Competition for attention in the information (overload) age

Simon P. Anderson and André de Palma*

March 2006, revised December 27 2008.

Abstract

Limited consumer attention limits product market competition: prices are stochastically lower the more attention is paid. Ads compete to be the lowest price with other ads from the same sector and they compete for attention with ads from other sectors: equilibrium sector ad shares under free entry follow a CES form. When a sector gets more attractive, its advertising expands: others lose ad market share but may increase in absolute terms if sufficiently attractive. The information hump shows highest ad levels for intermediate attention levels when there is a decent enough chance of getting the message across and also of not being undercut by a cheaper offer. The Information Age takes off when the number of sectors grows, but total ad volume reaches an upper limit. Overall, advertising is excessive, though the allocation across sectors is optimal. Nonetheless, both large sectors and small ones can be blamed for misallocation of ads in using up scarce attention.

JEL CLASSIFICATION: D11, D60, L13, IO.

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

*Simon P. Anderson: Department of Economics, University of Virginia, PO Box 400182, Charlottesville VA 22904-4128, USA, sa9w@virginia.edu. André de Palma: Institut Universitaire de France, Département Economie et Gestion, Ecole Normale Supérieure, 61 Ave du Président Wilson, 91235 FRANCE. depalma@ens-cachan.fr. The first author gratefully acknowledges research funding from the NSF under grant SES-0452864 and from the Bankard Fund at the University of Virginia. We thank the Autoridade da Concorrencia 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), and seminar participants at the Sauder School (UBC), Stern School (NYU), and the University of North Carolina (Chapel Hill) are gratefully acknowledged.

1 Introduction

According to a Wiki cite, perhaps the first person 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. 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 the dispersion of prices in the face of endogenous information congestion where information from each sector competes within the sector and with each other sector.

The Information Age is naturally captured in our framework as a result of several forces. One is a lower cost of sending information. This is in part because there are now more (and cheaper) channels to reach potential consumers. Traditional billboards and newspaper ads are now 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. These tend to be the ones that expand most. 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. However, scarcity of attention brings spillovers into other sectors, like raising their prices and making it more

likely interesting offers are missed.

New products are also responsible for churn in the other advertising sectors, and pricing churn too. A new sector tends to drive down existing sectors' ads, at least in relative terms (as a fraction of the total volume of messages), and it drives down weaker sectors in absolute terms. It may though 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. Insofar as it is cheaper to access information, and especially since both work and leisure time are spent increasingly on information intensive activities, it may be argued that consumer attention spans have risen. This may induce more or less information, as a function of the level of attention, an effect we term the information hump. When consumer attention is sparse, little information will be sent because there is not much chance of getting a look-in. This in turn implies that prices will be quite high because there is a small chance of running across a rival in the same production sector. Hence prices are nearer monopoly levels. With a lot of attention, again not much information is sent because there is a good chance the consumer will get a better offer from the same sector, meaning that the benefit from sending a message is low. In this case, prices will be low too. The middle ground 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.

Another interesting feature we can track as a consequence of improved attention is the distribution of messages across sectors. With low levels of attention, highly profitable sectors will be most prominently represented due to the high chance of making a sale despite a high price. Increasing consumer attention brings firms into more competition with each other, which drives down sector profitability and serves to equalize opportunities across sectors. Ultimately then, the distribution of ads becomes more even across sectors, and this is in conjunction with a general lowering of mark-ups.

Hence, we expect that more attention drives lower prices. The same holds for improved communication costs, though with the caveat that if improvements are sector specific, then lower prices prevail within the sector but the crowding of other sectors' messages can relax competition there (and raise prices).

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), who is best remembered for his symmetric analysis. 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 of the descriptor 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 (2007). Reception of messages is passive: the consumer does not search. This corresponds to 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, we suppose that the individual screens out some of the messages received. 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.

Anderson and de Palma (2007) 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. 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.[1] However, one concern with this conclusion is that message senders do not compete directly in the marketplace, they just compete for attention. That is, direct business-stealing effects are closed down in that model. A tax might *a priori* reduce price competition by reducing message volume, and so harm consumer welfare. This is one question investigated here by specifically modeling competition within each of several sectors vying for consumer attention. To focus on firm competition necessitates simplifying the consumer side of the model: it is assumed here the consumer's attention span is exogenously fixed.

[1] However, if the consumer's attention is not congested, a tax may worsen the allocation insofar as messages senders do not internalize the consumer surplus from contacting prospective clients.

We first characterize an equilibrium model with interaction both within and across sectors. Competition within a sector is reflected in the likelihood that a lower price is more likely to be the lowest 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. Lower transmission costs in one sector increase messages sent from there, which congests messages from elsewhere. This has ambiguous effects on the price dispersion in the other industries.

Surprisingly, the model endogenously generates a CES form for the number of messages sent across industries. This bears an intriguing parallel to the CES utility functional form so often used to parameterize Chamberlinian models. Information congestion gives a new 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.[2] 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 is one possible outcome in the Grossman and Shapiro (1983) model of informative advertising when the business stealing effect outweighs the consumer surplus one.[3] Along similar lines, Stegeman (1991) shows that the market allocation is insufficient when the Butters model is amended to allow demand to have some elasticity. This is because firms then tend to overprice without sufficient regard to the consumer surplus lost. In our context, over-advertising is quite natural as it serves to dissipate rents.

[2] This finding reinforces the conclusion of Anderson and de Palma (2007), where the optimal policy in the presence of congestion was a tax on transmission.

[3] Excessive advertising is also found in the controversial Dixit and Norman (1979) paper on persuasive advertising.

In the next Section, we describe the background model and solution technique. Section 3 derives the CES form for total advertising, derives the aggregates in the model and discusses their properties for information level changes. 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 optimal allocation property and the tax prescription to deal with over-advertising. Section 6 describes an extension to allow for distractions, while Section 7 discusses more general demand and endogenous examination levels. Section 8 concludes. The Appendix gives a quick reminder of the Butters (1977) model.

2 Message reception and transmission

2.1 Assumptions

There are $\Theta > 1$ active commercial sectors indexed by $\theta = 1, ..., \Theta$.[4] Each active sector comprises a continuum of firms. These firms post messages about their products, which can be seen by consumers (although the latter do not necessarily register the information). Each firm's message indicates the price at which the product can be bought. Products are homogeneous within sectors, and so a consumer buys at the lowest price she has registered, if at all (demand is split equally in case of a tie), and she cannot buy if she has not registered any price from the sector.

Consumers are assumed to be identical with respect to their reservation prices for each sector, and the cost of potentially reaching any consumer is the same. For example, messages could be sent by mail to consumers, or they could be posted on billboards seen by many consumers, or on TV programs. The cost of reaching a consumer from sector θ is γ_θ, which can therefore represent the cost of a letter, or the cost of a billboard divided by the number of consumers reached. 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 therefore treat the consumer as the unit of analysis and we therefore henceforth refer to a single consumer.[5]

[4] In Section 3.6 we discuss how this number can be endogenously determined.

[5] Allowing for multiple consumer types would be useful for extending the model to analyze consumer targeting.

Firms within each sector produce homogenous goods, and each sector transmits an endogenous number of messages n_θ for a total number of $N = \sum_{\theta=1}^{\Theta} n_\theta$ messages (per consumer). Each active advertiser sends just one message per consumer.[6] Hence n_θ also represents the number of firms in sector θ. Each message is an advertisement containing the price at which the consumer can buy the product from the sending firm. The consumer registers a fixed number of messages, $\phi \geq 1$, but does not emphasize any sector over the others. This reflects limited information processing capability. Hence, the ϕ messages are drawn at random from the N messages sent (the message anonymity assumption). In what follows we will assume for the sake of capturing information the advertising clutter / information congestion that $\phi < N$.

After registering the ϕ messages, the consumer chooses the lowest priced one 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. We later allow the conditional demand to depend on price, but for the moment demand is rectangular.

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.

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

2.2 Solution technique

A firm's expected demand depends on 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 therefore 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.

The highest price set by any firm (the consumer reservation price for the sector, b_θ) plays a key role because the only way a sender can make a profit at such a price is if it is the only message drawn from that

[6] Indeed, 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.

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.

In the sequel we exploit the sequential resolution of the model by first determining (in the next Section) aggregate numbers of messages per sector and total messages, and then describing price dispersion within each sector (in the following Section).

2.3 Message selection probability

We first seek the probability of registering one particular message and registering no other message from the sector, θ, it came from. Assume $n_\theta < N$ (so at least two sectors are active). In the development in the main text we consider choice with replacement. This corresponds (loosely) to being exposed to a constant stream of messages, with repetition (e.g., billboards on a commute repeated daily.) In a later footnote, we develop the appropriate expressions for choice without replacement; which might represent going through the day's bulk mail or email. We show that both formulations give the same choice probability, under the proviso that ϕ is small relative to N, which is the case we consider.

