. Nonetheless, the number of messages sent in other sectors can actually rise (see Proposition 7) because the reduced price competition can raise profits per firm (which then must be reduced by further entry).

Proposition 9 *An increase in the profitability, π_θ , of one sector decreases prices (and increases the support of price dispersion) in that sector and increases prices (and decreases the support of price dispersion) in the other sectors, in the sense of First-Order Stochastic Dominance. A proportional increase in the attractivity of all sectors leaves the price distribution unchanged.*

Proof. Recall $F(p, \theta) = \frac{1 - \left(\frac{N}{\phi\pi_\theta}\right)^{\frac{1}{\phi-1}} \left(\frac{b_\theta - c_\theta}{p - c_\theta}\right)^{\frac{1}{\phi-1}}}{1 - \left(\frac{N}{\phi\pi_\theta}\right)^{\frac{1}{\phi-1}}}$ by (15). Hence $\frac{dF}{d\pi_{\theta'}}$ (for $\theta' \neq \theta$) has the opposite sign from $\frac{dN}{d\pi_{\theta'}}$, which is positive, as already established. Hence $F(p, \theta)$ decreases in $\pi_{\theta'}$. However, $\frac{dF}{d\pi_\theta}$ has the opposite sign, since $\frac{N}{\pi_\theta}$ is decreasing in π_θ (from (9)). Hence, $F(p, \theta)$ increases in π_θ . If π_θ increases, \underline{p}_θ falls; if $\pi_{\theta'}$ increases, N rises so that \underline{p}_θ rises (see (16)). If all sectors increase proportionately in attractivity, $\frac{N}{\pi_\theta}$ is unchanged (by the homogeneity in Proposition 5) and so $F(p, \theta)$ is unchanged. ■

This means that advertised prices (and price dispersion) can be negatively correlated across sectors. If one sector becomes more desirable (in the sense of higher surplus), prices fall in that sector as competition intensifies. But the additional messages crowd out messages in other sectors, and this relaxes competition

³²Otherwise, the support of the price distribution will have a gap for prices such that profit is no lower at a lower price.

in those other sectors. On the other hand, across-the-board changes affecting all sectors can leave prices the same. This property underlies the result in the next Section that proportionately lower message transmission cost savings are dissipated fully: equivalently, a (proportional) tax might be raised without deadweight loss.

The *sales price distribution* differs from the advertised price distribution because lower prices are more likely to get sales, and also because even the lowest advertised price does not always make a sale. It is derived in the on-line version of the paper. We now follow through with the analysis of the symmetric case.

4.2 Dispersion and symmetric sectors

In the symmetric case, N is given by (12) as $N = \phi \left(\frac{\bar{\Theta}-1}{\bar{\Theta}} \right)^{\phi-1} \pi$, and so the cumulative distribution function for advertised prices (15) becomes

$$F(p, \theta) = \bar{\Theta} \left(1 - \frac{\bar{\Theta}-1}{\bar{\Theta}} \left(\frac{b-c}{p-c} \right)^{\frac{1}{\phi-1}} \right), \quad \text{for } p \in [\underline{p}, b], \quad (18)$$

where $\underline{p} = c + \left(\frac{\bar{\Theta}-1}{\bar{\Theta}} \right)^{\phi-1} (b-c)$ (by (16)). Hence, as ϕ rises, the lower bound \underline{p} falls, and so intra-sector competition rises in this respect. A tighter characterization is quite immediate.

Proposition 10 *Assume sectors are symmetric. A higher attention span, ϕ , lowers prices in the sense of First-Order Stochastic Dominance.*

Proof. From (18), $F(p, \theta, \phi_2) > F(p, \theta, \phi_1)$ as $\left(\frac{b-c}{p-c} \right)^{\frac{1}{\phi_2-1}} < \left(\frac{b-c}{p-c} \right)^{\frac{1}{\phi_1-1}}$, or $\phi_1 < \phi_2$. ■

Lower prices as attention goes up underpins the earlier comparative static results of the information hump. Even though the total message volume is not monotone (see Figure 2), the price effect is. For low ϕ , prices are high and few messages are sent: for high ϕ , prices are low and few messages are sent. In the first case, because few messages are registered, firms may as well set high prices and chance the low probability of another message from the same sector. In the second case, price competition intensifies because there is a strong likelihood another message from the same sector will be read.

Along similar lines, it is readily shown that higher $\bar{\Theta}$ stochastically increases prices (with more price dispersion). This is because the limited attention is more divided. We now turn to the normative analysis.

5 Normative properties

We first undertake a welfare analysis of the performance of the market equilibrium and emphasize the excess of information. In the following two sub-sections we consider cost changes and transmission taxes - even cost increases without any corresponding revenue collection can improve the allocation. These results stress the extent of the market failure, and also help indicate which sectors are particularly responsible.

One strong property of the Butters (1977) model is that the market allocation is optimal. However, this property crucially depends on his assumption that each message hits somewhere.³³ In our set-up, there is rent dissipation and socially wasteful duplication of messages.³⁴ Competition for attention imposes a congestion externality which leads to excessive advertising: this feature is perhaps more in tune (rather than optimality or under-advertising) with one's personal reaction to advertising clutter.

