Information Congestion

Simon P. Anderson and André de Palma*

June 2003; this version June 2006

Abstract

Advertising messages compete for scarce attention. "Junk" mail, "spam" e-mail, and telemarketing calls need both parties to exert effort to generate transactions. Message recipients supply attention depending on average message benefit. Senders are motivated by profits. Costlier message transmission may improve message quality so more messages are examined. Too many messages may be sent, or the wrong ones. A Do-Not-Call policy beats a ban, but too many individuals opt out. A monopoly gatekeeper performs better than personal access pricing if nuisance costs are moderate. "The medium is the message" with multiple channels, and there is excessive indiscriminate mailing.

JEL Classification: D11, D60, L13.

Keywords: information overload, congestion, advertising, common property resource, overfishing, two-sided markets, junk mail, email, telemarketing, Do Not Call List, message pricing, the Medium is the Message, market research.

---

*Simon P. Anderson: Department of Economics, University of Virginia, PO Box 400182, Charlottesville VA 22904-4128, USA, sa9w@virginia.edu. André de Palma: Senior Member of the Institut Universitaire de France, ThEMA, Université de Cergy-Pontoise, 33 Bd. du Port, 95011, Cergy Cedex, France. andre.depalma@u-cergy.fr. The first author gratefully acknowledges research funding from the NSF under grant SES-0137001 and from the Bankard Fund at the University of Virginia. We thank Bruno Jullien, Joshua Gans, and Rabah Amir for comments, seminar participants at various universities, and conference participants at EARIE 2004 (Berlin), the 2004 Chinese IO Society meeting (Beijing), the 2005 Internet conference (IDEI, Toulouse), the 2006 Summer Workshop in Industrial Organization (Auckland), and the 2006 Economics of ICT Workshop (ENST, Paris).

1 Introduction

Expenditure on advertising in the US amounted to some \$245 bn. (in 2003); which constitutes around 2.25% of GDP.[1] Of this, around \$49 bn. (19% of the total) was spent on direct mailing. A further \$47 bn. was spent on telephone marketing.[2] This spending is just the tip of a much larger economic activity that is facilitated by this marketing. The Direct Marketing Agency (admittedly not an impartial observer) estimates that Direct Marketing activity drives 10.3% of GDP. However, nearly half (46%) of the 4m tons of bulk mail delivered goes unopened to the landfill (or recycler). If this unread component could be more efficiently harnessed to generate even a fraction of the revenue the read component generates, then the Information Age junk mail congestion problem alone could be larger than the more traditional sector road congestion problem (which is around 1% of GDP).

It is not just bulk mail and telemarketing that suffer from congestion. Spam email is a curse on a new communication technology because a spammer can send 650,000 messages in an hour, at virtually no cost: spam filters cause people to lose important messages, or even valid commercial offers that they might have taken up had they not been lost in a morass of other propositions.[3] Advertisements in general often do not register their message with the prospective customer. Estimates of the number of advertising messages seen per day vary from 250 to 5000.[4] Nielsen Media Research reports a telephone survey (from 2000) in which they rang up households before 10 p.m. and then called back after 10 the same evening.[5] Under 15% of respondents could cite an ad from the last ad break in the program they were watching at the time of the call, and very few cited more than one, even after such a short delay.

The economics of such unsolicited advertising is characterized by a clutter of messages and the subse-

[1] These statistics are from Nielsen data at www.tvb.org and the Direct Marketing Association at www.the-dma.org .

[2] The other main categories for advertising are 25% on TV, 18% in newspapers, 8% on radio, and 6% in yellow pages. Note that telephone marketing is included in Direct Marketing figures (together with Direct Mail, it makes up 60% of the total), but not in Advertising figures.

[3] The term spam comes from an early anecdote in the annals of computer geekdom. Someone sent his friends a message which contained just the word "spam" (after the Monty Python Flying Circus song) repeated hundreds of times: www.templetons.com/brad/spamterm.html describes the "origin of the term spam to mean net abuse". Spam can be around 55% of email, or even rise to 80%: see www.obviously.com/junkmail/ and cobb.com/spam/numbers.html .

[4] See http://answers.google.com/answers/threadview?id=56750 for discussion: the Consumer Reports Website estimates 247 messages a day.

[5] Details are available at the CAB website: www.cabletvad-bureau.com

quent congestion of the consumer's limited attention span.[6] In response, the consumer rations attention by screening out information – and good goes out with the bad, like a spam filter that blocks out some worthwhile messages. This view indicates two externalities at work. Senders of messages trying to get their messages through the clutter do not account for crowding out other senders: the consumer's attention can be construed as a common property resource. Receivers of messages mentally screening out clutter do not account for lost sender benefits.

Various policy measures and institutions aim to address these problems. Telemarketing (junk telephone calling) in the US has seen a dramatic decline since the recent advent of the FTC-sponsored Do-Not-Call list. Ayres and Nalebuff (2003) suggested that receivers could set their own personal access prices. Bill Gates has suggested an email tax might help for spam: "At perhaps a penny or less per item, e-mail postage wouldn't significantly dent the pocketbooks of people who send only a few messages a day. Not so for spammers who mail millions at a time." Van Alstyne et al. (2004) propose a system whereby the sender must post a bond that can only be recouped if the receiver likes the message content. However, regarding bulk mail, the lowest rate charged by the US post office for bulk mailing is 13.1 cents per item (up from 8.8 cents in 2005), which is way below the current rate of 39 cents for first class mail.[7] When the message medium is owned by a profit maximizing entity, consumer attention can be enhanced by improving the attractiveness of the medium, such as screening interesting movies on TV.

To model the interaction, we consider two groups of economic agents, senders and receivers of messages. For concreteness here, think of them as firms and consumers in a two-sided market. Firms need to communicate their wares. They do so by sending messages (bulk mail, etc.) to prospective consumers on the other side of the market. Sending messages is costly and both sides need to exert effort to arrive at transactions: the firm must send a message and the consumer must "examine" it (answer the 'phone, say). In this market interaction, the number of messages sent depends on expected profit of the marginal sender, which in turn

[6] Unsolicited advertising includes billboards and radio/television. Classified ads and ads in specialist magazines may be more sought after. The distinguishing feature we consider is the crowding of attention. This feature applies to billboards and TV ads too.

[7] This lowest rate applies to non-profit organizations. The rate is up to 10.3 cents higher for private firms. For USPS rate information, see http://pe.usps.com/text/dmm300/ratesandfees.htm

depends on the number of messages read by the receiver. However, the number of messages the receiver examines depends on the average quality that the receiver expects. A higher cost to sending messages may increase the number of messages examined because the expected quality rises.

It has been recognized for decades that excess information is costly: the term Information Overload was coined by Toffler (1970), although the concept was recognized earlier. Miller (1956) presents evidence of an "inverse N" relation between information received and decision accuracy, which was later elaborated upon by Schroeder, Driver, and Streufert (1967). Eppler and Mengis (2004) review the literature on Information Overload from Organizational Science, Accounting, Marketing, and MIS. Interestingly, while they note that there are also contributions in Health Care, Psychology, and Mass Communication, there is curiously little work in economics.[8] A striking exception is Van Zandt (2004), who considers a model similar to ours (except that his receivers examine a fixed number of messages) but with a different emphasis: he is interested in targeted recipients of messages. His receivers have different worth to different senders, and therein lies the efficiency benefit in his model from a tax on messages. A small tax may be Pareto improving for all firms because such a tax will cause marginal firms to refrain from sending messages to those consumers unlikely to be much interested. Those firms will gain from becoming more prominent with consumers from whom they expect larger profits. This matching aspect does not arise in our main model.

The outline of the paper is as follows. The next section describes the behavior of the agents on the two sides of the potential transaction and derives the building blocks of the equilibrium analysis, namely the sender transmission function and the receiver examination function. Section 3 puts these together and compares equilibrium to the optimum. Section 4 looks at the crucial role of the receiver surplus in the equilibrium solution. Section 5 allows for intrinsic nuisance costs and looks at the possibility of receivers opting out completely (for example, the federal Do Not Call list in the US), the pros and cons of outright bans, and the ability of personal access pricing to solve the problem. Section 6 addresses the solution chosen by a monopoly information gatekeeper, relates the current work to the analysis of two-sided markets, and compares with personal pricing. Section 7 discusses competition among senders in the product market.