The probability that the first message drawn is the one under consideration is $1/N$. If draws are taken with replacement, the probability that none of the $n_\theta - 1$ other messages in the sector is registered on the subsequent $\phi - 1$ draws is $\left(1 - \frac{n_\theta - 1}{N}\right)^{\phi-1} \approx \left(1 - \frac{n_\theta}{N}\right)^{\phi-1}$.[7] The probability that the message under consideration is drawn is ϕ/N. If ϕ is small relative to N, then there is a negligible probability this same message is drawn twice (or more). Then, for $\phi << N$, the probability $\mathbb{P}_\theta$ that one (specific) message from

[7] Without replacement, the probability that the first of the $\phi - 1$ other draws is not from sector θ is $\left(1 - \frac{n_\theta - 1}{N-1}\right) \approx \left(1 - \frac{n_\theta}{N}\right)$. By extension, the probability that *none* of the other $\phi - 1$ messages gets through, assuming $\phi << N$, is

$$\left(1 - \frac{n_\theta - 1}{N-1}\right)\left(1 - \frac{n_\theta - 1}{N-2}\right) ... \left(1 - \frac{n_\theta - 1}{N - (\phi - 1)}\right) \approx \left(1 - \frac{n_\theta}{N}\right)^{\phi-1}.$$

With is a large number of messages, drawing one message does not noticably change the residual number of messages in the sector. On the other hand, because ϕ is an integer, one draw does significantly reduce the number of other draws left.

sector θ is registered (and no other message is registered from that sector) is thus:[8]

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

This is conveniently rewritten as $\mathbb{P}_\theta = \frac{\phi}{N-n_\theta}\left(1 - \frac{n_\theta}{N}\right)^{\phi}$, which can also be expressed as[9]

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

where

$$\mathbb{Q}_\theta = 1 - \left(1 - \frac{n_\theta}{N}\right)^{\phi}. \tag{2}$$

Here $\mathbb{Q}_\theta$ is readily seen as the probability that there is *at least one hit* in the sector (θ): the probability that each of the n_θ messages is missed on each of the ϕ draws. The second important link between the two probabilities is that $\frac{\partial \mathbb{Q}_\theta}{\partial n_\theta} = \mathbb{P}_\theta$. That is, 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.

3 Advertising levels

3.1 Advertising shares by sector

Consider an advertisement which is sent out for a price equal to the reservation price b_θ. As we argue in Section 4 below, there will be such an ad, and the probability of finding a second ad at the same price is zero. Since $\mathbb{P}_\theta$ (as given by (1)) is the probability of this ad being the only one found from its sector, the equilibrium zero profit condition reads:

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

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

$$\pi_\theta = \frac{\left(b_\theta - c_\theta\right) q_\theta}{\gamma_\theta} > 1,$$

[8] We will later use the notation $\mathbb{P}(p, \theta)$ to denote the probability of a sale at price p in sector θ: hence in terms of our later notation $\mathbb{P}_\theta = \mathbb{P}(b_\theta, \theta)$, 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.

[9] The probability of a single particular message being chosen is here seen as the probability that the sector is not chosen (second term) times the probability that the message is chosen given ϕ draws *outside* the sector (first term). This makes sense once one realizes that the individual message in question, being "small", can just as well be effectively housed initially outside the sector.

which measures the economic potential (surplus per \$ transmission cost) of sector θ. Let $\bar{\Theta}$ be the number of sectors for which $\pi_\theta > 1$: this is the number of *potentially active sectors* (this will be the number of sectors sending positive numbers of messages if ϕ goes to infinity). We rank these sectors such that π_θ is decreasing in the index θ, i.e. from highest to lowest 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, which we denote $\Theta \leq \bar{\Theta}$. 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 wish to incur the cost of a message, given that messages are not read with certainty.

The zero profit condition (3) for the equilibrium probability the highest-priced sender in active sector θ makes a sale is then written as

$$\mathbb{P}_\theta = \frac{1}{\pi_\theta}. \tag{4}$$

This probability depends only on the intrinsic economic performance index, π_θ, of the sector. Equating this probability, derived from the zero-profit condition, with the probability as derived from the individual's sampling that she gets no other message from the subset in the sector determines the ad market shares.

An equilibrium to the model maps the primitives $\left\{\phi, \{\pi_\theta\}_{\theta=1,\ldots,\bar{\Theta}}\right\}$ into a set of non-negative sector message numbers $\{n_\theta\}_{\theta=1,\ldots,\bar{\Theta}}$ which define the total message volume as $N = \sum_{\theta=1}^{\bar{\Theta}} n_\theta$. A sector is active if and only if $n_\theta > 0$. For each active sector, the equilibrium specifies sector purchase probabilities, $\mathbb{P}_\theta$, for the consumer, and a price distribution within each sector, and corresponding choice probabilities for each product $\mathbb{P}(\theta, p)$. We show below that equilibrium is unique, and involves the first Θ sectors, where the cut-off Θ between active and inactive sectors is determined below. We proceed by determining message volume by sector as a function of an arbitrary number, N, of total messages, for any given N. There are two interpretations to this exercise. First, it characterizes the sector numbers at any equilibrium N. Second, it sets up the conditions that must hold for N to be an equilibrium number since in equilibrium we must have $N = \sum_{\theta=1}^{\bar{\Theta}} n_\theta$. This second step is the object of Lemma 5 below.

Lemma 1 *Fix $N > \phi$. All θ such that $\pi_\theta > \frac{N}{\phi}$ are active sectors. The relative sector sizes in message*

distribution are

$$\frac{n_\theta}{N} = 1 - \left(\frac{N}{\phi}\frac{1}{\pi_\theta}\right)^{\frac{1}{\phi-1}} > 0, \qquad \theta = 1, ..., \Theta. \tag{5}$$

and $n_\theta = 0$ *otherwise.*

Proof. Equating (1) and (4) implies $\mathbb{P}_\theta = \frac{\phi}{N}\left(1 - \frac{n_\theta}{N}\right)^{\phi-1} = \frac{1}{\pi_\theta}$, which is rewritten as (5). Hence, sector θ sends a positive number of messages if and only if $\pi_\theta > \frac{N}{\phi}$. ■

If $\pi_\theta < \frac{N}{\phi}$, then even a single message sent from the sector at the highest price would not be expected to cover its costs: i.e., $(b_\theta - c_\theta)\, q_\theta \frac{\phi}{N} < \gamma_\theta$, where $\frac{\phi}{N}$ is the probability the message is registered.[10] We defer considering the overall comparative static properties of equilibrium because N is still to be determined in (5). However, we can use the expression to compare across sectors of different economic characteristics within an equilibrium.

Sectors with larger economic potential 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 An (non-)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 (a characteristic of the multinomial logit discrete choice model 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 (8) 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 competitors ($m_{-\theta} \equiv \frac{N-n_\theta}{N}$

[10] As we shall see below, this is also the condition for the lowest price in the price support to be below the reservation price. (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 low price is critically at b_θ, the condition follows immediately.)

where n_θ is the number of messages from sector θ). From (5), the present IIA property is:[11]

$$\frac{m_{-\theta}}{m_{-\theta'}} = \left(\frac{\pi_{\theta'}}{\pi_\theta}\right)^{\frac{1}{\phi-1}}. \tag{6}$$

This is a property of invariance of the ratio of *all rivals'* advertising levels as the appeal of any rival (outside the pair) changes. The property implies the following Proposition.[12]

Proposition 2 *At any equilibrium with Θ active sectors, the non-θ shares can be written in a logit form:*

$$\frac{m_{-\theta}}{(\Theta-1)} = \frac{\pi_\theta^{\frac{-1}{\phi-1}}}{\sum_{\theta'=1}^{\Theta} \pi_{\theta'}^{\frac{-1}{\phi-1}}} \equiv \Psi_\theta, \quad \theta = 1, ..., \Theta. \tag{7}$$

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

Proof. Inverting (6),

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

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

Some care must be taken with the interpretation of the result. In particular, the value of Θ is endogenous here (and is determined below), and so only the active sectors are counted: inactive sectors π_θ must be excluded from the summation.[13] The same caveat applies below.

As π_θ increases, the RHS of (7) falls: the more attractive is a sector, then the more its ads push out the ad shares of other sectors. That is, as profitability rises, the affected sector produces proportionately more ads while the others produce relatively less.[14] One reason for a higher "profitability" in a mature sector is a lower cost of communication in the sector (lower γ_θ). New media are more appropriate for advertising some goods when the delivery of the message is complemented by the medium. The model says that sectors which benefit from such improved communication costs get larger ad market shares at the expense of the others. Indeed, as shown in sections 3.6 and 3.7, weak sectors might be pushed out of the market entirely.

[11] It would be interesting to check empirically whether this property is satisfied when the standard IIA property is violated (see Train, 2003, for a discussion of the Hausman test for IIA in the context of discrete choice models).

[12] Therefore $\frac{n_\theta}{n_{\theta'}} = \frac{1-(\Theta-1)\Psi_\theta}{1-(\Theta-1)\Psi_{\theta'}}$, which indicates that IIA does not hold.

[13] This is true too in the standard logit insofar as only available options are included when determining choice probabilities across options.

[14] We see in Section 3.7 that the number of ads from sector θ' may actually rise if that sector is sufficiently attractive.

There is also an effect of raising ϕ on the distribution of messages by sector. The properties below are fundamentally those of the logit formulation (see for example Anderson, de Palma, and Thisse (1992)). However, the derivation of that form above is very different from the usual roots.

Proposition 3 *For Θ constant, as ϕ rises, the ad market share of sector 1 decreases with ϕ, and the share of sector Θ increases. As ϕ falls to 1, almost all messages are sent by sector 1.*

Proof. To show the first point, fix Θ. The relation in (7) gives the fraction of messages in sector θ as

$$\frac{n_\theta}{N} = 1 - (\Theta - 1)\,\Psi_\theta.$$

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

$$\frac{d\Psi_\theta}{d\phi} \overset{s}{=} -\ln\frac{1}{\pi_\theta} + \sum\nolimits_{\theta'=1}^{\Theta} \Psi_{\theta'} \ln \frac{1}{\pi_{\theta'}},$$

or

$$\frac{d\Psi_\theta}{d\phi} \overset{s}{=} \sum\nolimits_{\theta'=1}^{\Theta} \Psi_{\theta'} \left(\ln\frac{1}{\pi_{\theta'}} - \ln\frac{1}{\pi_\theta} \right).$$

Hence, the share decreases with ϕ for the first sector, and increases for the last viable one.