The welfare function is given by summing over sectors the total sector surplus times the probability a sale is made in the sector, and then subtracting the message costs. Define

$$\mathbb{Q}(n_\theta, N) = 1 - \left(1 - \frac{n_\theta}{N}\right)^\phi. \quad (19)$$

as the probability that there is *at least one hit* in sector θ : the probability that each of the n_θ messages is missed on each of the ϕ draws. Notice that

$$\frac{\partial \mathbb{Q}(n_\theta, N)}{\partial n_\theta} = \mathbb{P}_\theta. \quad (20)$$

Thus the increased chance of discovering a sector when an extra message is sent is the probability that the extra message is registered when no other message from the sector has registered. We can write the welfare function (for any values $n_\theta \geq 0$, $\theta = 1, \dots, \bar{\Theta}$) as³⁵

$$W(n_1, \dots, n_{\bar{\Theta}}; N) = \sum_{\theta=1}^{\bar{\Theta}} [(b_\theta - c_\theta) q_\theta \mathbb{Q}(n_\theta, N) - \gamma_\theta n_\theta], \quad (21)$$

where $N = \sum_{\theta=1}^{\bar{\Theta}} n_\theta$. This form (breaking out N as a separate argument) is convenient for what follows.

³³It also depends on the rectangular demand assumption. Stegeman (1991) shows that there is insufficient advertising if demand slopes down, because firms do not internalize the consumer surplus of lower prices.

³⁴Clearly the first best optimum comprises one message per sector, and the active sectors should be the ϕ for which the profit per message, $(b_\theta - c_\theta) q_\theta - \gamma_\theta$, is highest. If γ is the same for all θ , these are the first ϕ ones, those for which π_θ is highest.

³⁵When it comes later to including tax revenues, all we will need to assume is that they have some social value.

Lemma 11 *The social benefit from an extra message in sector θ is equal to*

$$\frac{dW}{dn_\theta} = \frac{\partial W}{\partial n_\theta} + \frac{\partial W}{\partial N} = ((b_\theta - c_\theta) q_\theta \mathbb{P}_\theta - \gamma_\theta) + \frac{\partial W}{\partial N}, \quad (22)$$

where the RHS terms are private sector profit and congestion externality respectively.

Proof. From (21), we have $\frac{dW}{dn_\theta} = \frac{\partial W}{\partial n_\theta} + \frac{\partial W}{\partial N} \frac{dN}{dn_\theta}$: noting that $\frac{dN}{dn_\theta} = 1$ (message anonymity) gives the first inequality in (22). Now, from (21) and (19), and then using (20), we have that

$$\begin{aligned} \frac{\partial W}{\partial n_\theta} &= (b_\theta - c_\theta) q_\theta \frac{\partial \mathbb{Q}(n_\theta, N)}{\partial n_\theta} - \gamma_\theta \\ &= (b_\theta - c_\theta) q_\theta \mathbb{P}_\theta - \gamma_\theta. \end{aligned} \quad (23)$$

This expression is the profit of a firm setting the top price in active sector $\theta \leq \hat{\Theta}$ given n_θ messages emanating from the sector (see (5)). Since this is zero in equilibrium, the remaining term, $\partial W/\partial N$, is naturally interpreted as the congestion externality from active sectors. ■

The next result shows the externality is negative, and quantifies it at the equilibrium allocation.

Proposition 12 *The total number of messages transmitted is excessive in equilibrium, and the (negative) congestion externality is measured as the average transmission cost, $\frac{1}{N^e} \sum_{\theta=1}^{\hat{\Theta}} \gamma_\theta n_\theta^e$.*

Proof. Recall that $\frac{dW}{dn_\theta} = \frac{\partial W}{\partial n_\theta} + \frac{\partial W}{\partial N}$, and $\frac{\partial W(n^e, N^e)}{\partial n_\theta} = 0$ (where the superscript e denotes that the variable is evaluated at its equilibrium value) by the zero profit condition of equilibrium for all active θ . Then we have that $\frac{dW(n^e, N^e)}{dn_\theta} = \frac{\partial W(n^e, N^e)}{\partial N}$. From (21), we have for active sectors

$$\begin{aligned} \frac{\partial W(n, N)}{\partial N} &= \sum_{\theta=1}^{\hat{\Theta}} (b_\theta - c_\theta) q_\theta \frac{d\mathbb{Q}(n_\theta, N)}{dN} \\ &= - \sum_{\theta=1}^{\hat{\Theta}} (b_\theta - c_\theta) q_\theta \frac{n_\theta}{N} \mathbb{P}_\theta. \end{aligned} \quad (24)$$

Using the zero profit condition (3) we get

$$\frac{\partial W(n^e, N^e)}{\partial N} = -\frac{1}{N^e} \sum_{\theta=1}^{\hat{\Theta}} \gamma_\theta n_\theta^e < 0, \quad (25)$$

i.e., the congestion externality is strictly negative and equal to (minus) the average transmission cost. ■

This result underscores the main problem with the market equilibrium: although (as we show next) the allocation is (second-best) optimal across sectors given the total equilibrium message volume, the *overall volume is excessive*. This is seen clearly from what we just argued in Lemma 11, namely that $\frac{\partial W(n^e, N^e)}{\partial n_\theta} = 0$ (i.e., evaluated at the equilibrium), while $\frac{dW(n^e, N^e)}{dN} < 0$. However, while optimal and private incentives are aligned in terms of allocation, the private choice ignores the message crowding externality on all other sectors, which is measured by $\frac{\partial W(n^e, N^e)}{\partial N} < 0$. This implies excessive messages are sent. The social cost of an extra message, as per (25), is the average sending cost. This relation holds because if extra messages have to be sent, they should be allocated across sectors in proportion to the sector representation in the population: one more message therefore costs the average transmission cost.