[8] Although Shirman (1996) and Willmore (1999) note that e-mail is excessive because it creates negative externalities.

Section 8 underscores the benefits of multiple media, Section 9 considers the insufficiency of equilibrium targeting, and Section 10 concludes.

2 Congestible information

For a (mutually beneficial) transaction to be consummated, information must be transmitted by a sender, and the receiver must both process it and react positively (by purchasing an advertised good, say, or joining a club). Only after these costly efforts from both participants can a successful transaction occur. We suppose that there is a single receiver and many senders. The surpluses to each party (conditional on a message being processed) depend only on the sender type: there is no "business stealing," so that messages only compete with each other insofar as they compete for the receiver's attention in examining them. Senders' expected profits differ because they have different mark-ups, and/or the probability that the receiver is interested in buying a product may differ across products. The expected receiver surplus from any particular message can reflect the probability of buying (being interested in an advertised product), as well as different conditional surpluses from buying.

Senders decide whether or not to send a message. The receiver chooses an attention span which is how many of the messages received to examine. For bulk mail, households decide how many letters to open. For telemarketing calls, they decide how often to answer the telephone. The receiver's decision considers the expected surplus from a message. Congestion arises when the receiver chooses not to examine all the messages received.

Equilibrium is described as the intersection of two curves that represent the behavior of the two sides of the market. Further technical details of these curves are given in the discussion paper version.[9]

2.1 Information Senders

Senders are indexed by $\theta \in [0,1]$ such that expected profit conditional on a message being examined, $\pi(\theta) > 0$, is increasing in θ and is independent of how many messages are examined. We approximate the corresponding step function with a continuous function $\pi(.)$, assumed to be strictly increasing. The

[9] Available online at www.virginia.edu/economics/papers/anderson/info060514.pdf

distribution of sender types is denoted by $F(\theta)$, with associated density $f(.) > 0$.[10]

The mass of senders is M. There is a cost, $\gamma \in (\pi(0), \pi(1))$, for sending a message.[11] For the welfare analysis, we treat this as equal to social cost; and we also treat $\pi(\theta)$ as the gross social benefit on the producer side.[12] We assume that at most one message can be sent by each sender to the receiver.[13] The relation between the number of messages examined, ϕ, and the number this induces to be sent, $N(\phi)$, is the Sender Transmission Function (STF).

If ϕ is high enough (if receivers were prepared to examine many messages, and at least as many as are sent) there will be no congestion. The marginal sender type then attains its least value, $\theta^{\min}$, and the number of senders attains its largest value, $n^{\max}$. This marginal type is determined from the break-even condition as $\pi\left(\theta^{\min}\right) = \gamma$; since all higher sender types transmit, the corresponding number of senders is $n^{\max} = M\left[1 - F\left(\theta^{\min}\right)\right]$. This gives the vertical segment in Figure 1.

INSERT FIGURE 1. The Sender Transmission Function $N(\phi)$.

If ϕ is low ($\phi < n^{\max}$), there is congestion in the messages sent. The likelihood that any given sent message is examined is $\frac{\phi}{n}$. The profitability of sending a message is $\pi(\theta)$ weighted by this examination probability, and so the marginal sender type is uniquely defined by

$$\pi(\theta^*)\frac{\phi}{n} = \gamma, \tag{1}$$

with $n = M[1 - F(\theta^*)]$. This relation generates the curve in Figure 1 in the congested region ($\phi < n$). The chord from the origin to the curve (the ratio ϕ/n) is rising along the curve: only a rise in the examination probability can induce more messages to be sent. This implies that the STF is convex and slopes up.[14] A

[10] We could, without loss of generality, set $f(\theta) = 1$ and have the $\pi(\theta)$ function pick up relative clustering of types (alternatively, we could set $\pi(\theta)$ linear and have $f(\theta)$ do the work). It is useful for exposition to retain both $f(\theta)$ and $\pi(\theta)$.

[11] This implies that there are always some active senders and some senders "waiting in the wings" in equilibrium.

[12] Think, for example, of purely informative advertising that tells prospective consumers of the existence of products. The welfare economics of "persuasive" advertising is more contentious: see the discussion in Bagwell (2006) and in issues of the RAND journal following Dixit and Norman (1978).

[13] Anderson and de Palma (2006) address multiple messages sent to the same receiver. In terms of the notation below, all senders will transmit only one message if a second message is not profitable for the highest profit type, $\theta = 1$, i.e., $\pi(1) < \pi(\theta_1)/[1-\gamma/\pi(\theta_1)]$ where θ_1 solves $\pi(\theta_1)\phi/M[1-\theta_1] = \gamma$. Loosely, each sender transmits only one message as long as senders are quite homogenous and the transmission cost is close to the average profit level.

[14] In summary, $N(\phi) = n^{\max}$ for $\phi \geq n^{\max} = M\left(1 - F^{-1}\left(\pi^{-1}(\gamma)\right)\right)$ and for $\phi < n^{\max}$ its inverse function is given from (1) as $\phi = \frac{n\gamma}{\pi\left(F^{-1}\left(1-\frac{n}{M}\right)\right)}$.

higher transmission rate, γ, means that fewer messages are sent for any given ϕ, so that the STF moves to the left.

2.2 Receiver Attention Span

The receiver's relation between the number of messages sent and the number examined is the Receiver Examination Function (or REF), denoted $\Phi(n)$. Assume the receiver has a strictly convex (and twice differentiable) examination cost $C(\phi)$. The decision of how many messages to examine, which we call the receiver's *attention span*, is determined from the equality of marginal cost with the marginal benefit from examining a further message. Given the receiver cannot tell a priori which messages contain which offers, she examines them at random.[15] Since we assume message independence (how and whether she responds to any message does not depend on which other messages are received), her marginal benefit is just the *average* expected surplus over the configuration of messages received, $s_{av} = \int_{\theta^*}^{1} s(\theta)\, dF(\theta) / [1 - F(\theta^*)]$, where $s(\theta) \geq 0$ is the expected surplus enjoyed by the receiver after examining a message of type θ. The receiver may be constrained by the number of messages received, so that

$$\Phi(n) = \min\left\{n, C'^{-1}(s_{av})\right\}. \tag{2}$$

In the sequel we consider $s(\theta)$ either constant, positively or negatively correlated with $\pi(\theta)$. Positive correlation implies that s_{av} is decreasing with n because receiving more messages means adding those of lower expected value to the receiver, and so $\Phi(n)$ is decreasing when $\phi < n$. Negative correlation implies $\Phi(n)$ is increasing. The REF is illustrated for $s(\theta) = \bar{s}$ constant in Figure 2.

INSERT FIGURE 2. Receiver Examination Function $\Phi(n)$ when $s(\theta) = \bar{s}$.

The receiver is effectively supply constrained (she would examine more messages if they were sent) up to $n = C'^{-1}(\bar{s})$, and beyond that point she will always examine the same number of messages.

[15] We allow for different media below. The fact that the sender has paid a higher price provides information before the letter is opened. For example, there may be two types of letter (with different postage costs, like registered and bulk mail).

3 Information Overload

3.1 Equilibrium (Constant receiver benefits)

Equilibrium is a consistency condition that the agents on each side rationally and correctly anticipate the actions of the agents on the other side of the market. This is simply where the sender transmission function intersects the receiver examination function. Thus, an equilibrium will be described by a pair (ϕ^e, n^e) such that $N(\phi^e) = n^e$ and $\Phi(n^e) = \phi^e$.

Suppose $s(\theta)$ is constant. There is one equilibrium (apart from the trivial equilibrium at which no messages are sent and none are examined). Depending on parameter values, either all messages are examined, or else only a subset of those sent are. An increase in γ has no effect on the receiver examination function. The sender transmission function shifts left though. Therefore, $\gamma_1 > \gamma_2 > \gamma_3$ in Figure 3.