Now consider the limit case $\phi \downarrow 1$. We show that the limit involves a single viable sector. First note from (7) that

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

Hence, $\lim\limits_{\phi\downarrow 1} \Psi_1 = 0$: almost no messages are sent from sectors other than sector 1. ■

When $\Theta > 2$, the advertising shares of the intermediate sectors are not necessarily monotonic in ϕ. To see this, consider 3 sectors. Sectors 1 and 2 have very high profits, with 2 slightly less efficient 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 in this example.

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. As we shall see below, the number

of messages sent from this sector tends to π_1.[15] This corresponds to pure dissipation of the monopoly profit in sector 1. 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. 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.[16]

At the other extreme, when the attention span is extensive, any message greater than 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.

Expression (7) in turn begets the choice probabilities themselves, and gives rise to a familiar functional form.

3.3 Aggregate advertising

The next step is to determine the equilibrium market size, N. Expressions (5) and (7) give two different expressions for m_{-i}. Equating them yields the expression in the next proposition.[17]

Proposition 4 *The equilibrium total message size given $\Theta > 1$ active sectors is a CES form:*[18]

$$N = \phi\,(\Theta - 1)^{\phi-1} \Big/ \left(\sum_{\theta'=1}^{\Theta} \pi_{\theta'}^{\frac{-1}{\phi-1}}\right)^{(\phi-1)} . \tag{8}$$

Thus N is increasing in each profitability, π_θ, and homogenous of degree one in the sector profitabilities.

The CES form has well-known properties. The first property means that raising the profitability of any sector causes the total volume of messages to rise: the extra clamor causes a larger total without a fully compensating backlash from the other sectors. Similarly, adding another viable sector raises N. To see

[15] Indeed, from (8) below, as ϕ falls to 1, the number of messages sent from a single sector tends to $N = \pi_1$.

[16] As we shall see below, if all sector transmission costs fall proportionately to zero, 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.

[17] This can also be derived from summing up the expressions for market shares in (5).

[18] Therefore, it is maximal at symmetry (under the constraint that the sum of the inverse π_θ's is constant).

the second point, consider introducing a "barely viable" sector s with $n_s = 0$: by the argument following Lemma 1, the corresponding attractivity of such a new sector s is $\pi_s = N/\phi$. We now verify (as is true by construction of (8)) that introducing this barely viable sector s leaves (8) unchanged.

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

Thence, introducing a strictly viable sector, with $\pi_s > N/\phi$, will cause N to increase.[19]

The homogeneity property in Proposition 4 implies that total message volume doubles if all communication costs are halved.[20] 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.[21]

We next consider the symmetric case before going into more detail into the asymmetric one.

3.4 The Information Age

One driver of the information age is lower communication costs, another is a larger set of viable sectors. Under symmetry ($\pi_\theta = \pi$ for all θ), the expression (from (8)) for the total number of messages, N, reduces to[22]

$$N = \phi \left(\frac{\Theta - 1}{\Theta}\right)^{\phi-1} \pi. \tag{9}$$

Having more sectors, Θ, raises the total number of messages. The number N is a logistic function of the number of sectors: it is first convex (for $\Theta < \phi/2$), and then concave, for $\Theta > \phi/2$. 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. Indeed, the amount of

[19] That is, $\frac{N}{\phi} = \frac{\Theta^{\phi-1}}{\left(\sum_{\theta'=1}^{\Theta} \pi_{\theta'}^{\frac{-1}{\phi-1}} + \left(\frac{N}{\phi}\right)^{\frac{-1}{\phi-1}}\right)^{\phi-1}} < \frac{\Theta^{\phi-1}}{\left(\sum_{\theta'=1}^{\Theta} \pi_{\theta'}^{\frac{-1}{\phi-1}} + \pi_s^{\frac{-1}{\phi-1}}\right)^{\phi-1}}$ ..

[20] No 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 $\pi_s > N/\phi$.

[21] This is reminiscent of Zahavi's Law in transportation, which says that average travel times have remained constant over several decades, despite substantial increases in travel speed.

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

information has an asymptote of $\bar{N} = \phi\pi$, which is the bound to the amount of information the system can sustain.[23]

The average number of messages per sector, $n_\theta = N/\Theta$, is increasing in Θ if and only if $\phi > \Theta$, so it is eventually decreasing (for Θ large enough), coherently with what we have above. The interesting feature here is the initial increase. This is explained by the idea that more sectors mean less competition, so higher prices and more incentive to send messages.

The logistic function in (9) is sketched in Figure 1, for $\pi = 20$ and $\phi = 20$ (the asymptote of the function is at $N = 400$, the maximal value of N/Θ is attained at $\Theta = 20$, and the inflection point is at $\Theta = 10$).

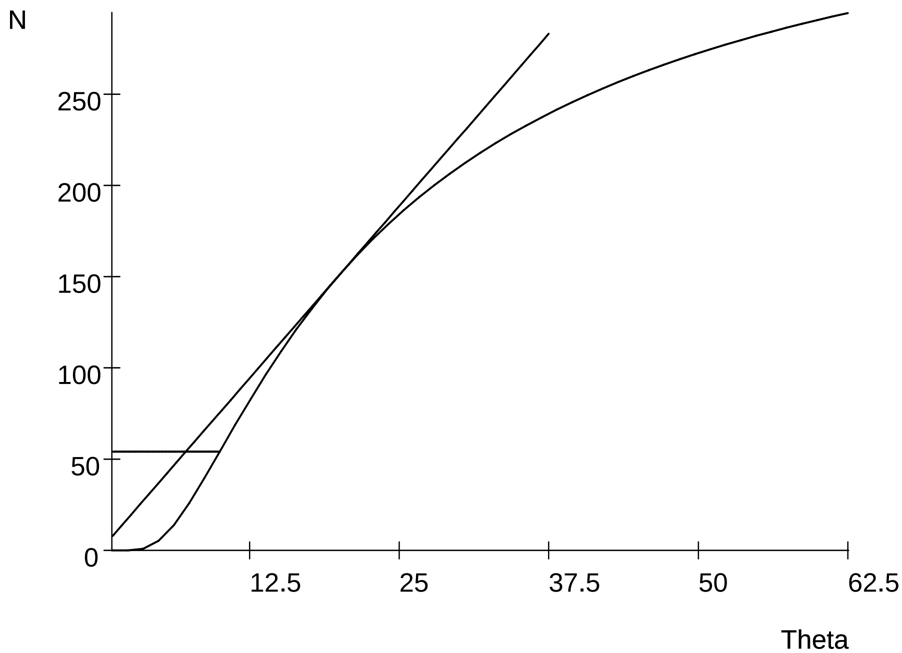

Figure 1. Total messages as a function of number of sectors.

The other comparative static property of N, with respect to ϕ, is described next.

[23] At the limit, monopoly prices, b_θ, are set in each sector, returning π when the message is chosen. The probability of being chosen is $\phi/\bar{N}$, which therefore equals $1/\pi$.

3.5 The information overflow hump

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

From the symmetric analysis above, (see (9)), we can derive how the information level, N, varies with the consumer span, ϕ. Indeed, N is decreasing in attention span ϕ if and only if $\phi > \hat{\phi} \equiv \frac{1}{\left(\ln \frac{\Theta}{(\Theta-1)}\right)}$, and so it is necessarily decreasing for $\phi > \Theta$ (since $\Theta > \ln \frac{\Theta}{(\Theta-1)}$). Likewise, $\frac{N}{\phi}$ is falling in ϕ, and therefore N increases more slowly than the number of messages examined, ϕ.

Figure 2 plots the relation of N as a function of ϕ for $\pi = 100$ and $\Theta = 10$ (hence $N = 100\phi\left(\frac{9}{10}\right)^{\phi-1}$ attains its maximum at $\hat{\phi} \equiv \frac{1}{\left(\ln \frac{10}{9}\right)}$, which is slightly less than 10). The dashed line is the line $\phi = N$.

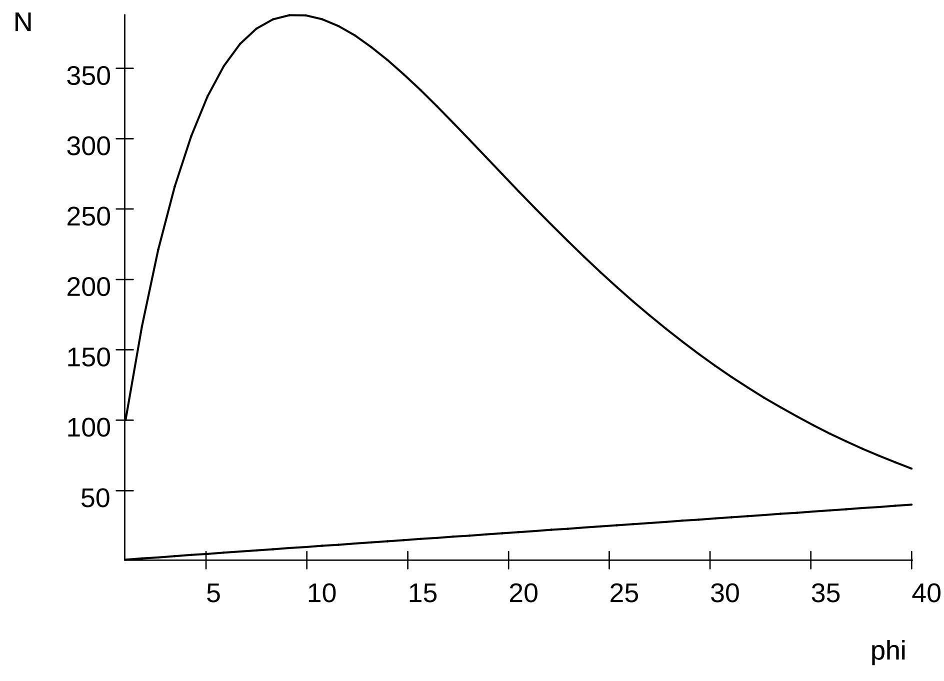

Figure 2. Total number of messages sent, N, as a function of examination, $\phi \geq 1$.