Proposition 13 *The equilibrium allocation of messages across sectors is socially optimal given the number of messages transmitted at the equilibrium.*

This is proved in the Appendix. The key feature that generates the optimality result is (20): the marginal change in the choice probability holding fixed the total number of messages, $\frac{\partial Q(n_\theta, N)}{\partial n_\theta}$, which is crucial to the social problem, is equal to the probability \mathbb{P}_θ that the highest-priced firm makes a sale in the private problem. The equivalence holds because the probability that an extra message is examined and nothing else was examined from the sector reflects both its social contribution and the private incentive for sending it.

5.1 Increasing transmission costs uniformly

We first establish a strong neutrality result for across-the-board cost changes. Uniform transmission cost decreases raise advertising levels (and industry sizes) proportionately, and so are a strong driver for increased information, but they do not affect the real outcome. Indeed, the economics of lower transmission rates are the economics of rent dissipation. Halving the cost in each sector simply doubles the number of ads sent per sector. The intuition comes from the fact that both N and n_θ are homogeneous of degree minus one. The sector choice probabilities (n_θ/N) are then homogeneous of degree zero in the percentage cost increase. The advertised price distribution, $F(p, \theta)$ is then also independent of such cost changes. This also explains why no sectors enter in the face of a common cost increase: halving transmission costs also halves the chance the

highest priced sender makes a sale (since it faces twice the competition). Because optimal taxes are positive, we phrase the next proposition in terms of a cost rise.

Proposition 14 *A uniform percentage increase in transmission costs leaves welfare unchanged. Hence a uniform percentage tax raises welfare. Price dispersion remains unchanged, as does the fraction of messages sent per sector, while the number of messages per sector (and therefore the total) goes down in proportion to the percentage cost increase. The number of active sectors remains the same.*

Proof. A common percentage transmission cost increase, s , raises each γ_θ to $\gamma_\theta(1+s)$ and so reduces each π_θ proportionately to $\frac{\pi_\theta}{1+s}$. From (9), such a common cost increase means $N(s)(1+s)$ is constant, where $N(s)$ is the equilibrium aggregate message volume under common cost increase s . Equivalently, the original $N(0)$ falls to $N(s) = \frac{N(0)}{1+s}$. Recall $\frac{n_\theta}{N} = 1 - \left(\frac{1}{\phi} \frac{N}{\pi_\theta}\right)^{\frac{1}{\phi-1}}$ from (6). Since the ratio $\frac{N}{\pi_\theta}$ (on the RHS) is unaltered by the cost increase, then so is the ratio $\frac{n_\theta}{N}$ (on the LHS). Likewise, since $\frac{N}{\pi_\theta}$ is unchanged, the price support and the cumulative price distribution stay the same. Recall that a sector is active iff $(b_\theta - c_\theta) q_\theta \frac{\phi}{N} > \gamma_\theta$. With a common cost increase s , the condition becomes $(b_\theta - c_\theta) q_\theta \frac{\phi}{N(s)} > \gamma_\theta(1+s)$. However, since $N(s)(1+s) = N(0)$, the condition remains unaltered. Consumer welfare therefore is unchanged, profits remain zero, and so welfare remains unchanged. The tax result is an immediate corollary. ■

If tax revenues were discarded, a tax would have no effect on welfare. Any tax not lost in the collection is therefore a social gain, and gets transferred purely from costs. Since profits are zero, consumers are just as well off since they face the same situation (same distributions, but fewer overall messages). The tax is therefore raised *without deadweight loss*.

Proposition 13 showed that the base allocation of messages was optimal for the equilibrium message volume, N^e . By Proposition 14, an equal percentage tax on transmission scales back messages proportionately. However, unless transmission costs are the same across sectors, the scaled-back message levels induced by a non-negligible tax are not optimal for the new (given) total volume of messages. Indeed, the partial welfare derivative (23) is $\frac{\partial W(n, N)}{\partial n_\theta} = (b_\theta - c_\theta) q_\theta \mathbb{P}_\theta - \gamma_\theta$, which expression still holds in the presence of a tax which is fully redistributed back to consumers (although the arguments in \mathbb{P}_θ are proportionately smaller). These

partial derivatives are still to be equalized across sectors at any constrained optimal allocation for given N^e . However, the market equilibrium condition in the presence of a proportional tax, τ , on transmission becomes $(b_\theta - c_\theta) q_\theta \mathbb{P}_\theta = \gamma_\theta (1 + \tau)$. Substituting,

$$\frac{\partial W(n^e, N^e)}{\partial n_\theta} = \tau \gamma_\theta. \quad (26)$$

This means that the allocation is constrained optimal (all the $\partial W(n^e, N^e) / \partial n_\theta$ are equal) either if $\tau = 0$ (where we evaluated the earlier welfare derivative), or if all the transmission costs, γ_θ , are equal. Otherwise, ramping up the transmission cost with a tax causes an allocative distortion: from (26), the higher-cost messages ought to be provided more (and the lower-cost ones less). This means that the *cheaper messages tend to be overused in equilibrium* (in the presence of the tax). These are the ones associated with the most dissipation, *ceteris paribus*.