INSERT FIGURE 3. Equilibrium with constant receiver surplus, $\bar{s}$.

The first case, with STF1 and high γ, involves all messages examined. Receivers are willing to examine more but no more are sent even though messages are examined with probability one. The marginal sender's decision neglects the positive net surplus to the receiver. This logic indicates that there is insufficient transmission. Equivalently, a small subsidy on message transmission will raise total surplus.

In the second case, with intermediate γ, senders examine all messages, but would not examine more if more were sent. Again, the senders wish to send no more even though they are examined with probability one. This is a knife-edge case defined by

$$\pi(\theta_2) = \gamma_2, \quad [1 - F(\theta_2)]\, M = \phi_2^c, \text{ and } C'(\phi_2^c) = s_{av}(\theta_2) \tag{3}$$

where in this section $s_{av}(\theta_2) = \bar{s}$. Subsidies on transmission decrease total surplus for the reason above; taxes reduce it for the reason below. Nevertheless, as shown at the end of the section, the allocation is not first-best optimal.

In the third case, with low γ $(< \gamma_2)$, the receiver examines fewer messages than are sent. Senders would send more messages if more were examined, and there is message congestion at equilibrium. This "over-fishing" will be diminished by raising γ: there will be less rent dissipation by senders, and better θ types will transmit. There is a clear welfare gain here to the "better" senders from eliminating the worse rivals, so higher profit senders are more likely to get attention. Overall sender benefits may rise even if they are not compensated with the extra revenues because the senders with the lowest benefits are foreclosed, rendering the remainder more prominent.[16] The receiver though is unaffected because the examination decision is unchanged.

This discussion is summarized in the following result.

Proposition 1 *For high transmission cost ($\gamma > \gamma_2$), all messages sent are examined and a small subsidy on transmission raises total surplus. For low transmission cost ($\gamma < \gamma_2$), only a fraction of the messages sent are examined, and a small tax on transmission raises total surplus.*

This means a tax is indicated if senders are overactive, and a subsidy if they are underactive.

3.2 Optimal examination: bored receivers and hyperactive senders

The first best optimum is unattainable if the equilibrium has congestion.[17] The optimum has no congestion. If (higher) pricing were used to price out congestion (reaching the "elbow" point in Figure 3, at $n_2^{\max}$), transmission volume would be too low because the receiver examination decision does not account for sender profits. On the other hand, if the status quo has no congestion, then it may be possible to attain the optimum by reducing the transmission rate and so inducing more messages to be sent (which will be examined as long as the receiver is not sated).

Formally, since the optimum necessarily has no congestion (and so is at a point on the $\phi = n$ locus in Figure 3), the welfare function is

$$W = M \int_{\hat{\theta}}^{1} \pi(\theta)\, dF(\theta) - n\gamma - C(n) + M \int_{\hat{\theta}}^{1} s(\theta)\, dF(\theta), \tag{4}$$

[16] Van Zandt (2004) allows targeting of different consumer types. Then a rate increase may benefit all senders because low profit opportunities are crowded out. This raises the profits of the remaining senders (which now have better prospects for being examined). All sellers' profits may rise if different senders have high profits with different receivers.

[17] The same reasoning applies to the set-up of the next section.

where $\hat{\theta} = F^{-1}\left(1 - \frac{n}{M}\right)$. The welfare derivative is

$$\frac{\partial W}{\partial n} = \left[\pi\left(\hat{\theta}\right) - \gamma\right] - \left[C'(n) - s\left(\hat{\theta}\right)\right]. \tag{5}$$

Clearly, the optimum equates the benefit to the marginal sender and net marginal cost to the receiver, i.e., $\pi\left(\hat{\theta}\right) - \gamma = C'(n) - s\left(\hat{\theta}\right)$.[18]

If the equilibrium is congested (STF3 in Figure 3), the marginal cost is zero at the elbow, $n_2^{\max}$, and marginal benefit is positive since this point is above STF3. Hence welfare is locally rising along the $\phi = n$ locus. However, at the point where STF3 crosses the $\phi = n$ locus, welfare is locally falling since there marginal benefit is zero and marginal cost is negative (it is above the REF). Hence the optimum is on the 45 degree line between $n_2^{\max}$ and $n_3^{\max}$. This implies the next result.

Proposition 2 *For $\gamma < \gamma_2$, the first best optimum can be attained with a tax on senders and a subsidy to the receiver.*

The economic problem here is that receivers do not account for firm surplus and so examine too few messages. At the same time, senders have open access, and do not account for deleterious effects on other senders. While this suggests that the optimum should have less sent and more examined, it may also be that more should be both sent and examined. Nonetheless, the direction of the corrective taxes is unambiguous. It is perhaps difficult to envisage subsidizing message examination since the receiver could claim to have examined to collect the subsidy. The platform intermediary market organization (discussed further below for the pricing of access) affords some solutions. Television broadcasters can increase program quality to raise attention span, and radio disc jockeys can announce prizes to attentive listeners.[19]

3.3 Heterogeneous receivers

Households in some zip codes get more bulk mail than others. They are more attractive to senders because the households are more likely to buy, or buy more (or they could be more responsive in terms of opening the

[18] This relation underscores the two biases in the congested equilibrium conditions. These are the two equations that marginal examination cost equal *average* surplus, and *probability-weighted* marginal sender profits equal transmission costs, i.e., $\frac{\phi}{n}\pi\left(\hat{\theta}\right) - \gamma = 0 = C'(\phi) - s_{av}\left(\hat{\theta}\right)$.

[19] See Anderson and Coate (2005) for a description of the business model of advertising-financed media.

mail).[20] A simple way to introduce heterogeneity into the model is to let households be different sizes (or have different likelihood of purchasing). Write the profit of sender θ matching with household h, conditional on the message being examined, as $\tilde{\pi}(\theta, h)$. The simplest size relation is a separable form $\tilde{\pi}(\theta, h) = a(h)\pi(\theta)$, where h is an index on $[0, 1]$ and $a(.)$ is an increasing function, so bigger h households are more attractive.

Suppose that households are targeted, meaning that the senders distinguish on the basis of h. Let too the number of messages examined per household be $\bar{\phi}$. Then the volume of messages sent to a household of type h is determined by the generalized version of (1) as

$$a(h)\pi\left(\hat{\theta}_h\right)\frac{\bar{\phi}}{n_h} = \gamma.$$

To see the effects of size, divide both sides of this equation by $a(h)$ to get $\pi\left(\hat{\theta}_h\right)\frac{\bar{\phi}}{n_h} = \frac{\gamma}{a(h)}$. The LHS has the familiar form (1), and the RHS just scales the transmission cost. Thus a larger household means a lower effective transmission cost, and it follows directly from our earlier analysis that larger households will be more prone to congestion. Figure 3 can therefore be used to illustrate the situation facing different households. STF1 now represents a small household, and STF3 a large one.[21]

A rise in the transmission rate can now be analyzed given our earlier results. There is a decrease in the social surplus associated to small households that are "supply-constrained" by not getting as much mail as they would read. Such a household is already underserved as the sender's calculus disregards the receiver's surplus. At the other extreme, there is an increase in the social surplus associated to households that are already congested, for the reason that higher transmission rates deter the least socially desirable senders. The optimal transmission rate trades off the numbers of households of each type. It would exceed marginal cost if relatively few households are uncongested and their surplus is small. Of course, a transmission rate tailored to household type would be better than one-size-fits-all (similar to the personal price proposed by Ayres and Nalebuff, 2003, although allowing individuals to choose their own prices will lead to excessive

[20] Targeting is more precise than zip code. For example, households are further broken down into whether they are pet-owners, computer users, whether they have previously given to charities, etc. The current material deals with the case where households are ranked by size ("vertical differentiation"): a later section deals with housholds that differ by relative tastes ("horizontal differentiation").

[21] We assume here that the households retain the same examination rate independent of size. This happens if they face the same examination cost function that is low up to $\bar{\phi}$ and exorbitant beyond. Otherwise, if the households have the same costs, smaller ones might examine less due to lower average benefit (which is an extra reason to send them less). On the other hand, "smaller" households may well have lower examination costs.

prices, following the logic below).