The Figure shows the quasi-concave function, i.e., first increasing, then decreasing with the reception

level, ϕ. This we term the *information overflow hump.* However, the number of messages only increases for low $\phi < \hat{\phi}$ ($< \Theta$). More reception has two conflicting consequences. On the one hand, this means that the individual reads more, and more examination in a sector raises the probability of a message from the sector being seen, which raises profitability, and hence the number of messages sent, ceteris paribus. But it also has the effect of increasing price competition insofar as 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 Θ, 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 therefore decreases sector attractivity.

The *degree* of information overflow, N/ϕ, is the extent to which messages are missed.[24] This overflow impacts the degree of competitiveness in the individual markets, and the equilibrium price dispersion per market, as elaborated in the next section. Markets are linked through overall competition for attention.

3.6 Sector viability

When sectors are asymmetric, some may be precluded from the market by the strength of those in the market. We now make precise the conditions for sectors to be active.

Let $\bar{\Theta}$ denote the number of sectors for which $\pi_\theta > 1$, and assume that $\bar{\Theta} > 1$.[25] Any sector with $\pi_\theta \leq 1$ is not viable, and so can be eliminated from the discussion (even if a message sent at the sector monopoly price were examined exclusively with probability one, it would not generate a profitable sale).

Lemma 5 *Assume that $\bar{\Theta} > 1$. Then there exists a unique equilibrium where all sectors $1, ..., \Theta$ are active, with $\Theta \in \left[2, \bar{\Theta}\right]$, and the total volume of messages is given by (8).*

Proof. From Lemma 1, a sector is active in equilibrium if $\pi_\theta > \frac{N_\Theta}{\phi}$, where we (temporarily) let N_Θ denote

[24] The degree of information overflow can also be initially rising with ϕ, but for even smaller values of ϕ.

[25] The model is degenerate if there is a single sector. From (8), there are zero messages.

the number of messages for Θ active sectors as given by (8). We first show that if there are two sectors, then they are both active. From (8), $\frac{N_2}{\phi} = \frac{1}{\left(\pi_1^{\frac{-1}{\phi-1}} + \pi_2^{\frac{-1}{\phi-1}}\right)^{\phi-1}}$, and the RHS is less than π_2 (as is readily seen by cross-multiplying), so this implies $\pi_2 > \frac{N_2}{\phi}$, as desired for the second sector to be active.

Next, we show there is a unique sector cut-off, Θ. The condition for a sector to be active is $\pi_\theta > \frac{N_\Theta}{\phi}$. Given the ranking of sectors, the LHS decreases in the marginal sector, Θ, while we showed in the argument following Proposition 4 that the RHS increases as sectors are added. Note that all $\bar{\Theta}$ sectors are active if $\pi_{\bar{\Theta}} > \frac{N_{\bar{\Theta}}}{\phi}$ (which condition we showed to hold in the symmetric equilibria analyzed previously).

Finally, it remains to show that the equilibrium does indeed follow the ranking: that is, there cannot be an equilibrium with some sector θ excluded while some sector $\theta' > \theta$ is included. Suppose 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 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.[26] ■

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

3.7 Raising a sector's profitability

We noted in Section 3.2 that an increase in a sector's profitability will increase the total number of messages sent (see (8)). 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 $\frac{n_\theta}{N} = 1 - \left(\frac{N}{\phi}\frac{1}{\pi_\theta}\right)^{\frac{1}{\phi-1}}$ from (5). 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

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

positive). Indeed, since $\frac{dN}{d\pi_{\theta'}} > 0$, then $\frac{dn_\theta}{d\pi_{\theta'}} = \frac{dn_\theta}{dN}\frac{dN}{d\pi_{\theta'}}$ has the sign of $\frac{dn_\theta}{dN}$. From (5), we have the derivative[27]

$$\begin{aligned}\frac{dn_\theta}{dN} &= 1 - \frac{\phi}{\phi-1}\left(\frac{N}{\phi}\frac{1}{\pi_\theta}\right)^{\frac{1}{\phi-1}} \\ &= 1 - \frac{\phi}{\phi-1}\frac{(\Theta-1)}{\Theta}\frac{\pi_\theta^{\frac{-1}{\phi-1}}}{\left(\sum_{\theta'=1}^{\Theta}\pi_{\theta'}^{\frac{-1}{\phi-1}}\right)/\Theta},\end{aligned}$$

where we have substituted $\frac{N}{\phi}$ from (8). Define $\chi_\theta = \pi_\theta^{\frac{-1}{\phi-1}}$ and so

$$\frac{dn_\theta}{dN} = 1 - \frac{\phi}{\phi-1}\frac{(\Theta-1)}{\Theta}\frac{\chi_\theta}{\bar{\chi}} \tag{10}$$

where the average value of χ_θ, denoted by $\bar{\chi}$, is homogenous of degree $\left(\frac{-1}{\phi-1}\right)$ in the π_θ.

From a symmetric starting point (where $\chi_\theta = \bar{\chi}$ for all θ), $\frac{dn_\theta}{dN}$ has the sign of $1 - \frac{\phi}{\phi-1}\frac{(\Theta-1)}{\Theta}$, which is negative if and only if $\phi > \Theta$. If though $\phi < \Theta$, 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 (which arises for low ϕ) and for the early "take-off" part of the Information Age evolution depicted in Figure 1 (for low Θ). In all cases, 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, (10) indicates that there is a cut-off value of χ_θ for which $\frac{dn_\theta}{dN}$ is negative for higher χ and positive for lower χ. Since π is inversely related to χ, this means that *larger sectors are more likely to see an increase* in the number of messages sent. A summary Proposition:

Proposition 6 *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 absolutely.*

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

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

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 (although the aggregate message volume, N, and attention span, ϕ, both matter). 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

$$\mathbb{P}(p,\theta) = \frac{\gamma_\theta}{(p-c_\theta)\, q_\theta} = \frac{(b_\theta - c_\theta)}{(p-c_\theta)} \frac{1}{\pi_\theta}, \tag{11}$$

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 (4), 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. We first argue that the support of the equilibrium advertised price distribution (for any active sector) is a compact interval $\left[\underline{p_\theta}, b_\theta\right]$ with no atoms nor gaps, where the lower bound, $\underline{p_\theta}$, 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_θ because at price c_θ the transmission cost cannot be recuperated.

Lemma 7 *Prices in industry θ are distributed on a compact support $\left[\underline{p_\theta}, b_\theta\right]$ and there are no atoms, where $\underline{p_\theta} > c_\theta + \gamma_\theta$.*

We now derive the lowest price in the support along with the equilibrium advertised price distribution. Let $F(p,\theta)$ denote the *fraction* of messages in sector θ sent at price p or below. (Then $F\left(\underline{p_\theta},\theta\right) = 0$ 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

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

where we simply note that the number of messages sent from the sector with a price no higher than p is $n_\theta F(p, \theta)$.[28] Since $\mathbb{P}(p, \theta)$ is given by the zero profit condition (11), we have

$$\frac{(b_\theta - c_\theta)}{(p - c_\theta)} \frac{1}{\pi_\theta} = \frac{\phi}{N} \left(1 - \frac{n_\theta F(p, \theta)}{N}\right)^{\phi - 1}, \tag{12}$$

where n_θ / N is given above by (5).

Proposition 8 *The equilibrium advertised price density in sector θ is decreasing and convex on $\left[\underline{p_\theta}, b_\theta\right]$, with distribution*

$$F(p, \theta) = \frac{1 - \left(\frac{N}{\phi \pi_\theta}\right)^{\frac{1}{\phi - 1}} \left(\frac{b_\theta - c_\theta}{p - c_\theta}\right)^{\frac{1}{\phi - 1}}}{1 - \left(\frac{N}{\phi \pi_\theta}\right)^{\frac{1}{\phi - 1}}} \tag{13}$$

where N is given by (8) and $\underline{p_\theta}$ is given by

$$\underline{p_\theta} = c_\theta + \left(\frac{N}{\phi}\right) \frac{\gamma_\theta}{q_\theta}. \tag{14}$$

Proof. The equilibrium advertised price distribution is given from the relation (12) 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 (5), we can write (13). It is readily checked that $F(b_\theta, \theta) = 1$.

Since $F\left(\underline{p_\theta}, \theta\right) = 0$, the lowest price in sector θ is determined by (12) as:

$$\frac{(b_\theta - c_\theta)}{\left(\underline{p_\theta} - c_\theta\right)} \frac{1}{\pi_\theta} = \frac{\phi}{N}.$$

Then (14) follows immediately. The corresponding density, $f(p, \theta)$ is strictly positive on $\left[\underline{p_\theta}, b_\theta\right]$, where it is decreasing and convex (as shown by differentiation of (13)). ■

As in Butters (1977), lower prices are advertised more heavily. In the Butters model, the corresponding lowest price, $\underline{p}$, would be simply $c_\theta + \gamma_\theta / q_\theta$,[29] because such a price just covers the cost of production plus

[28] Clearly, expected sales per consumer of the cheapest priced product are $q_\theta \phi / N$ for sector θ.