This suggests that the low transmission-cost sectors are over-represented in the population of messages (in the sense that they ought to be scaled back more than proportionately). Although the proportional tax does not effect choice probabilities, the fact that the allocation is no longer optimal if transmission costs are different means that the optimal tax (given a target N) is not a proportional one. Instead, the optimal tax should fall more heavily on the cheaper message communications: from (26), the sector-specific tax rate that ensures all sectors have the same marginal social benefit entails τ_θ inversely proportional to γ_θ .³⁶

5.2 Specific cost increases

Proposition 13 suggests that low transmission-cost sectors do not inflict more damage on high transmission-cost ones, or vice versa, at equilibrium. All sectors are in excess, but no group should be singled out.

This result leads us to ask whether a deterioration in a sector - say an increase in the sector's sending cost (like a tax with the proceeds discarded) - can reduce welfare. As we shall show, such an increase cannot help if all sectors are the same, but it can if they are sufficiently asymmetric. Loosely, cost increases may help on sectors with low transmission costs (relative to surplus) and those with low surpluses.

³⁶Indeed, the first-best optimum entails just one message per sector, which also suggests more than proportional scaling back through taxes of low-cost sectors. Low-cost sectors are also the sectors with small tax raised per message: the high-cost sectors have the additional social benefit of larger tax revenue per message.

Proposition 15 *Welfare can rise when transmission costs increase in sectors with low profitability or with low transmission costs. Hence welfare rises from a transmission tax on such sectors.*

Proof. From (21), and since $\frac{\partial W}{\partial n_{\theta'}} = 0$ at equilibrium for active sectors, the relevant welfare derivative is

$$\frac{dW}{d\gamma_{\theta}} = -n_{\theta} + \frac{\partial W}{\partial N} \frac{dN}{d\gamma_{\theta}}.$$

This expression indicates that there is a trade-off. From (25), $\frac{\partial W}{\partial N} = -\frac{1}{N} \sum_{\theta'=1}^{\hat{\Theta}} n_{\theta'} \gamma_{\theta'}$; from (9), we have $\frac{dN}{d\gamma_{\theta}} = -N \frac{1}{\gamma_{\theta}} \frac{1}{\bar{\Theta}} \frac{\chi_{\theta}}{\bar{\chi}}$, where we recall that $\bar{\chi}$ is the average value of $\chi_{\theta} = \left(\frac{1}{\pi_{\theta}}\right)^{\frac{1}{\phi-1}}$. Pulling these expressions together, the derivative condition is:

$$\frac{dW}{d\gamma_{\theta}} = -n_{\theta} + \frac{1}{\gamma_{\theta}} \frac{1}{\bar{\Theta}} \frac{\chi_{\theta}}{\bar{\chi}} \sum_{\theta'=1}^{\hat{\Theta}} n_{\theta'} \gamma_{\theta'}.$$

Under symmetry, a rise in one sector's transmission costs has no effect at the margin, since $dW/d\gamma_{\theta} = 0$.

To deal with asymmetric sectors, it helps to rewrite the above expression as

$$\frac{dW}{d\gamma_{\theta}} \stackrel{s}{=} \frac{-n_{\theta} \gamma_{\theta}}{\sum_{\theta'=1}^{\hat{\Theta}} n_{\theta'} \gamma_{\theta'} / \hat{\Theta}} + \frac{\chi_{\theta}}{\bar{\chi}} = -\frac{\Gamma_{\theta}}{\bar{\Gamma}} + \frac{\chi_{\theta}}{\bar{\chi}}, \quad (27)$$

where $\Gamma_{\theta} = n_{\theta} \gamma_{\theta}$ is the aggregate transmission cost for sector θ , and $\bar{\Gamma}$ is the average of these. (27) shows that it will typically be beneficial to increase costs on some sectors: the two effects in $\frac{dW}{d\gamma_{\theta}}$ go in different directions. Consider two special cases. First, suppose that two sectors have the same transmission cost, and one is more profitable than the other, so it also has a higher equilibrium industry size (number of messages). Then Γ_{θ} is smaller for the less profitable one, and χ_{θ} is larger, so $\frac{dW}{d\gamma_{\theta}}$ is larger for the less profitable one. Second, suppose that two sectors are equally profitable, so they have the same equilibrium industry size (number of messages). If one has a higher transmission cost than the other, its Γ_{θ} is larger, and so $\frac{dW}{d\gamma_{\theta}}$ is smaller.³⁷ Thus, higher transmission costs are beneficial, *ceteris paribus*, in less profitable or in low transmission cost sectors. The tax result is an immediate corollary. ■

The analysis of this sub-section indicates the low-profit products and those with low transmission costs as being socially harmful. This holds despite them having a small foothold: one might have otherwise

³⁷This is because a cost increase in the sector with the larger transmission cost has a relatively smaller effect: profit is not changed much so there is little reduction in congestion, but the higher cost is borne over a large market base. The same cost increase in the smaller cost sector has a larger effect on profit, π_{θ} , and so causes a much larger reduction in message congestion, while being borne over a smaller base since the sector contracts more.

suspected high-profit products because they are responsible for the most crowding. The previous subsection also points the finger at low transmission-cost products as being over-represented when all messages are scaled back proportionately by a proportional tax. Thus these results take different perspectives on the “blame” issue, but reach similar conclusions.