The targeting model described above leads to disproportionate overcongestion of the larger households. There are three types of outcome: recipients receive more messages than they examine, they examine all messages received, or they receive no messages (although they would examine everything they got if they got any!)[22] The larger households have a higher equilibrium congestion rate, while smaller ones may face no congestion: a welfare improvement would divert message volume from large to small.

4 Receiver Effort: correlation of surpluses

Suppose that $s(\theta)$ is increasing (decreasing) so that profits and consumer surpluses are positively (negatively) correlated.

4.1 Increasing receiver benefits

Positive correlation is more likely when receivers have a high probability of not liking offers proposed in messages: profits and receiver surpluses are only earned when receivers are interested in taking up the offers. In this case, the average benefit (s_{av}) increases with θ^*. An interior solution to the receiver's problem is given by $s_{av}(\theta^*) = C'(\phi)$. Consider the Receiver Examination Function, starting with high n. As n falls, θ^* rises, and so ϕ rises. With a low enough number of messages sent, the constraint $\phi = n$ is reached. Thereafter, a lower n (higher θ^*) leads to a smaller ϕ. Thus (reading from right to left), the receiver's choice relation traces out an increasing curve until the constraint $\phi = n$ is attained, and then it follows a declining path with $\phi = n$ (see Figure 4). The kink point is for γ_2 defined by (3).

INSERT FIGURE 4. Equilibrium with increasing receiver surplus.

The interesting feature of this case is the beneficial effects of a higher γ when there is congestion. Then, the number of messages examined rises because a higher γ crowds out senders with lower θ. This raises the average surplus from examination, causing the receiver to examine more messages.[23]

[22] A similar situation occurs for referee reports: some only respond to some requests, others do all they are asked, and some would do them but are never asked.

[23] It is possible for aggregate sender surplus to go up too.

Proposition 3 *Suppose that receiver surplus, $s(\theta)$, is increasing. For $\gamma > \gamma_2$, all messages sent are examined. For $\gamma < \gamma_2$, only a fraction of the messages sent are examined and a small tax on γ causes fewer messages to be sent and more to be examined. This raises receiver welfare by improving the quality of the messages for both the sender and the receiver.*

This Proposition highlights the possibility of an extra social benefit ensuing a transmission price rise. The price rise causes higher quality messages to be sent, which leads to higher examination rates.[24] The higher examination rate somewhat curtails the reduction in transmission, but not so much as to overturn the initial reduced transmission.

Consider now the socially optimal choice of intermediary price. A higher γ moves the STF left and leads to a socially better selection of messages (the higher profit ones). Hence, welfare rises by pricing out congestion. This is a fortiori true when $s'(\theta) > 0$ since receivers also benefit from a better selection induced from better senders (and they respond with an improved examination rate). Indeed, as long as the free-access equilibrium has congestion, the (second-best) optimum price is at the REF kink in Figure 4, i.e., at γ_2. The full optimum though has a higher examination rate.

4.2 Decreasing receiver benefits

This case leads to an upward-sloping receiver examination function. It may join the constraint, $\phi \leq n$ and leave it, then join it again, etc. Such an upward-sloping relation allows for multiple equilibria. Figure 5 illustrates four such equilibria.[25]

INSERT FIGURE 5. Equilibria with decreasing receiver surplus.

The stable equilibria are the two where the REF cuts the STF from above. The logic is akin to that for duopoly reaction functions. The zero-equilibrium in the Figure is unstable (true whenever the REF

[24] A similar finding (though in the opposite direction) arises in Engers and Gans (1998): paying referees is not as effective as might be thought since referees may be more likely to refuse knowing that other referees are induced by payments.

[25] The Figure is drawn with a positive vertical intercept for the REF. This arises if $C'(0) < s(1)$, meaning the receiver is interested if only the top profit sender were present. If instead $C'(0) > s(1)$, the REF hugs the horizontal axis until sufficient senders transmit that the average surplus is high enough to make it worthwhile to start examining messages.

meets the STF from above), and the one at $(\phi^c, n^{\max})$ is stable.[26] A low equilibrium level of transmission is sustained when the receiver rationally anticipates a low average surplus from the highest profit sender types. The receiver examines few messages, inducing few senders to transmit. A higher level of transmission can be sustained when the receiver examines many messages in rational anticipation of high numbers sent, and thence high average surplus. Senders respond to high examination with high transmission.

Consider the (stable) second equilibrium in Figure 5. For a small rise in the transmission price, γ, the STF moves left and this equilibrium moves down the REF. A higher transmission price causes fewer messages to be sent, and this in turn leads to lower expected receiver surplus, causing even fewer messages to be sent. This is a vicious circle for the receiver. Indeed, for high enough γ, the top two equilibria disappear: the market can collapse down to a much lower level of transmission (and examination).[27] Decreasing receiver benefits may entail that the only equilibrium has no messages at all. This happens if the REF lies everywhere below the STF, as occurs in the example below.

Equilibria with higher levels of messages involve higher ϕ/n, from the properties of the STF. Active senders are better off when they have a better chance of examination. Higher sender numbers also behoove the receiver because the average expected surplus (s_{av}) is higher. We summarize the discussion with:

Proposition 4 *Multiple equilibria are possible when receiver surplus, $s(\theta)$, is decreasing. These equilibria can be Pareto ranked: equilibria with higher levels of messages transmitted are Pareto superior for both senders and the receiver. At any stable equilibrium, an increase in γ causes both examination and transmission to fall. This fall may be drastic.*

The equilibrium with the highest level of transmission is an obvious candidate for selection by dint of it being Pareto superior.[28] The next example shows that the market may be closed down when it ought optimally to be functioning.

[26] This logic also implies that the no examination/no send equilibrium is unstable for the constant and increasing sender benefits cases.

[27] A lower transmission rate may also have drastic consquences: the middle two equilibra disappear and the message volume may jump up.

[28] The coordination problem is essentially on the side of the senders insofar as the receiver does examine sequentially in practice – and would therefore discover the average quality of messages. Sequential search makes no difference (to the equilibria) because each sender is too small to influence the average quality.

Example. (Market close-down) Suppose that $s(\theta) = 1$ for $\theta \in [0, 1/2]$, and $s(\theta) = 0$ for $\theta \in (1/2, 1]$. Let $F(\theta) = \theta$ and $M = 1$. Then, $s_{av}(\theta^*) = \frac{\frac{1}{2} - \theta^*}{1 - \theta^*}$ for $\theta^* \in [0, 1/2]$, and $s_{av}(\theta^*) = 0$ for $\theta^* \in (1/2, 1]$. Suppose that $C(\phi) = \frac{\phi^2}{2}$. Then the REF, as a function of θ^*, satisfies $\phi^r(\theta^*) = 1 - \frac{1}{2[1-\theta^*]}$ for $\theta^* \in [0, 1/2]$, and $\phi^r(\theta^*) = 0$ for $\theta^* \in (1/2, 1]$. The first part of the curve is a decreasing and convex function. Suppose that $\pi(\theta) = k\theta$, with $k > \gamma$. The STF is given by inverting the relation $k\theta\frac{\phi}{1-\theta} = \gamma$, as long as $\phi \leq 1 - \theta$, i.e., $\theta \geq \gamma/k$. Hence, the STF is $\phi^s(\theta^*) = \frac{\gamma[1-\theta]}{k\theta}$ for $\theta \geq \gamma/k$. For k large enough, the only intersection with the REF is at $\phi = 0$, and the only equilibrium has the market unravel.[29] However, for low k (and γ not too large) the social optimum involves the lowest set of products.