[29] This allows for a quantity effect, which Butters does not have.

sending the message. In contrast to the Butters model, here the lowest price in any sector does not make a sale with probability 1. 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. Here, information overflow pushes up the lowest price in the support: this is needed to compensate for the likelihood that the message may not be received.

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 (14).

The simplest measure of price dispersion is the breadth of the support of the equilibrium prices. This is $b_\theta - \underline{p_\theta}$, where $\underline{p_\theta} = c_\theta + \left(\frac{N}{\phi}\right)\frac{\gamma_\theta}{q_\theta}$, from above. 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 (8)).

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 not collapse. That is $\underline{p_\theta} < b_\theta$. Writing out the condition, it means that $\frac{N}{\phi} < \pi_\theta$; this is the same condition from (5) for $n_\theta > 0$ in equilibrium.

The next two sub-sections stress the properties of the equilibrium price distribution with respect to two key variables of emphasis in the paper, sector profitability and consumer attention span.

4.1 Advertised price dispersion and sector profitability

Greater sector profitability impacts the affected sector by increasing the volume of messages sent (Proposition 4). 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. Nonetheless, the number of messages sent in other sectors can

actually rise (see Proposition 6) because the reduced price competition can raise profits per firm (which then must be reduced by further entry).

Proposition 9 *An increase in the attractivity 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 (13). 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 (8)). Hence, $F(p,\theta)$ increases in π_θ. If π_θ increases, $\underline{p_\theta}$ falls; if $\pi_{\theta'}$ increases, N rises so that $\underline{p_\theta}$ rises (see (14)).

If all sectors increase proportionately in attractivity, $\frac{N}{\pi_\theta}$ is unchanged (by the homogeneity in Proposition 4) 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 in those other sectors. On the other hand, across-the-board changes affecting all sectors can leave prices the same. The last property underlies the result in the next Section that benefits from proportionately lower message transmission costs 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 connection. It is derived in the Appendix. In the meantime, we follow through with the analysis of the symmetric case.

4.2 Dispersion and symmetric sectors

In the symmetric case, N is given by (9) as $N = \phi\left(\frac{\Theta-1}{\Theta}\right)^{\phi-1}\pi$, and so the cumulative distribution function for advertised prices (13) becomes

$$F(p,\theta) = \Theta\left(1 - \frac{\Theta-1}{\Theta}\left(\frac{b-c}{p-c}\right)^{\frac{1}{\phi-1}}\right), \quad \text{for } p \in [\underline{p}, b], \tag{15}$$

where $\underline{p} = c + \left(\frac{\Theta-1}{\Theta}\right)^{\phi-1}(b-c)$ (by (14)).[30] 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 examination rate, ϕ, lowers prices in the sense of First-Order Stochastic Dominance.*

Proof. From (15), $F(p,\theta,\phi_1) > F(p,\theta,\phi_2)$ as $\left(\frac{b-c}{p-c}\right)^{\frac{1}{\phi_1-1}} < \left(\frac{b-c}{p-c}\right)^{\frac{1}{\phi_2-1}}$, or $\phi_2 < \phi_1$. ■

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 examined, firms may as well set high prices and chance the low probability of another message from the same sector. In the second case, firms tend to set low prices because there is a strong likelihood another message from the same sector will be read, which intensifies price competition.

Along similar lines, it is readily shown that higher Θ 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

One strong property of the Butters (1977) model is the optimality of the market allocation. However, this property is crucially dependent on the fact in his set-up that each message hits somewhere.[31] In our set-up,

[30] As $\left(\frac{\Theta-1}{\Theta}\right)^{\phi-1}\tilde{\pi} \to 1$, then $N \to \phi$, and, (by (14)), $\underline{p}$ goes to $c + \frac{\gamma}{q}$: this is like the Butters price.

[31] It also depends on the rectangular demand assumption. Stegeman (1991) shows that there is insufficient advertising if demand slopes down, because the pricing distortion has firms not internalizing the consumer surplus of lower prices. We discuss downward sloping demand below.

there is rent dissipation and socially wasteful duplication of messages (to the tune of $N-\phi$).[32] Competition for attention imposes a congestion externality which leads to excessive advertising, which result is perhaps more in tune (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. Recall that $\mathbb{Q}_\theta$ is the probability of at least one hit in sector θ (see (2)), and write this as $\mathbb{Q}(n_\theta, N) = 1 - \left(1 - \frac{n_\theta}{N}\right)^\phi$, which depends only on own message fraction and the attention span. Let n denote the vector $n_1, ..., n_\Theta$; then we can write the form of the welfare function as

$$W(n, N) = \sum_{\theta=1}^{\Theta} \left[(b_\theta - c_\theta) q_\theta \mathbb{Q}(n_\theta, N) - \gamma_\theta n_\theta\right]. \tag{16}$$

This form is convenient for the demonstration that follows.

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

$$\frac{dW}{dn_\theta} = \frac{\partial W}{\partial n_\theta} + \frac{\partial W}{\partial N} \tag{17}$$

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

Proof. From (16), 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 (17).

Now, from (16), and then using (1), we have that

$$\begin{aligned} \frac{\partial W}{\partial n_\theta} &= (b_\theta - c_\theta) q_\theta \frac{\partial \mathbb{Q}(n_\theta, N)}{\partial n_\theta} - \gamma_\theta \\ &= (b_\theta - c_\theta) q_\theta \mathbb{P}_\theta - \gamma_\theta. \end{aligned} \tag{18}$$

This expression is the profit of a firm setting the top price in sector θ given n_θ messages emanating from the sector (see (4)). Since this is zero in equilibrium, the remaining term, $\partial W/\partial N$, is naturally interpreted as the congestion externality. ■

Applying this property yields the following result.

[32] Clearly the first best optimum comprises one message per sector, and the active sectors should be the first ϕ ones, meaning the ones for which π_θ is highest.

Proposition 12 *The equilibrium allocation of messages across sectors is optimal given the number of messages transmitted at the equilibrium. However, the total number of messages transmitted is excessive, and the congestion externality is measured as the average transmission cost.*

Proof. Let N be given at the equilibrium level stipulated by (8), $N = N^e$, and we wish to show that the division of these messages effectuated in equilibrium is optimal. Optimal choice of the message levels by sector under the constraint that they must sum to the equilibrium level, N^e, implies that $\frac{dW}{dn_\theta}$ must be equalized across sectors. From (17), it follows that the partial derivatives, $\frac{\partial W}{\partial n_\theta}$, are to be equalized, and the corresponding n_θ's must sum to N^e. As per (18), $\frac{\partial W}{\partial n_\theta} = 0$ by the zero profit condition for the highest-priced sender in sector θ, and so the equalization condition is guaranteed.

We must also show that there is no candidate solution other than the market equilibrium where the partial derivatives are equalized for the same $N = N^e$. This is impossible because $\frac{\partial^2 W}{\partial n_\theta^2} < 0$: if one n_θ were lower, then at least one other must be higher to not violate the constraint. But then the first of these would involve $\frac{\partial W}{\partial n_\theta} > 0$, and the second $\frac{\partial W}{\partial n_\theta} < 0$, which means they cannot be equalized.

To prove the second part (excessive messages), recall that $\frac{dW}{dn_\theta} = \frac{\partial W}{\partial n_\theta} + \frac{\partial W}{\partial N}$, and we have just argued that $\frac{\partial W(n^e,N^e)}{\partial n_\theta} = 0$ for all θ. Then we have that $\frac{dW(n^e,N^e)}{dn_\theta} = \frac{\partial W(n^e,N^e)}{\partial N}$. From (16), we have

$$\begin{aligned} \frac{\partial W(n,N)}{\partial N} &= \sum\nolimits_{\theta=1}^{\Theta} (b_\theta - c_\theta) q_\theta \frac{d\mathbb{Q}(n_\theta, N)}{dN} \\ &= -\sum\nolimits_{\theta=1}^{\Theta} (b_\theta - c_\theta) q_\theta \frac{n_\theta}{N} \mathbb{P}_\theta. \end{aligned} \tag{19}$$

Using the zero profit condition (3) we get

$$\frac{\partial W(n,N)}{\partial N} = -\frac{1}{N} \sum\nolimits_{\theta=1}^{\Theta} \gamma_\theta n_\theta. \tag{20}$$

This congestion externality is therefore strictly negative. ■

The first key feature here is the one that generates the optimality result. This is that the marginal change in the choice probability holding fixed the total number of messages, $\frac{\partial \mathbb{Q}(n_\theta,N)}{\partial n_\theta}$, which is instrumental in the social problem, is equal to $\mathbb{P}_\theta$, the probability 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 both reflects its social contribution and the private incentive for sending it. In neither case are we concerned about it crowding out other messages from the sector: in the private case, any other message takes precedence by dint of its lower price, and, in the social case, again only the extra likelihood of being examined counts.

The presentation above underscores the main problem with the market equilibrium: although the allocation is optimal across sectors given the total message volume, the *overall volume is excessive.* This is seen clearly from the condition we just argued, namely that $\frac{\partial W(n^e,N^e)}{\partial n_\theta} = 0$ (i.e., evaluated at the equilibrium N), while $\frac{dW(n^e,N^e)}{dn_\theta} < 0$. 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 means excessive messages are sent. The social cost of an extra message, as per (20), 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.

5.1 Taxing transmission (uniformly)

The results above suggest that taxing transmission will raise welfare. The next Proposition clarifies in which sense.