6 Distractions

The strong neutrality property of Proposition 14, that uniform cost changes have no real effects, relies critically on the lack of outside competition for attention, which also implies the homogeneity property in (9). One natural way to relax this property is to introduce another source of competition for attention. This amendment retains a basic CES form and the broad comparative static properties, but now implies that transmission cost changes have real effects: a lower cost increases the likelihood of finding any sector, and decreases prices. However, taxing messages remains optimal.

Think of consumers as having a limited amount of time, or a limited attention span. They cannot process all the information coming at them. Jostling with the price of MicroSoft Word or a supermarket flyer for pork chops is a really important email from a Dean or a crying child. We model this outside competition for attention as further “distractions” to attention. Formally, suppose there are n_0 (exogenous) other messages (or activities) which compete for attention along with the advertising messages. Hence now we have $N = \sum_{\theta=0}^{\hat{\Theta}} n_{\theta}$. We associate an exogenous social value $\pi_{\theta} > 0$ to outside *each* message (or activity) examined (the positive value reflects the fact that the individual allows herself to be distracted).

6.1 Message volume with distractions

With distractions, it is still true that each active sector’s message share is $\frac{n_{\theta}}{N} = 1 - \left(\frac{N-1}{\phi \pi_{\theta}}\right)^{\frac{1}{\phi-1}}$, $\theta = 1, \dots, \hat{\Theta}$ (see (6)). However, to find the total number of messages, N , now means adding in the outside sector, so the earlier CES form is amended to yield the implicit form:

$$N = n_0 + N \sum_{\theta=1}^{\hat{\Theta}} \left(1 - \left(\frac{N-1}{\phi \pi_{\theta}} \right)^{\frac{1}{\phi-1}} \right).$$

Writing this out, we have a quasi-CES form

$$n_0 + N \left(\hat{\Theta} - 1 \right) = N^{\frac{\phi}{\phi-1}} \sum_{\theta=1}^{\hat{\Theta}} \left(\frac{1}{\phi} \frac{1}{\pi_{\theta}} \right)^{\frac{1}{\phi-1}}. \quad (28)$$

The LHS is linear in N (with a positive intercept), and the RHS is convex (and starts at 0), so that there is one intersection with $N > 0$, which is thus the unique solution. The comparative static properties of the equilibrium are quite simple, and concur with previous results. One new one: a higher value of n_0 leads to a higher N , and n_{θ} falls in all other sectors.

However, now there is a real effect of uniform cost changes. To see this, suppose that γ_{θ} falls to $\gamma_{\theta}(1+s)$ for all $\theta > 0$ (with $s < 0$). Then (28) becomes $\frac{n_0}{N} + \left(\hat{\Theta} - 1 \right) = \left(N(1+s) \right)^{\frac{1}{\phi-1}} \sum_{\theta=1}^{\hat{\Theta}} \left(\frac{1}{\phi} \frac{1}{\pi_{\theta}} \right)^{\frac{1}{\phi-1}}$, and clearly N rises when s falls. From the same equation, a higher N also entails a lower value of $N(1+s)$. This means that a lower cost per message now raises the number of messages *less than proportionately*. From (6), each active sector's share of the larger message total is bigger, as well as being larger in absolute terms. This is reflected too in the equilibrium price distribution: from (15) and the arguments of Proposition 9, noting that N/π_0 rises, the lower cost decreases prices (in the sense of First-Order Stochastic Dominance).

6.2 Welfare analysis

We first show that there is still the right allocation of N (cf. Proposition 13), but too many messages. The welfare function is now written as

$$W = \sum_{\theta=1}^{\bar{\Theta}} [(b_{\theta} - c_{\theta}) q_{\theta} Q_{\theta} - \gamma_{\theta} n(\theta)] + n_0 \pi_0 \frac{\phi}{N}, \quad (29)$$

where n_0 distractions vie for the attention span of ϕ given N total competitors. The proof follows the lines of the earlier one. For any given N , the partial derivative marginal benefit expressions (which are to be equalized across all sectors in the second-best problem of choosing the optimal allocation of N messages) are the same as those given before, and hence the equilibrium still has the “right” allocation of the messages across sectors, but too many messages (given $\pi_0 > 0$).

We already showed a uniform cost increase reduces the number of messages, and has real effects which harm consumers since prices rise. Nonetheless, we now show a uniform percentage tax on all sectors (except

the “untaxed” sector, n_0) still raises welfare. From (29), the effect of a tax is $\frac{dW}{d\tau} = \sum_{\theta=1}^{\hat{\Theta}} \frac{\partial W}{\partial n_{\theta}} \frac{dn_{\theta}}{d\tau} + \frac{\partial W}{\partial N} \frac{dN}{d\tau}$.³⁸ Evaluating at $\tau = 0$ yields again the result that the equilibrium entails, $\frac{\partial W}{\partial n_{\theta}} = 0$, where the zero comes from the zero profit condition, as seen before. Hence, $\frac{dW}{d\tau} = \frac{\partial W}{\partial N} \frac{dN}{d\tau}$, and we know $\frac{dN}{d\tau} < 0$. Also, $\frac{\partial W}{\partial N} < 0$ since each Q_{θ} term is decreasing in N and the additional term, $n_0\pi_0\frac{\phi}{N}$, is decreasing in N (given that $\pi_0 > 0$). Hence welfare increases locally from a uniform percentage tax. Here, a tax has the additional social benefit of rendering more prominent the “distractions.”