4.3 Gresham's Law of Junk Mail

When profits are negatively correlated with social surplus the wrong products may be emphasized because the market sorts out senders on the basis of profits (see also Spence, 1976). To illustrate, suppose that there are two different classes of product. Let π_i denote the sender benefit of each product in class i, $i = 1, 2$, and similarly let s_i denote the receiver benefit. We assume that $\pi_1 > \pi_2$ and $s_1 < s_2$, so that the first class has higher sender benefit and lower receiver benefit than the second class. There is a large enough number of independent products in each class. (This is similar to the main model except now products are clustered on two points.) In equilibrium, only the high profit senders survive: any low profit sender is driven out of the market since a high profit sender has a bigger incentive to send a message (advertise). Put another way, if the high-profit senders are earning zero expected profits from sending messages, then the low-profit ones cannot enter the market given that the consumer chooses at random which messages to examine. However, the optimum arrangement will have only the low profit senders sending messages if the sum $b = \pi + s$ is higher. In equilibrium, the low-profit senders are chased from the market by the others, even though the social surplus associated with them is higher.

This is reminiscent of Gresham's law - bad junk mail crowds out good. The receiver would examine more messages if she got more of the low-profit ones, but she does not rationally expect to get them. An

[29] For higher k, there are two other equilibria where REF intersects STF. For example, if $k/\gamma = 20$, there are two solutions: $\theta^* = 0.22$ and $\theta^* = 0.42$. From the argument in the text, the former solution describes the stable equilibrium; the no examination/no send equilibrium is also stable.

extreme form of this phenomenon is when the receiver surplus on the high profit messages is below marginal examination cost. Then the high-profit ones crowd out all others, and the market unravels completely because no receiver finds it worthwhile to examine any messages. This is the "lemons" problem of e-mail - some people have closed their accounts because of the preponderance of spam.[30]

Proposition 5 *Suppose that there are two product classes, with many products in each class. Benefits satisfy* $\pi_1 > \pi_2 > \gamma$ *and* $s_1 < s_2$. *If* $C'(0) < s_1$, *and* $\pi_1 + s_1 < \pi_2 + s_2$, *the only equilibrium at which a positive number of messages is sent involves only messages of type* 1, *but the optimum entails only messages of type* 2 *being sent. If* $C'(0) \geq s_1$, *the only equilibrium has no messages sent.*

This suggests that the market failure is likely greater the bigger the negative correlation: few messages are examined but more "worst" ones are trying to get through.

5 Do-Not-Call

5.1 Pas de publicité s.v.p., and the Federal Do-Not-Call List

Some Belgians and Frenchmen have a little sign on their letterboxes saying they do not want advertising flyers. In the US, the Do Not Call List is a successful initiative orchestrated by the Federal Trade Commission that allows people choose not to receive calls from telemarketers. If there is a nuisance cost to receiving the messages, the receiver may refuse to accept them.

Let S^* denote the equilibrium value of expected receiver surplus from examining messages, and let C^* denote the corresponding examination cost. Assume that receiving each message has a constant annoyance cost, ω. The surplus and examination cost are independent of the nuisance cost, which is sunk if receivers intrinsically dislike receiving messages (telemarketing calls especially). Then the private benefit while receiving messages is $S^* - C^* - n\omega$. This benefit is to be compared to the zero benefit the receiver gets by opting out.[31]

[30] 16% of email address changes have been ascribed to excessive spam (htpp://spam-filter-review.toptenreviews.com/spam-statistics.html).

[31] The zero here reflects the surplus from no message transmission: the basic message medium has a positive benefit that is netted out on both sides of any comparison. We suppose that the receiver does not close down the message medium entirely (in the status quo of no restrictions and free access). She does not disconnect her telephone to block out telemarketers, nor close her e-mail account to stop spam.

Suppose that ω is distributed in the receiver population (which has unit mass) with support $[\underline{\omega}, \bar{\omega}]$ with distribution $G(\omega)$, so that different individuals face different annoyance costs, but are otherwise identical. An equilibrium with some, but not all, individuals opting out is then a critical value, $\hat{\omega} \in (\underline{\omega}, \bar{\omega})$ with

$$\hat{\omega} = \frac{S^* - C^*}{n}.$$

The social benefit associated to a receiver with nuisance cost ω accepting messages is

$$\Pi^* - n\gamma + S^* - C^* - n\omega.$$

The two extra terms added here are the gross sender benefits and the transmission costs. A blanket ban is preferred to allowing the nuisance if

$$B^A \equiv \Pi^* - n\gamma + S^* - C^* - n \int_{\underline{\omega}}^{\bar{\omega}} \omega dG(\omega) < 0,$$

so that the social perspective need not prescribe a blanket ban. Allowing individuals to opt out is better than a total ban if

$$B^{oo} \equiv [\Pi^* - n\gamma + S^* - C^*] G(\hat{\omega}) - n \int_{\underline{\omega}}^{\hat{\omega}} \omega dG(\omega) > 0,$$

where $G(\hat{\omega})$ is the fraction of receivers opting in. This is necessarily positive since $S^* - C^* - n\omega$ must be positive for all those opting in (by revealed preference) and also $\Pi^* > n\gamma$, which also holds by revealed preference because senders only transmit when they get non-negative net benefits.

Conversely, allowing opt-out from a status quo of no opt-out improves welfare if $B^{oo} > B^A$, or

$$[\Pi^* - n\gamma + S^* - C^*] [1 - G(\hat{\omega})] - n \int_{\hat{\omega}}^{\bar{\omega}} \omega dG(\omega) < 0.$$

This condition may or may not hold. For example, if $\hat{\omega}$ is close to $\bar{\omega}$ then nearly all individuals opt out, but if $\Pi^* - n\gamma$ is large, allowing opt-out is socially disadvantageous. The reason is that "too many" individuals opt out: the optimal opt-out cut-off, ω^o, is below the privately-chosen one. This is because the individual does not account for the profit of senders at the margin.

Proposition 6 *Allowing opt-out is socially preferable to banning messages entirely, but may be worse or better than allowing free access.*

The first part follows because opting improves welfare by letting in information when both receiver and senders effectively agree; the second part depends simply on whether surplus is greater with or without messages, given that the receiver side achieves a negative total benefit at the status quo of receiving all that senders wish to transmit.

The Do-Not-Call opt-out may also change the profile of messages received. For example, suppose messages are sent by a producer with increasing returns to scale. This likely implies its equilibrium price decreases with volume. Then consumers who opt out of receiving messages cause a higher price for those remaining consumers since less information gets through. A decreasing returns to scale technology has the opposite impact. It could also be that the consumers who exclude themselves have a more inelastic demand (for example) and so price falls when they opt out. If so, the remaining consumers expect higher surplus and so end up examining more messages.

5.2 Want to Call Me? Pay Me

Ayres and Nalebuff (2003) have suggested individual receivers could set their own personal prices to be contacted. This price would reflect the individual's cost of time and nuisance. Opt-out is equivalent to an infinite price and staying in is like a zero price. For those individuals who choose to opt out under an all-or-nothing scheme, allowing them to choose a price at which they can be contacted cannot make either them or senders worse off, and will make senders better off whenever the price induces some message transmission. The drawback is that individual pricing (of gatekeeper access to the individual's attention) puts the market power in the hands of the individual consumer and underplays sender surplus. This effect of personalized pricing overly restricts message volume because the individual acts as a monopolist against the demand curve for accessing her attention. She then overprices relative to the optimum.

To see this more formally, note that the social welfare associated to an individual with nuisance cost ω is the same as in (4) except for subtracting an additional $n\omega$ for the nuisance. The welfare derivative is then (cf. (5))

$$\frac{\partial W}{\partial n} = \left[\pi\left(\hat{\theta}\right) - \gamma\right] - \left[C'\left(n\right) - s\left(\hat{\theta}\right) + \omega\right].$$

The bracketed terms are simply the demand price (above the base level γ) for sending messages for a marginal sender type, and the net marginal cost to the receiver, all given that the receiver examines all messages.

The individual receiver's problem is to maximize

$$p\left(n\right)n - C\left(n\right) + M\int_{\hat{\theta}}^{1} s\left(\theta\right)dF\left(\theta\right) - n\omega,$$

where the first term represents the revenue from the personalized price. This price is the access demand price of the marginal individual sender, which is $p\left(n\right) = \pi\left(\hat{\theta}\right) - \gamma$, given that the cost γ is paid by the sender for transmission. The individual therefore sets marginal cost, as given above for the social welfare problem, equal to marginal revenue. The solution therefore satisfies $\partial\left[\left[\pi\left(\hat{\theta}\right) - \gamma\right]n\right]/\partial n = \left[C'\left(n\right) - s\left(\hat{\theta}\right) + \omega\right]$.[32] Since marginal revenue is below the demand price, the volume of messages is sub-optimal (as is standard with monopoly pricing).