Proposition 13 *A percentage transmission tax increases 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 tax.*

Proof. A common percentage tax, τ, on transmission raises each γ_θ to $\gamma_\theta (1+\tau)$ and so reduces each π_θ proportionately to $\frac{\pi_\theta}{1+\tau}$. From (8), such a common tax means $N(1+\tau)$ is constant: equivalently, the original N falls to $\frac{N}{1+\tau}$. Recall $\frac{n_\theta}{N} = 1 - \left(\frac{1}{\phi}\frac{N}{\pi_\theta}\right)^{\frac{1}{\phi-1}}$ from (5). Since the ratio $\frac{N}{\pi_\theta}$ (on the RHS) is unaltered by the tax, 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 remain unaltered too. Consumer welfare therefore remains unchanged, profits

remain zero, and so welfare rises by the amount of tax raised. ■

The economics of raising transmission rates are the economics of rent dissipation. Doubling the cost in each sector simply halves the number of ads sent per sector. The intuition comes from the fact that both N and then n_θ are homogeneous of degree minus one. The sector choice probabilities (n_θ/N) are then homogeneous of degree zero in the percentage transmission tax. The advertised price distribution, $F(p,\theta)$ is then also independent of the tax rate.

5.1.1 Crowding out by higher qualities?

We just showed that percentage (across the board) cost increases have no effect locally on total transmission costs as borne by senders, due to adjustment to the zero profit equilibrium. 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 12 showed that the base allocation of messages was optimal for the equilibrium message volume, N^e. An equal percentage tax on transmission scales back messages proportionately. However, this does NOT mean that the scaled-back message levels induced by a non-negligible tax are also optimal for the new (given) total volume of messages. Indeed, the proof of Lemma 11 gives the partial welfare derivative (19)

$$\frac{\partial W(n,N)}{\partial n_\theta} = (b_\theta - c_\theta)\, q_\theta \mathbb{P}_\theta - \gamma_\theta,$$

and this expression still holds in the presence of a tax (although the arguments are proportionately smaller). These partial derivatives are still to be equalized across sectors at any optimal allocation for given N. However, the market equilibrium condition in the presence of a proportional tax on transmission becomes $(b_\theta - c_\theta)\, q_\theta \mathbb{P}_\theta = \gamma_\theta (1+\tau)$. Substituting,[33]

[33] Loosely, this is akin to an envelope result: here is the revenue raised on the last message in the sector when sector sizes are optimal.

$$\frac{\partial W(n,N)}{\partial n_\theta} = \tau\gamma_\theta. \tag{21}$$

This means that the allocation is constrained optimal (all the $\partial W(n,N)/\partial n_\theta = 0$) 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 (21), 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.

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. The results above instead suggest that the optimal tax should fall more heavily on the cheaper message communications: from (21), the sector-specific tax rate that ensures all sectors have the same marginal social benefit entails τ_θ inversely proportional to γ_θ.[34]

5.1.2 Crowding out by lower qualities?

Proposition 12 suggests that the low surplus sectors do not inflict more damage on the high surplus ones, or vice versa, at the equilibrium. All sectors are in excess, but no group should be singled out. However, the discussion at the end of the last subsection indicates that if they were all scaled back to half size (by the expedient of a 100% tax on transmission say), then the high surplus sectors would be overrepresented.[35] This suggests that the higher surplus alternatives might be over-represented in the population of messages (in the sense that they ought to be scaled back more than proportionately).

This result leads us to ask whether an 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

[34] Indeed, the first-best optimum entails just one message per sector, which is also not consistent with proportional scaling back through taxes (unless all sectors are symmetric).

[35] If half the messages were discarded across the board, then the marginal benefit of a message in a sector would be (retaining the original number values as arguments): $\frac{\partial W(n,N/2)}{\partial n_\theta} = (b_\theta - c_\theta)\, q_\theta \frac{2\phi}{N}\left(1 - \frac{n_\theta}{N}\right)^{\phi-1} - \gamma_\theta$. Since $(b_\theta - c_\theta)\, q_\theta \frac{\phi}{N}\left(1 - \frac{n_\theta}{N}\right)^{\phi-1} = \gamma_\theta$ at the original equilibrium, then $\frac{\partial W(n,N/2)}{\partial n_\theta} = \gamma_\theta$. This also indicates a benefit from scaling back the low-cost end relatively more.

help if all sectors are roughly similar, but it can if they are sufficiently asymmetric and a low-surplus sector gets worse (or even becomes inviable).

From (16), the relevant welfare derivative is

$$\begin{aligned}\frac{dW}{d\gamma_\theta} &= -n_\theta + \sum\nolimits_{\theta'=1}^{\Theta} \frac{\partial W}{\partial n_{\theta'}} \frac{dn_{\theta'}}{d\gamma_\theta} + \frac{\partial W}{\partial N} \frac{dN}{d\gamma_\theta} \\ &= -n_\theta + \frac{\partial W}{\partial N} \frac{dN}{d\gamma_\theta}\end{aligned}$$

since each $\frac{\partial W}{\partial n_{\theta'}} = 0$ at equilibrium. This expression indicates that there is a trade-off. From (20), $\frac{\partial W}{\partial N} = -\frac{1}{N} \sum\nolimits_{\theta'=1}^{\Theta} n_{\theta'} \gamma_{\theta'}$. The other desired term is $\frac{dN}{d\gamma_\theta}$: from (8)) we have $\frac{dN}{d\gamma_\theta} = -N \frac{1}{\gamma_\theta} \frac{1}{\Theta} \frac{\chi_\theta}{\bar{\chi}}$, 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:

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

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

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

$$\begin{aligned}\frac{dW}{d\gamma_\theta} &\stackrel{s}{=} -\frac{n_\theta \gamma_\theta \Theta}{\sum\nolimits_{\theta'=1}^{\Theta} n_{\theta'} \gamma_{\theta'}} + \frac{\chi_\theta}{\bar{\chi}} \\ &= -\frac{\Gamma_\theta}{\bar{\Gamma}} + \frac{\chi_\theta}{\bar{\chi}},\end{aligned}$$

where the s label denotes that the derivative has the sign of the expression, and where $\Gamma_\theta = n_\theta \gamma_\theta$ is the aggregate transmission cost for sector θ, and $\bar{\Gamma}$ is the average of these. A marginal sector has Γ_θ close to zero because it delivers few messages: if the sector has higher costs per message than average, then its π is low, so its χ is high. Therefore, for a marginal sector, the first term is close to zero and the second term dominates: a weak sector's rise in costs (which effectively can bring about its demise) is socially beneficial. The two sets of results above are summarized as:

Proposition 14 *Under a positive uniform percentage transmission tax, the market allocation is biased*

against sectors with low transmission costs. However, welfare can rise when transmission costs increase in weak sectors.

The sense above indicates the weak products as those which are socially most harmful. This holds even though they have a small foothold: one might have otherwise suspected the high-profit products because they are responsible for the most crowding. The result of the previous subsection, which instead pointed the finger at the low-cost products as being over-represented, was derived under the assumption that all messages were scaled back proportionately (equivalently, the tax was a proportional one). Here it is the cost of a message that is being increased without a corresponding tax receipt. This underscores the reasons for the differing results: taxing a high-profit sector yields a lot of revenue except if the proceeds are lost.

6 Distractions

Several of the strong properties in the normative analysis above relied critically on the homogeneity property of the numbers of messages sent. One natural way to relax this property is to introduce another source of competition for attention.

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 an email from a Dean or a crying child. We model this outside competition for attention as further "distractions" to attention. Formally, this means there are n_0 other messages (or activities) which compete for attention along with the messages sent from the advertising sectors. Hence now we have $N = \sum_{\theta=0}^{\Theta} n_\theta$. We will further associate an exogenous social value π_0 to each message (or activity) examined from the outside sector, and we assume that this value accrues on EACH such message examined.

This amendment relaxes some of the stronger properties of the equilibrium configuration, but retains other key ones, most notably that the market equilibrium is still optimal given N^e. However, raising τ in all sectors (except the distraction sector, which in that sense can be viewed as an untaxed sector) no longer causes the price spread to remain unchanged for all sectors. The key property before was that $N/\pi(1+\tau)$ was independent of τ. Now that is no longer true, and some sectors get evicted as tax rates rise.

6.1 Message volume with distractions

With distractions, it is still true that each sector's message share is $\frac{n_\theta}{N} = 1 - \left(\frac{N}{\phi}\frac{1}{\pi_\theta}\right)^{\frac{1}{\phi-1}}$, $\theta = 1, ..., \Theta$ (see (1)). 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:

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

Writing this out, we have

$$n_0 + N(\Theta - 1) = N^{\frac{\phi}{\phi-1}} \sum\nolimits_{\theta=1}^{\Theta} \left(\frac{1}{\phi}\frac{1}{\pi_\theta}\right)^{\frac{1}{\phi-1}}. \tag{22}$$

The LHS is linear in N (with a positive intercept), and the RHS is convex (and starts at 0), so that there is one and only one intersection with $N > 0$. Hence there is a unique solution $N > 0$. The comparative static properties of the equilibrium are quite simple. For example, a higher value of n_0 leads to a lower N, and n_θ falls in all other sectors.

6.2 Welfare analysis

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

$$W = \sum\nolimits_{\theta=1}^{\Theta} \left[(b_\theta - c_\theta) q_\theta \mathbb{Q}_\theta - \gamma_\theta n(\theta)\right] + n_0 \pi_0 \frac{\phi}{N},$$

where π_0 denotes the net social benefit per distraction, and n_0 distractions vie for the attention span of ϕ given N total competitors. The result of Proposition 12 with $n_0 > 0$ still holds true: with a distraction, the equilibrium allocation is still constrained optimal.

The proof follows the lines of the earlier one: again, for the active message-sending sectors, all the marginal benefits are equalized, 1...Θ. 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 constrained optimal problem still has the "right" allocation of the messages across the Θ sectors.