7 Conclusions

The Information Age is characterized by a surfeit of information sent at relatively low cost. Modern economies involve many media which can be used to catch the attention of prospective consumers, so the attention span of consumers is likely larger than ever before. Yet modern economies also involve many product classes. These factors interact to determine information congestion and the consequent degree of competitiveness of sectors, as reflected in the degree of price dispersion. Below we bring together some of the key comparative static properties and how they are transmitted.

First, new product classes may displace others by crowding information spans. As profitable new opportunities arise, or as the cost of communicating them through new media falls, less profitable classes are displaced. Total information volume rises, and new (or improved) sectors carve out advertising market shares at the expense of the others. Nevertheless, sufficiently strong other sectors may see a rise in size because crowding relaxes price competition leading to stochastically higher prices. This encourages messages when the enhanced profit effect dominates the direct crowding effect.

Second, *ceteris paribus*, increasing the number of product classes causes an initial acceleration in the volume of messages as crowding raises prices making more ads profitable. Eventually this tails off, in a classic S-shaped (logistic) volume relation over time, with an upper bound to message volume.

Third, if consumer attention rises, prices fall stochastically as competition is enhanced. This gives rise to the Information Hump: information volume initially rises as it becomes easier to get messages across. But

³⁸The welfare function(29) is written in terms of real resource costs and benefits: implicitly, any tax raised from suppliers is being redistributed to consumers with unit marginal utility of money.

the lower prices eventually come to dominate as it becomes less profitable to send messages as it is likely that other offers register with the consumer. This suggests that both more attention and more product classes raise the volume of information. Eventually though the attention span effect reduces information volume and increases competition. Thus, whether prices get lower depends crucially on whether attention rises “faster” than the range of (desirable) goods.

The model borrows heavily from Butters (1977) in using a zero-profit condition to derive equilibrium price distributions. But it differs in key respects in assumptions and conclusions. While Butters’ model assumes that each message is read by some consumer, here some messages are “lost” because they are not read at all. We stress too the competition for attention across sectors, which gives rise to cross-sector effects in pricing and message volume. While Butters finds that the overall level of advertising is optimal, we have too much advertising, though a constrained optimality result is retained in the sense that the allocation across sectors is optimal, given the equilibrium message level.

The intuition for our optimal allocation of ads across sectors, given the total (excessive) volume, is as follows. First, the congestion externality of the overall ad level is the same regardless of which sector sends an extra ad (the term $\partial W/\partial N$ in the normative analysis). Second, the individual sector contribution to welfare from an extra ad is the probability it is seen, weighted by its social contribution, from which is subtracted the sending cost. As with the Butters model, this is the profit of the top firm, and so is zero for all sectors.

The model delivers a detailed picture of equilibrium price distributions across asymmetric sectors competing for attention. Equilibrium message ratios are shown to obey an *inverse* IIA property. The equilibrium total volume of advertising messages is a CES function of the individual sectors’ profitability measures. This constitutes a novel derivation of such a CES function, and is instrumental in delivering sharp predictions.

A CES form is still central when we allow for “distractions” to the attention paid to ads. This device relaxes the homogeneity property that proportional decreases in communication costs raise ad levels proportionately, and gives rise to a modified CES form for ad levels, whereby lower costs across the board now may cause weaker sectors to exit. However, a tax on ads still raises welfare despite the introduction of an “untaxed” sector (there is still over-advertising), and the allocation of ads across sector is still optimal under

the constraint of the equilibrium total volume of messages.

Some caveats to the analysis constitute further extensions. The model is one of firms seeking (passive) consumers through ads, which can be thought of as the pure Couch Potato model. The converse case has consumers seeking opportunities through search. Indeed, both sides can be active, as in Baye and Morgan (2001). One step in this direction is to allow the attention span to be endogenously determined by equating the expected surplus from an extra ad to the marginal cost of paying more attention: the current specification can be viewed as a simple version of this with prohibitive marginal cost at ϕ .

The model also views all media as equally delivering messages for attention, and is not immediately equipped to deal with which messages might be better suited to which media. Nor indeed is media pricing of message delivery given much shrift, though this is the topic of the (platform) economics of broadcasting. Instead, perhaps like billboards, web-sites and bulk mail, access price is exogenous. The crucial marketing dimension of targeting of messages to consumers (for example through the use of specific media) has been closed down through the device of a single representative consumer. Likewise, messages are assumed to be sampled randomly, so there is no allowance for the consumer to pay more attention to particular message types or filter out others. Information congestion and the Economics of Attention have yet to be fully fleshed out in these broader directions.

8 Appendix

8.1 A. Comparison to Butters model

Butters (1977) supposes M consumers, and a *single* sending sector (so we can suppress the subscript θ in what follows). Letters are sent randomly, and each message reaches only a *single* consumer (ours potentially reach all consumers). Consumers examine *all* the messages received, and each buys at the lowest price received. As with our model, the equilibrium price support has no atoms, no holes, and runs up to b . It starts at $c + \gamma$, because a message at that price is surely read by whoever receives it, and it is a winner (in our model, it must start higher because even the best deal may be unread).