In summary, personalized pricing does not necessarily out-perform the Do-Not-Call list. Welfare on the account of individuals who opt out is higher with pricing because mutually profitable deals are effectively struck with senders. Welfare on the account of those who do accept calls may be lower with pricing because those individuals may excessively restrict access to enjoy monopoly access rents. That situation is to be compared to the excessive volume of calls when those individuals are priced at zero (which induces excess calls).

6 Pricing Access

6.1 A Monopoly Gatekeeper

The open access market organization of earlier sections may arise spontaneously in the marketplace. An alternative market system has an intermediary controlling the volume of messages transmitted through a conduit. Then, the price of transmitting messages may be determined by pure profit maximization concerns. For example, there could be a profit maximizing Broadcast Company, telephone company, Internet Service Provider, or Post Office. Such an entity is the intermediary (or "platform") in a two-sided market, in the parlance of the recent literature. The present context emphasizes the common property problem (and induced

[32] The individual wants to examine all messages sent ex-post as long as $C'\left(n\right) \leq s_{av}$.

congestion) of an open access system. Since access can be priced by a platform which will account for the common property problem, pricing may have a positive effect of reducing congestion. But the intermediary is also interested in the volume of messages since it takes its mark-up on the total number of messages sent. A priori, it is unclear whether it will encourage mailings or just concentrate on the high willingness-to-pay senders. We show below that the intermediary will fully price out the common property problem.

Assume that the receiver examines at most $\bar{\phi}$ messages. This case arises when $s(\theta)$ is constant, so that $C'\left(\bar{\phi}\right) = \bar{s}$. Let $\pi(\theta) = \bar{\pi}\theta$ (without loss of generality[33]) and suppose that $\frac{f}{1-F}$ is an increasing function (this is the monotone likelihood ratio property; equivalently, the log-concavity of $1-F$).

Suppose that there were congestion, so the price a critical sender type $\hat{\theta}$ is willing to pay to be included in the mailing is $p = \bar{\phi}\bar{\pi}\hat{\theta}/n$, where $n = M\left[1 - F\left(\hat{\theta}\right)\right]$. With congestion, some mail is unexamined and $n > \bar{\phi}$. The constant marginal cost of delivering a message is γ.[34] The gatekeeper's profit is then

$$M\left[1 - F\left(\hat{\theta}\right)\right](p - \gamma) = \left(\bar{\phi}\bar{\pi}\hat{\theta} - M\left[1 - F\left(\hat{\theta}\right)\right]\gamma\right),$$

which is increasing in $\hat{\theta}$.[35] If there is congestion, the monopoly is in the inelastic part of its demand curve and can therefore raise profits by pricing higher. Hence, a monopolist would price out any congestion (assuming indeed that all consumers are the same!); hence, $\bar{\pi}\hat{\theta} = p$. Then the gatekeeper faces a simple monopoly problem of maximizing $(p - \gamma) M (1 - F(p))$, which has a straightforward solution.[36] It either prices so high that the message-examination constraint is slack, or else holds with equality.[37]

[33] We standardize $\pi(\theta)$ and look at a distribution of types since this is familiar from monopoly pricing analysis. A later footnote gives the parallel analysis for $f(\theta)$ normalized.

[34] This is a gross over-simplification. The cost structure is an important determinant of the Post Office's pricing policy. Costs and tariffs are lower in bulk mailings that group similar destinations and when the sender uses bar-codes. Non-profit organizations also benefit from lower tariffs.

[35] If each message had a 50-50 chance of being read, a (risk-neutral) sender would pay twice as much to be read for sure. This means that the monopoly could ration message delivery by half and keep revenues the same. It would save on costs and raise profits.

[36] When $s'(\theta) > 0$, the receiver either examines all messages sent or else an amount that decreases in the number sent (see Figure 4). Again the monopoly prices out congestion. Now a higher message price improves the receiver's selection. Pricing causes the receiver to examine more (and better) messages.

[37] We can rephrase this analysis using the equivalent representation with $f(\theta)$ uniform and $\pi(\theta)$ as the profit function. Under congestion, the intermediary's profit function, $(M\left(1-\hat{\theta}\right)(p-\bar{\gamma}) = \bar{\phi}\pi\left(\hat{\theta}\right) - \bar{\gamma}M\left(1-\hat{\theta}\right))$, is increasing in $\hat{\theta}$ (since $\pi'(\theta) > 0$). Without congestion, the intermediary's problem is $\max_{\{\hat{\theta}\}}\left[\pi\left(\hat{\theta}\right) - \bar{\gamma}\right] M\left(1-\hat{\theta}\right)$. The first order condition to this problem yields $\frac{\pi'\left(\hat{\theta}\right)}{\pi\left(\hat{\theta}\right)} = \frac{1}{\left(1-\hat{\theta}\right)}$. The LHS is decreasing under the assumption that the function $\pi\left(\hat{\theta}\right)$ is log-concave (while the RHS is increasing), so there is a unique solution.

We showed earlier that the (second-best) optimal choice of γ also involves pricing out any congestion. Despite this similarity, the monopoly does not necessarily implement the optimum arrangement since it tends to price too high. However, both monopolist and optimum may price at the kink in the REF, i.e., where congestion just ceases. Clearly, there are cases where either monopoly gatekeeper or open access would be preferred in a binary comparison. The monopoly platform likely restricts access too much but the common-property solution has too much access when it is congested (which is the interesting case).[38]

6.2 Comparison with personalized pricing

The monopoly platform disregards the nuisance costs to the receiver (except insofar as it must price out congestion). The individual accounts for her personal costs (nuisance costs and examination costs, as well as any expected receiver surplus) but then exacts a monopoly mark-up. This means that monopoly platform pricing may or may not welfare dominate individual pricing.

The demand price for the monopoly intermediary is that of the marginal sender, namely $\pi\left(\hat{\theta}\right)$ and the gatekeeper monopoly seeks to maximize $\left[\pi\left(\hat{\theta}\right)-\gamma\right]n$, where the term in brackets is the mark-up. It sets the corresponding marginal revenue to zero. The individual sets the same marginal revenue equal to personal marginal cost, $C'(n)-s\left(\hat{\theta}\right)+\omega$. Whether there is more output under monopoly gatekeeper or personalized pricing depends simply on the sign of this cost. If it is negative, personal pricing leads to a higher message volume and is necessarily welfare superior to a monopoly gatekeeper. This outcome is more likely, ceteris paribus, if ω is low.[39] With higher ω, a positive marginal cost may result in a solution closer to the welfare maximand (so that the monopoly gatekeeper is welfare superior). Even higher costs place the monopoly gatekeeper solution with excessive message volume relative to the social optimum.[40] The monopoly ignores the receiver costs and so underprices (allows too much message volume) when these are high enough. With

[38] The two-sided market literature usually considers access pricing for *both* sides of the market. In the current setting, this might mean charging receivers for access too: an access price would extract all receiver surplus. One insight from the two-sided market literature is that the platform may not want to charge for access (even if it could): getting "on board" with sufficient gusto the side that is more desirable to the other side may enable the platform to charge more for access. Thus it could be that the optimal price from the monopoly intermediary could still be zero (or even a subsidy, just like "free" entertainment on the television or radio could be seen as a subsidy to entice prospective customers to advertisers).

[39] Low ω are suggested by figures cited in Beard and Abernethy (2005) that most consumers were unwilling to pay nominal fees for the State do-not-call plans that preceded the Federal one. Positive net costs are suggested by the large number of subscribers to the Federal list.