Now consider a uniform percentage tax on all sectors, except the "untaxed" sector, n_0. From the welfare function above, the effect of a tax is $\frac{dW}{d\tau} = \sum_{\theta=1}^{\Theta} \frac{\partial W}{\partial n_\theta}\frac{dn_\theta}{d\tau} + \frac{\partial W}{\partial N}\frac{dN}{d\tau}$.[36] Evaluating at $\tau = 0$ yields again the result that the equilibrium entails the optimal allocation, $\frac{\partial W}{\partial n_\theta} = 0$, 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 $\frac{\partial W}{\partial N} < 0$ as argued before, since each $\mathbb{Q}_\theta$ term is decreasing in N (it becomes harder to find a sector) and the additional term, $n_0\pi_0\frac{\phi}{N}$, is decreasing in N as long as $\pi_0 > 0$. Since $\frac{dN}{d\tau} < 0$, welfare increases from a uniform percentage tax.

7 Further Extensions

7.1 Elastic demand

So far we have supposed that demand is rectangular. We now argue that the positive analysis remains tractable when we replace this assumption by a downward-sloping demand curve. We can still fully determine the shape of the equilibrium price distribution when there are many sectors, each with a specific conditional demand function. The key property is that we can still back out the number of messages in the sector from the calculus for the highest-priced firm, and thence determine the entire price distribution, Details are below. Surprisingly, we also retain the key property that a percentage tax on message transmission does not change the distribution, but simply scales back the number of messages proportionately.

Suppose then that sector θ is associated to a conditional demand, $q_\theta(p)$, with the understanding that the consumer will buy this number of units at the lowest price, p, held. Assume that demand begets a quasi-concave profit function with a maximizing price $\hat{p}_\theta$. The corresponding conditional (on being the only one found from the sector) profit 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 did π_θ in the earlier analysis with rectangular demand. In equilibrium, no firm will charge more than $\hat{p}_\theta$ because profits can be increased by charging $\hat{p}_\theta$.

The parallel analysis to that above yields the equilibrium price distribution as

[36] To find $N(\tau)$, and hence $n_\theta(\tau)$, use the previous expressions and note that $n_0 + N(\Theta - 1) = N^{\frac{\phi}{\phi-1}}(1+\tau)^{\frac{1}{\phi-1}} \sum_{\theta=1}^{\Theta} \left(\left(\frac{1}{\phi}\frac{1}{\pi_\theta} \right)^{\frac{1}{\phi-1}} \right)$, so that total messages go down with τ.

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

(cf. (13)), where $N=\phi\frac{(\Theta-1)^{\phi-1}}{\left(\sum_{\theta=1}^{\Theta}\left(\frac{1}{\hat{\pi}_\theta}\right)^{\frac{1}{\phi-1}}\right)^{\phi-1}}$ (which is the same expression as (8) except with $\hat{\pi}_\theta$ replacing π_θ). Now $\underline{p_\theta}$ is given implicitly by (cf. (14))

$$\left(\underline{p_\theta}-c_\theta\right)q_\theta\left(\underline{p_\theta}\right)=\left(\frac{N}{\phi}\right)\gamma_\theta,$$

which has a unique price solution for $\underline{p_\theta}<\hat{p}_\theta$ under the assumption that profit is strictly quasi-concave.[37] Compared to the earlier distribution for rectangular demand, if we set $\hat{p}_\theta=b_\theta$, the distribution is now stochastically lower (FOSD) because lower prices are relatively more attractive than before because of the demand expansion effect.

It is clear from the price distribution given above that it is independent of a percentage tax on message transmission costs if N is proportional to these costs. But this is true by the linear homogeneity property of (8).

Proposition 13 addressed the optimal allocation of a given number of messages. In the earlier context, price levels within a sector are irrelevant for total surplus (though not for surplus distribution). Now price levels matter. We could, of course, assume first-best pricing at marginal cost, but this would scarcely reflect the market situation. We can scarcely assume the equilibrium price distribution is given and then vary message numbers by sector since this would violate the assumed equilibrium zero-profit condition. This means that we cannot meaningfully perform a similar exercise to the earlier one.

7.2 Endogenous Examination

So far we have treated the examination decision as effectively exogenous. The consumer is completely passive and just processes a fixed window of the information that flows past. This can be thought of as the pure Couch Potato model. One way to treat a more active consumer is to assume there is a cost to examination, say $C(\phi)$, and the consumer chooses how much attention to pay.

[37] If the profit function is not quasiconcave, the support of the price distribution will have a gap for any price such that profit is no lower at a lower price.

For any given level of attention, the equilibrium on the firm side looks as in the analysis above, and the level of consumer welfare is simply given by the welfare function (16) (since profits are zero there), from which $C(\phi)$ is to be subtracted. However, it would not be reasonable to assume that the consumer internalizes the zero profit condition, so the consumer should not be modeled as maximizing such a function. Instead, the consumer should be modeled as taking as given the number of messages sent.[38] The analysis of the model above constitutes the special case where $C(\phi)$ is zero up to a "capacity" level.[39]

8 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. This means that likely the attention span of consumers is larger than ever before. Yet modern economics also involve many product classes, and the volume of information is large. These factors interact to determine the degree of competitiveness of sectors, which is reflected in the degree of price dispersion. These are the factors that are brought together in the model, and 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 indeed, 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 their absolute message volume because crowding relaxes price competition leading to stochastically higher prices. This can encourage 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.

[38] The consumer then maximizes her welfare given the price distribution predicated upon an expected attention span (see (13)), and the equilibrium is a fixed point where the desired and anticipated attention spans coincide.

[39] The simplest case to analyze is when sectors are symmetric. Then, for given N, symmetrically divided among sectors, the consumer's optimal ϕ can be determined, and intersected with the relation in Figure 2 (which determines the senders' actions as a function of a given ϕ).

Third, as consumer attention rises through new outlets reaching consumers, 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 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 level.

The intuition for Butters' optimality result is this. As we have seen, the highest-price firm's calculus drives the equilibrium number of messages. It only gets a sale if it is the only message in a mailbox, so the zero-profit condition applied to it means that the social benefit times the probability of being alone must equal the cost of sending the message. The optimal calculus is that the net social benefit of a further message should be zero. A message nets $b - c$ with probability that it is the only one in the box, and costs γ, just like a private message. This implies that the equilibrium decentralizes the optimum allocation.

The intuition for our optimal allocation of ads across sectors (given the total 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 of an extra ad there 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 this is zero for all sectors.

The model delivers a detailed picture of equilibrium price distributions across asymmetric sectors com-

peting for attention. The equilibrium total volume of advertising messages is shown to be a CES function of the individual sectors' profitability measures. This constitutes a novel derivation of such a CES function, and is instrumental in being able to derive 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. The converse case is 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 the 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. 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, though this remains a crucial marketing dimension. Likewise, messages are assumed to be sampled randomly, so there is no allowance for the consumer to pay more attention to particular message types. The Economics of Attention has yet to be fully fleshed out in these broader directions.

9 Appendix

9.1 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 $1-\left(1-\frac{1}{M}\right)^N$. When M is large, this is approximately $1-\exp\left(-N/M\right)$ $(=\Lambda)$. Hence, $\frac{d\Lambda}{dN} = \frac{1}{M}\exp\left(-N/M\right) = \frac{1}{M}\left(1-\Lambda\right)$, from which it follows that the number of uninformed at the optimum is $\frac{M\gamma}{b-c}$, the same as in equilibrium.[40]

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 $\exp\left(-A\left(p\right)/M\right)$ which must equal $\frac{\gamma}{p-c}$ by the zero profit condition. The form of $A\left(p\right)$ and its properties (decreasing, concave) follow directly.

[40] The interpretation is that the business stealing and consumer surplus appropriation externalities net out.

9.2 Sales price distribution

The advertised price distribution is $F(p,\theta)$ as given in Proposition 8; denote the corresponding density by $f(p,\theta)$. Then the sales price density at p is the advertised price density, $f(p,\theta)$, times the probability, $\mathbb{P}(p,\theta)$, that the advertised price makes a sale (up to a multiplicative constant, k_1). Since the latter is $\frac{\gamma_\theta}{p-c_\theta}$, this means that the sales price density is:

$$g(p,\theta) = k_1 f(p,\theta)\,\mathbb{P}(p,\theta) = k_2\left(\frac{1}{p-c_\theta}\right)^{\frac{1}{\phi-1}+2},$$

where k_2 is a constant to be determined. The corresponding cumulative distribution of the sales price (cf. (13)) is

$$G(p,\theta) = k_3\left[\left(\frac{1}{\underline{p}-c_\theta}\right)^{\frac{1}{\phi-1}+1} - \left(\frac{1}{p-c_\theta}\right)^{\frac{1}{\phi-1}+1}\right].$$

The constant k_3 can be determined from the relation $G(b_\theta,\theta) = 1$. Differentiating gives the following properties (cf. Proposition 8).

The equilibrium sales price density in sector θ is decreasing and convex on $[\underline{p_\theta}, b_\theta]$, with cumulative distribution given by

$$G(p,\theta) = \frac{\left(\frac{1}{\underline{p_\theta}-c_\theta}\right)^{\frac{1}{\phi-1}+1} - \left(\frac{1}{p-c_\theta}\right)^{\frac{1}{\phi-1}+1}}{\left(\frac{1}{\underline{p_\theta}-c_\theta}\right)^{\frac{1}{\phi-1}+1} - \left(\frac{1}{b_\theta-c_\theta}\right)^{\frac{1}{\phi-1}+1}}, \tag{23}$$

where $\underline{p_\theta}$ is given by (14).