We follow Butters in equating the probability of a sale from two different perspectives. The first is the

zero profit condition, $\mathbb{P} = \frac{\gamma}{p-c}$. The second is the finding probability for the price p . For the price b , the likelihood of finding an empty letter box (the only way for the highest price to make a sale) must therefore equal $\frac{\gamma}{b-c}$. This is thus the fraction of the market unserved, and so is a key statistic in comparing equilibrium to optimum.

The corresponding welfare function is $W = (b - c)M\Lambda - \gamma N$ if N messages are sent, where Λ is the fraction of consumers informed. Hence the optimal number of ads is determined from $(b - c)M \frac{d\Lambda}{dN} - \gamma$: this equation suggests that an exponential form for the probability of finding an empty letterbox will give equivalence with the equilibrium. This remark underscores the formulation of Butters' letterbox technology. To derive this, note that the probability that at least one of N letters sent reaches a particular one of the M letterboxes is $1 - (1 - \frac{1}{M})^N$. When M is large, this is approximately $1 - \exp(-N/M)$ ($= \Lambda$). Hence, $\frac{d\Lambda}{dN} = \frac{1}{M} \exp(-N/M) = \frac{1}{M} (1 - \Lambda)$, from which it follows that the number of uninformed at the optimum is $\frac{M\gamma}{b-c}$, the same as in equilibrium.³⁹

Finally, consider the equilibrium advertised price distribution in the Butters model. Let the number of letters priced below p be $A(p)$ (which therefore replaces N in the logic of the previous paragraph). Hence the probability of a letter missing all lower-priced letters in a mailbox is $\exp(-A(p)/M)$ which must equal $\frac{\gamma}{p-c}$ by the zero profit condition. The form of $A(p)$ and its properties (decreasing, concave) follow directly.

8.2 B. Proof of Proposition 13

Let N be given at the equilibrium level stipulated by (9), that we denote as N^e , and we wish to show that the *division* of these messages effectuated in equilibrium is optimal.

First, note that maximization of $W(\cdot)$ under the constraint that the non-negative n_θ 's sum to a given value of N is a maximization problem of a continuous function on a compact set and therefore must have a solution. Therefore at least one of the n_θ must be positive: call this sector j .

Second, substituting the constraint $n_j = N - \sum_{\theta \neq j}^{\bar{\Theta}} n_\theta$ into $W(n_1, \dots, n_j, \dots, n_{\bar{\Theta}}; N)$ enables us to write the maximand as $\tilde{W}(n_1, \dots, [n_j], \dots, n_{\bar{\Theta}}; N)$, and we now show that $\tilde{W}(\cdot)$ is concave in the arguments $n_1, \dots, [n_j], \dots, n_{\bar{\Theta}}$

³⁹The interpretation is that the business stealing and consumer surplus appropriation externalities net out.

(for given N), where the notation $[n_j]$ denotes that the corresponding argument, n_j , is excluded. Indeed,

$$\begin{aligned}\tilde{W}(n_1, \dots, [n_j], \dots, n_{\bar{\Theta}}; N) &= (b_j - c_j) q_j \mathbb{Q} \left(N - \sum_{\theta \neq j}^{\bar{\Theta}} n_{\theta}, N \right) - \gamma_j \left(N - \sum_{\theta \neq j}^{\bar{\Theta}} n_{\theta} \right) \\ &\quad + \sum_{\theta \neq j}^{\bar{\Theta}} [(b_{\theta} - c_{\theta}) q_{\theta} \mathbb{Q}(n_{\theta}, N) - \gamma_{\theta} n_{\theta}].\end{aligned}$$

Recall that the sum of concave functions is concave. The terms in the transmission costs γ are linear in $n_1, \dots, [n_j], \dots, n_{\bar{\Theta}}$, while for $\theta \neq j$, the $\mathbb{Q}(n_{\theta}, N)$ terms are concave in own n_{θ} . There remains the term $\mathbb{Q} \left(N - \sum_{\theta \neq j}^{\bar{\Theta}} n_{\theta}, N \right) = 1 - \left(\frac{\sum_{\theta \neq j}^{\bar{\Theta}} n_{\theta}}{N} \right)^{\phi}$ (by definition (19)): the summation term is linear in the n_{θ} , given N ; hence raising this to a power $\phi > 1$ gives a convex function, and one minus a convex function is concave, as desired.

Third, since $\tilde{W}(\cdot)$ is concave, and is maximized over a compact and convex set, it has a unique maximal value. Let a solution be denoted $\{n_1^o, \dots, [n_j^o], \dots, n_{\bar{\Theta}}^o\} \geq 0$, with $n_j^o = N^e - \sum_{\theta \neq j}^{\bar{\Theta}} n_{\theta}^o$, and let $\{\mu_1^o, \dots, [\mu_j^o], \dots, \mu_{\bar{\Theta}}^o\} \geq 0$ be the corresponding Lagrangian multipliers. The solution maximizes \tilde{W} if and only if $\{n_1^o, \dots, [n_j^o], \dots, n_{\bar{\Theta}}^o; \mu_1^o, \dots, [\mu_j^o], \dots, \mu_{\bar{\Theta}}^o\}$ solves the Karush-Kuhn-Tucker conditions. This means that:

$$\frac{\partial \tilde{W}}{\partial n_{\theta}} = \begin{cases} = 0 & \text{if } n_{\theta}^o > 0, \\ \leq 0 & \text{if } n_{\theta}^o = 0. \end{cases}$$