[40] The gatekeeper solution can, fortuitously, coincide with the optimum. This happens if the gatekeeper price against the sender demand curve equals the sum of transmission cost plus receiver marginal cost.

large enough (disregarded) costs, the extent of underpricing may be so great as to render the personalized price (overpricing) preferable again. In summary,

Proposition 7 *Personalized pricing leads to higher social surplus than a monopoly gatekeeper if and only if nuisance costs are either very high or else very low.*

Of course, there are distributional benefits of personalized pricing, namely that the individual keeps the revenue instead of the intermediary. Another issue that may favor the individual solution is that the gatekeeper is one-size-fits-all (i.e., the telephone company does not extensively offer different rates for contacting different individuals), while the personal solution is by construction individually tailored.

7 Firm Competition

The models above have assumed that each ad is sent by a firm producing a different new good. However, most junk mail is for credit cards, much spam concerns Viagra or mortgages, and many telemarketers call about time-sharing. This leads us to consider the effects of competition when there are many firms that offer the products for sale.

Suppose that all junk mail is from credit card companies and all credit cards are perfectly homogeneous except for possibly their price.[41] Assume that consumers open messages randomly but there is a cost to each additional message examined. The equilibrium has consumers rationally anticipate the prices announced in the mail, and firms maximize profits rationally anticipating other firms' prices and consumer mail-opening. Then, as per the Diamond (1971) paradox, the only equilibrium is that all firms set the monopoly price.[42] Now though, firms will enter to dissipate all rents. Raising the message transmission price necessarily raises welfare by decreasing the amount of rent dissipation of the monopoly profit. With a higher rate, fewer messages are sent to vie for the fixed profit.

[41] US households received just over 6 bn. credit card offers in 2005 (http://core.synovate.comMAILVOL.asp). (The response rate was 0.3%.)

[42] This analysis supposes that all credit cards are homogeneous. Introducing product heterogeneity tempers the extreme results of the Diamond Paradox. Consumers will typically open several envelopes to find a suitable product. This brings firms into competition and brings equilibrium prices down (the original set-up has no competition because only one letter is opened). The rent dissipation problem is muted because consumers typically choose the best of several offers. However, as shown in Anderson and Renault (1999), the number of firms is excessive, implying that an increase in the postage rate is optimal.

This suggests that the junk message problem may be a double common property resource problem when there is competition within product classes. First there is over-fishing for a consumer's attention and second there is over-fishing in any product class (business stealing). The case for high postage rates on junk mail (email) is the strongest when most consumers do not open all of their mail and there are high rents so that there are many competing products within any class. What comes to mind is credit card ads through the regular mail and Viagra through email.

There is though a caveat to this conclusion. When there are multiple senders within a product class and multiple products, the logic of the Diamond paradox breaks down because a receiver may get a second (or further) price quote while searching for other product class offers. This breaks the monopoly price equilibrium because consumers may then have several price quotes before choosing (as in Burdett and Judd, 1983, although that paper considers a single product class). The consequent pricing and transmission/examination equilibrium are left for future investigation.

8 The medium is the message[43]

When there are multiple media for reaching the consumer, the medium through which a message arrives may furnish useful information to the receiver. Different media cost the sender differently, and incur different congestion levels in equilibrium. For example, two types of letters may arrive, with different postage costs, like registered and bulk mail, and the receiver may pay more attention to the former. The fact that the sender has paid a higher price provides information before the message is examined. In equilibrium, we would expect different media to be used by different sender types. Spam email appears to be used by the lowest expected gross profit senders, while telemarketing calls or notifications by FedEx presumably have higher gross returns than bulk mail.

We introduce alternative media with the minimal departure from our base model. We retain the assumption that each sender transmits a single message, and now must choose which medium to use (or not

[43] The aphorism is due to Marshall McLuhan (1964). McLuhan and Fiore (1967), in "The Medium is the Massage" (which title apparently began as a typo that tickled the authors) gives a precis of this and related work.

to transmit at all).[44] For simplicity, each medium entails the same examination cost function, $C(\phi)$. We also suppose that examination costs on any medium are independent of the levels of examination on other media. For comparisons of different media such as telemarketing calls, email messages, bulk postal mail, television commercials and exposure to billboards on the evening commute, this is perhaps not too egregious an assumption.

Let there be J media. The different media entail different transmission costs, and for clarity we ignore ties by assuming

$$\gamma_1 < \gamma_2 < ... < \gamma_J.$$

In equilibrium, lower priced media will be more congested because senders allocate themselves across media and a lower access price must be offset by a higher congestion level – otherwise senders would join any medium (platform) with both low access price and low congestion. Thus we expect email spam to have a lower equilibrium response rate (i.e., lower ϕ/n) than telemarketing.

This self-selection result implies that higher priced media will be used by senders with higher expected profits from reaching receivers. That is, the expected profit differential between two media is higher for a high θ type (and in favor of the more expensive medium) than for the low θ type. To see this property, note that the net benefit of a sender of type θ using medium j is

$$\pi(\theta_j)\frac{\phi_j}{n_j} - \gamma_j,$$

where ϕ_j is the equilibrium examination rate on medium j and n_j is the mass of senders using that medium. Here $\frac{\phi_j}{n_j}$ portrays the "quality" to the sender of the medium, and, just as in models of vertical product differentiation, higher quality is most appreciated (and therefore bought) by those with highest willingness to pay for it. Mathematically, the net benefit expression rises fastest across options available in different media for the highest θ_j.[45]

[44] As before, if transmission costs are high enough relative to profit differences across senders, this will be an equilibrium in an extended context where multiple messages are allowed.

[45] The equilibrium relation is simply illustrated by graphing net benefit for each option on the vertical axis as a function of θ. For given transmission prices and congestion rates, this yields a series of rays with increasing slopes (ϕ_j/n_j) and lower intercepts ($-\gamma_j$) as j rises. Senders choose the highest net benefit options, indicating the upper envelope of rays as the equilibrium

The indifferent sender type, $\hat{\theta}_j$ that determines the cross-over point between media $j-1$ and j satisfies

$$\pi\left(\hat{\theta}_j\right)\frac{\phi_{j-1}}{n_{j-1}} - \gamma_{j-1} = \pi\left(\hat{\theta}_j\right)\frac{\phi_j}{n_j} - \gamma_j$$

which may be rearranged as

$$\pi\left(\hat{\theta}_j\right) = \frac{\gamma_j - \gamma_{j-1}}{\frac{\phi_j}{n_j} - \frac{\phi_{j-1}}{n_{j-1}}}$$

and these $\hat{\theta}_j$ are increasing in j. Hence more expensive media attract more profitable senders: email is used by senders with low expected profits while telemarketing attracts those with high expected profits.

More expensive media are more attractive to the receiver as along as consumer surplus and sender profit are positively correlated (i.e., $s'(\theta) > 0$). This is because the most expensive media carry the most profitable messages and are therefore the most desirable ones. In response, the receiver examines more messages.

Proposition 8 *Higher priced communication channels are less congested and used by higher-profit senders. The receiver pays more attention to such channels if $s'(\theta) > 0$.*

Even though the receiver's attention span is larger for more expensive media, they do not necessarily attract more senders. To see this, consider the special case $s(\theta) = \bar{s}$ so that all messages carry the same expected surplus. The receiver will exert the same examination effort across all media ($\phi = C'^{-1}(\bar{s})$). Since congestion must be lowest at the top, the transmission rates across media satisfy

$$n_1 > n_2 > ... > n_J.$$

The cheapest media are used most and expensive ones least.

One social benefit of multiple media is to enable a better triage of messages because they are sorted into more homogenous groups. As long as receiver surplus is positively correlated with sender profit, receivers will pay more attention to messages in the higher-cost media, which will reinforce the incentives for the high profit types to use them. We now show how this sorting mechanism works with an example of two channels.

benefit level. The horizontal distance between the intersections of the rays on the envelope indicates the mass of senders using each medium. In equilibrium, this is consistent with the congestion levels that determine the rays' slopes, and the examination levels themselves are consistent with the receiver's choice of examination.

The simplest case has $f(\theta) = 1$ and constant surplus per message, $\bar{s}$, and so the same number of messages ($\phi = C'^{-1}(\bar{s})$) is examined on each medium.