This is the distribution of the actual transaction prices for sector θ, i.e., conditional on a sale being made. Since the probability of a sale being made in sector θ is $\mathbb{Q}_\theta = 1-\left(1-\frac{n_\theta}{N}\right)^\phi$, then $\mathbb{Q}_\theta G(p,\theta)$ represents the (unconditional) probability of a sale being made (or, indeed, the fraction of consumers buying) in sector θ at a price below p. This statistic allows us to calculate the expected consumer surplus from the sector.[41]

[41] We can use $G(p,\theta)$ to calculate the *size distribution of firms* within sector θ. In particular, we can simply replace $p = c_\theta + \frac{\gamma_\theta}{Q}$ (from the zero profit condition) where $Q = \mathbb{Q}q_\theta$ is the size of the firm, in terms of units sold. Substituting in (23) gives:

$$H(Q,\theta) = 1 - G\left(c_\theta + \frac{\gamma_\theta}{Q},\theta\right) = \left[\frac{\left(\frac{Q}{\gamma_\theta}\right)^{\frac{1}{\phi-1}+1} - \left(\frac{Q_L}{\gamma_\theta}\right)^{\frac{1}{\phi-1}+1}}{\left(\frac{Q_H}{\gamma_\theta}\right)^{\frac{1}{\phi-1}+1} - \left(\frac{Q_L}{\gamma_\theta}\right)^{\frac{1}{\phi-1}+1}}\right],$$

where Q_H $(=\frac{\gamma_\theta}{\underline{p_\theta}-c_\theta})$ is the highest output (corresponding to the lowest price, $\underline{p_\theta}$) and Q_L $(=\frac{\gamma_\theta}{b_\theta-c_\theta})$ is the lowest output (corresponding to the highest price, b_θ). The corresponding density, $h(Q,\theta)$, is proportional to $Q^{\frac{1}{\phi-1}}$ and so is increasing, and concave for $\phi > 2$. Note that aggregate congestion interactions DO enter here (and are simply described through the N effect), since $\underline{p_\theta} = c_\theta + \left(\frac{N}{\phi}\right)\frac{\gamma_\theta}{q_\theta}$ enters into Q_H $(=\frac{\gamma_\theta}{\underline{p_\theta}-c_\theta})$.

The next Figure shows the difference: the advertised price distribution, $F(p,\theta)$ is given as the solid line. (The parameter values used are: $b = 1$, $c = 0$, $N/\phi = 10$, $\phi = 10$ and $\gamma = 0.025$.)[42] The distribution of actual transactions conditional on a sale being made, $G(p,\theta)$, is the dotted line on top. The dashes represent the cumulative distribution of actual sales in the population, $\mathbb{Q}_\theta G(p,\theta)$, for sector θ. This sales price distribution lies above the advertised price distribution for low p. This means simply that more sales are made at low prices (as high priced ads are beaten). This must happen throughout the whole range of prices in Butters' model because each ad is received by someone. However, in our context, there is a probability that ads are not received at all. This feature is reflected in the fact that $\mathbb{Q}_\theta G(b) < 1$ in our model (see Figure 3): the consumer may get no ads from the sector and no sale is made at all.

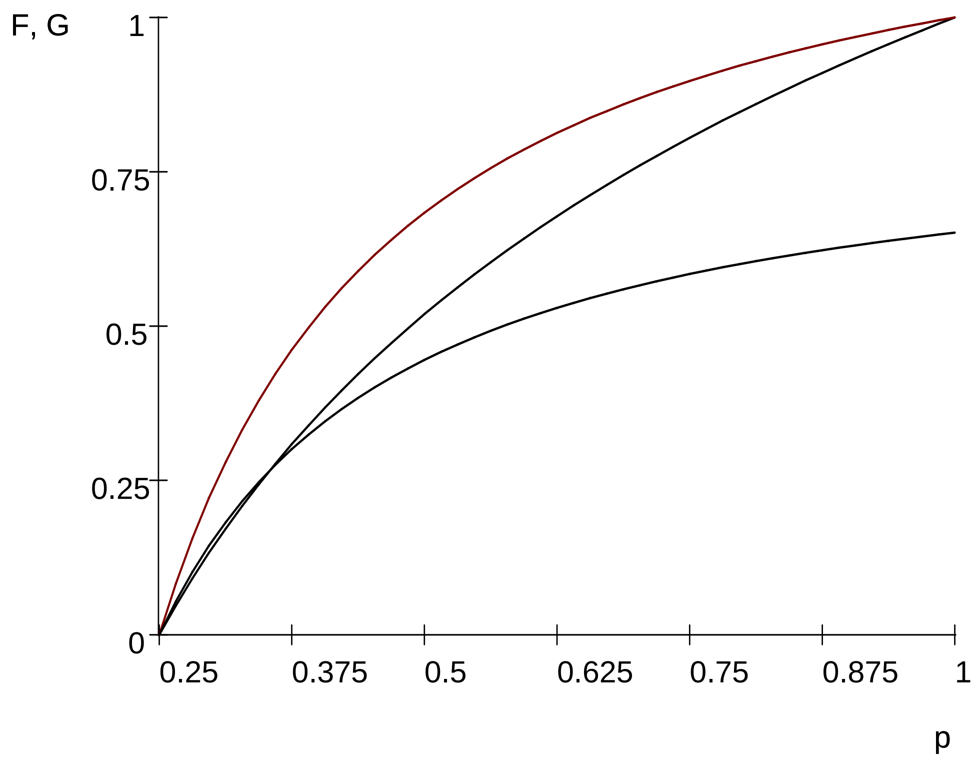

Figure 3. Advertised price functions (solid) and sales price distributions.

[42] The value $N/\phi = 10$ can be consistent with a specific number of other sectors.

References

[1] Anderson, Simon P. and Stephen C. Coate (2005) Market Provision of Broadcasting: A Welfare Analysis. *Review of Economic Studies*, 72, 947-972.

[2] Anderson, Simon P. and Andre de Palma (2007) Information Congestion: open access in a two-sided market. http://www.virginia.edu/economics/downablepapers.htm#anderson

[3] Anderson, Simon P., Andre de Palma and jacques-Francois Thisse (1992) *Discrete Choice Theory of Product Differentiation*. MIT Press.

[4] Anderson, Simon P. and Regis Renault (1999) Pricing, Product Diversity and Search Costs: a Bertrand-Chamberlin-Diamond Model. *RAND Journal of Economics*, 30, 719-735.

[5] Ayres, Ian and Barry Nalebuff (2003) Want to call me? Pay me! *Wall Street Journal,* October 8th, 2003.

[6] 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.

[7] Baye, Michael R. and John Morgan (2001) Information Gatekeepers and the Competitiveness of Homogeneous Product Markets, *American Economic Review*, 91(3), 454-474.

[8] Beard, T. Randolph and Avery M. Abernethy (2005) Consumer Prices and the Federal Trade Commission's "Do-Not-Call" Program. *Journal of Public Policy and Marketing*, 24, 253-259.

[9] Brown, Tom J., and Michael L. Rothschild (1993) Reassessing the Impact of Television Advertising Clutter, *Journal of Consumer Research*, 20(1), 138-146.

[10] Burdett, Kenneth and Kenneth L. judd (1983) Equilibrium Price Dispersion, *Econometrica*, 51(4), 955-969.

[11] Butters, Gerard R. (1977) Equilibrium Distributions of Sales and Advertising Prices. *Review of Economic Studies*, 44, 465-491.

[12] Debreu, Gerard (1960) Review of R. D. Luce, Individual Choice Behavior: A Theoretical Analysis. *American Economic Review*, 50, 186-188.

[13] Diamond, Peter A. (1971) A Model of Price Adjustment, *Journal of Economic Theory*, 3, 156-168.

[14] Dixit, Avinash K. and Victor D. Norman (1978) Advertising and Welfare. *Bell Journal of Economics.* 9, 1-17.

[15] 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.

[16] Grossman, Gene M. and Shapiro, Carl (1984) Informative Advertising and Differentiated Products. *Review of Economic Studies*, 51, 63-81.

[17] Lanham, Richard A. (2006) *The Economics of Attention: Style and Substance in the Age of Information.* University of Chicago Press.

[18] Luce, R. Duncan (1959) *Individual Choice Behavior: A Theoretical Analysis.* New York, Wiley.

[19] Miller, George A. (1956) The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review,* 63, 81-97.

[20] Schroeder, Harold M., Michael J. Driver, and Siegfried Streufert (1967) *Human Information Processing: Individual and Groups Functioning in Complex Social Situations.* Holt, Rinehart, and Winston, New York.

[21] Shirman, D. R. (1996) When E-mail becomes Junk Mail: the Welfare Implications of the Advancement of Communications Technology. *Review of Industrial Organization,* 11, 35-48.

[22] 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.

[23] TRAIN, KENNETH (2003) *Discrete Choice Methods with Simulation*, Cambridge University Press.

[24] VAN ALSTYNE, MARSHALL, THEDE LODE, AND RICK WASH (2004) Information asymmetry and thwarting spam, *Center for eBusiness, Research Brief III(3),* 1-4, MIT Sloan School.

[25] VAN ZANDT, TIMOTHY (2004) Information Overload in a Network of Targeted Communication, *RAND Journal of Economics*, 35(3), 542-560.

[26] WEBB, PETER H., AND MICHAEL L. RAY, (1979) Effects of TV Clutter, *Journal of Advertising Research*, 19(3), 7-12.

[27] WILLMORE, C. (1999) A Penny for your Thoughts: E-mail and the Under-Valuation of Expert Time. *Working Paper,* University of Warwick.