By (23) we have $\frac{\partial \tilde{W}}{\partial n_{\theta}} = ((b_{\theta} - c_{\theta}) q_{\theta} \mathbb{P}_{\theta} - \gamma_{\theta}) - ((b_j - c_j) q_j \mathbb{P}_j - \gamma_j)$. Therefore

$$((b_{\theta} - c_{\theta}) q_{\theta} \mathbb{P}_{\theta} - \gamma_{\theta}) \quad \begin{cases} = ((b_j - c_j) q_j \mathbb{P}_j - \gamma_j) & \text{if } n_{\theta}^o > 0, \\ \leq ((b_j - c_j) q_j \mathbb{P}_j - \gamma_j) & \text{if } n_{\theta}^o = 0. \end{cases} \quad (30)$$

By the zero profit condition for active firms (3), $(b_{\theta} - c_{\theta}) q_{\theta} \mathbb{P}_{\theta} = \gamma_{\theta}$ if $n_{\theta} > 0$; but $(b_{\theta} - c_{\theta}) q_{\theta} \mathbb{P}_{\theta} \leq \gamma_{\theta}$ for inactive sectors (see (4)). This means that the market allocation solves (30), and so induces the maximal $\tilde{W}(\cdot)$ and hence the maximal $W(\cdot)$ under the constraint. In other words, as per (23), $\frac{\partial W}{\partial n_{\theta}} = 0$ by the zero profit condition for the highest-priced sender in sector θ , and so the equalization condition is guaranteed at the equilibrium N^e .

References

- [1] ANDERSON, SIMON P. AND ANDRE DE PALMA (2009) Information Congestion. *RAND Journal of Economics*, 40(4), 688-709.

- [2] ANDERSON, SIMON P., ANDRE DE PALMA AND JACQUES-FRANCOIS THISSE (1992) *Discrete Choice Theory of Product Differentiation*. MIT Press.
- [3] BAGWELL, KYLE (2007) The Economic Analysis of Advertising. In Mark Armstrong and Rob Porter (eds.) *Handbook of Industrial Organization*, Vol. 3, 1701-1844. Elsevier, Amsterdam, North Holland.
- [4] BASS, FRANK (1969) A new product growth model for consumer durables. *Management Science*, 15(5), 215-227.
- [5] BAYE, MICHAEL R. AND JOHN MORGAN (2001) Information Gatekeepers and the Competitiveness of Homogeneous Product Markets. *American Economic Review*, 91(3), 454-474.
- [6] BAYE, MICHAEL R. AND JOHN MORGAN (2002) Information Gatekeepers and Price Discrimination on the Internet. *Economic Letters*, 76, 47-51.
- [7] BURDETT, KENNETH AND KENNETH L. JUDD (1983) Equilibrium Price Dispersion. *Econometrica*, 51(4), 955-969.
- [8] BUTTERS, GERARD R. (1977) Equilibrium Distributions of Sales and Advertising Prices. *Review of Economic Studies*, 44, 465-491.
- [9] CHAMBERLIN, EDWARD H. (1933) *The Theory of Monopolistic Competition*. Cambridge, Mass.: Harvard University Press.
- [10] DEBREU, GERARD (1960) Review of R. D. Luce, Individual Choice Behavior: A Theoretical Analysis. *American Economic Review*, 50, 186-188.
- [11] DIXIT, AVINASH K. AND VICTOR D. NORMAN (1978) Advertising and Welfare. *Bell Journal of Economics*. 9, 1-17.
- [12] EPPLER, MARTIN J. AND JEANNE MENGIS (2004) The Concept of Information Overload: a Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines. *The Information Society*, 20, 325-344.

- [13] FALKINGER, JOSEF (2008) A welfare analysis of "junk" information and spam filters. Working Paper 0811, University of Zurich.
- [14] GROSSMAN, GENE M. AND SHAPIRO, CARL (1984) Informative Advertising and Differentiated Products. *Review of Economic Studies*, 51, 63-81.
- [15] JOHNSON, JUSTIN (2008) Targeted Advertising. Working Paper, Cornell University.
- [16] LANHAM, RICHARD A. (2006) *The Economics of Attention: Style and Substance in the Age of Information*. University of Chicago Press.
- [17] LUCE, R. DUNCAN (1959) *Individual Choice Behavior: A Theoretical Analysis*. New York, Wiley.
- [18] SHENK, DAVID (1997) *Data Smog: Surviving the Information Glut*. HarperEdge.
- [19] SHAPIRO, CARL (1980) Advertising and Welfare: Comment. *Bell Journal of Economics*, 11 (2), 749-752.
- [20] SIMON, H. A. (1971) Designing Organizations for an Information-Rich World. In Martin Greenberger, *Computers, Communication, and the Public Interest*. Baltimore, MD: The Johns Hopkins Press, ISBN 0-8018-1135-X.
- [21] STEGEMAN, MARK (1991) Advertising in Competitive Markets. *American Economic Review*, 81(1), 210-23.
- [22] VAN ZANDT, TIMOTHY (2004) Information Overload in a Network of Targeted Communication, *RAND Journal of Economics*, 35(3), 542-560.