The higher priced medium, 2, will attract the top θ senders, in number $n_2 = M(1 - \theta_2)$, while medium 1 will attract $n_1 = M(\theta_2 - \theta_1)$ senders. Assume that the media have no social cost, and we want simply to show that differential pricing is better than equal pricing. The surplus function is then measured by aggregate profit. This is

$$W = \frac{M \int_{\theta_1}^{\theta_2} \pi(\theta)\, d\theta}{n_1} \phi + \frac{M \int_{\theta_2}^{1} \pi(\theta)\, d\theta}{n_2} \phi,$$

where the first term is the average gross profit to the sellers in the low-θ group and the second term that in the high group. Suppose we take the number of messages as fixed and then vary the composition across media. This entails varying θ_2, with a surplus effect of

$$\frac{1}{M\phi} \frac{dW}{d\theta_2} = \frac{\pi(\theta_2)}{n_1} - \frac{\pi(\theta_2)}{n_2} - \frac{M \int_{\theta_1}^{\theta_2} \pi(\theta)\, d\theta}{n_1^2} + \frac{M \int_{\theta_2}^{1} \pi(\theta)\, d\theta}{n_2^2}.$$

At the limit as γ_2 approaches γ_1 from above, n_2 approaches n_1 and so the support of the high profit senders has the same size as that of the low profit ones ($\theta_2 - \theta_1 = 1 - \theta_2$). The welfare derivative above then reduces to the last two terms, i.e.,

$$\frac{M}{n^2} \left[-\int_{\theta_1}^{\theta_2} \pi(\theta)\, d\theta + \int_{\theta_2}^{1} \pi(\theta)\, d\theta \right].$$

This is necessarily positive because $\pi(\theta)$ is increasing and $\theta_2 - \theta_1 \simeq 1 - \theta_2$: starting from an allocation induced by equal transmission costs, surplus (profits) must rise by moving these costs apart in such a manner that the total volume of messages remains the same. The reason is that better quality messages get more prominence - even though the examination rate is the same in their group - because the group size falls and relieves the congestion by dropping out the lowest quality messages into the other group. Having multiple media therefore enables a better sorting of messages in the market.

9 Indiscriminate mailing (excess blanketing)

While van Zandt (2004) emphasizes different receiver types, we have so far treated the same receiver type. Van Zandt's model is one of targeted transmission, so a sender will not send to a receiver who has low likelihood of reacting to its message. We now consider market research that allows senders to determine which receivers (households) are more likely to be interested in their products. A sender who does not discriminate between households imposes a greater externality by increasing congestion on all other senders and to all households. This implies there is likely insufficient investment in market research (coupled with our overall finding of excessive message transmission).

For simplicity, let there be two classes, the m-agents and the f-agents, with equal mass of each class. Profitability for senders differs across the two types. We assume that receiver types are diametrically opposed in terms of their profitability – what is a high profit product if it reaches one consumer type could be a very low profit type if it reaches another (think of the classic Hotelling linear city model). For the m-agents, profit is given by $\pi_m(\theta)$ which is an increasing function on $[0,1]$ with $\pi_m(0) = 0$. Correspondingly, $\pi_f(\theta)$ is a decreasing function, and we impose symmetry by assuming that $\pi_f(\theta) = \pi_m(1-\theta)$. This means that the profits earned from the m-types are higher for senders with higher θ whereas the f-types yield higher profits for products/senders with lower θ.[46] The extreme θ senders have more incentive to undertake market research to determine agent type since they gain more from excluding the unprofitable class. Conversely, the middling θ's gain little from knowing agent types: they want to sell to both since both have similar valuations.

The expected profit to a sender of type θ who does not invest in market research is

$$\frac{\pi_f(\theta) + \pi_m(\theta)}{2}\frac{\phi}{n} - \gamma, \tag{6}$$

where we have normalized the *total* receiver population to unity, and used symmetry to let n and ϕ denote the common mass of messages sent to each agent type and the common examination rate (these endogenous variables are determined below).

Now suppose that the sender can pay κ to successfully identify the receivers. Let κ be low enough that

[46] As one example, demand for one type could be a multiple of the demand for the other. As long as marginal cost is constant, the price the sender would want to charge each receiver type is the same, and also the same as when the sender does not distinguish the type (the price discrimination motive for determining types is discussed later). Then the differences in the $\pi(.)$ functions fully reflect differences in quantity demanded: if the quantity demanded by an m-agent for sender $\hat{\theta}$'s product were twice that of an f-agent, we would have $\pi_m(\hat{\theta}) = 2\pi_f(\hat{\theta})$.

An alternative underpinning is to assume both m-agents and f-agents have the same demands conditional on being interested at all in the product. Then, different interest probabilities give rise to different values for $\pi(.)$: if the m-agents were statistically twice as likely to be interested in sender $\hat{\theta}$'s product than f-agents, we would again have $\pi_m(\hat{\theta}) = 2\pi_f(\hat{\theta})$.

some senders do use the discriminatory strategy. The total profit from discriminating for a high θ-type (who wants to find the m-agents if $\pi_f(\theta) < \gamma$ since the f-agents do not cover the transmission cost) is

$$\left[\pi_m(\theta)\frac{\phi}{n} - \gamma\right]\frac{1}{2} - \kappa, \tag{7}$$

where the term in square brackets is the expected profit per m-agent. The marginal sender type is defined by $\tilde{\theta}$, obtained by equating (7) with (6), i.e.,

$$\frac{\pi_f(\tilde{\theta}) + \pi_m(\tilde{\theta})}{2}\frac{\phi}{n} - \gamma = \left[\pi_m(\tilde{\theta})\frac{\phi}{n} - \gamma\right]\frac{1}{2} - \kappa, \tag{8}$$

where we note that the LHS exceeds the RHS for $\theta < \tilde{\theta}$, so that lower θ types necessarily prefer not to discriminate. Solving out from (8) yields

$$\pi_f(\tilde{\theta}) = \frac{n}{\phi}\left[\gamma - 2\kappa\right]. \tag{9}$$

For starters, suppose that ϕ is fixed (for example, if all senders' products are worth $\bar{s}$ to the receiver). Then the only other endogenous variable to determine is n. The set of senders to an f-type is all those below $\tilde{\theta}$; equivalently, the set of senders to an m-type is all those above $1 - \tilde{\theta}$. In both cases, we have

$$n = M\tilde{\theta}. \tag{10}$$

In conjunction with (9), this clearly yields a unique interior solution.

Now consider the welfare properties of the solution. Continue to treat ϕ as fixed, and the receiver surplus as the same across messages. Then the total surplus is the same as total profit, up to a constant. Total profit is

$$\begin{aligned}\Pi\left(\tilde{\theta}, \phi, n\right) &= 2\int_0^{1-\tilde{\theta}} \left\{\left[\pi_f(\theta)\frac{\phi}{n} - \gamma\right]\frac{1}{2} - \kappa\right\} d\theta \\ &\quad + 2\int_{1-\tilde{\theta}}^{1/2} \left\{\frac{\pi_f(\theta) + \pi_m(\theta)}{2}\frac{\phi}{n} - \gamma\right\} d\theta,\end{aligned}$$

where we have used symmetry. In $\Pi\left(\tilde{\theta}, \phi, n\right)$, n is given by (10), but we leave it as a separate argument to represent the negative externality across senders: $\Pi\left(\tilde{\theta}, \phi, n\right)$ is clearly decreasing in n in the expression above. The effect on $\Pi\left(\tilde{\theta},\phi,n\right)$ of raising $\tilde{\theta}$ is

$$\frac{d\Pi\left(\tilde{\theta},\phi,n\right)}{d\tilde{\theta}}=\left\{\gamma+2\kappa+\pi_m\left(1-\tilde{\theta}\right)\frac{\phi}{n}-2\gamma\right\}+\frac{\partial\Pi\left(\tilde{\theta},\phi,n\right)}{\partial n}M.$$

When evaluated at the equilibrium as given by (9), the term in curly brackets is zero, so there remains the pure externality effect, $\frac{\partial\Pi(\tilde{\theta},\phi,n)}{\partial n}M<0$. This indicates that there is too little investment in screening since senders disregard the benefits to other senders of withdrawing their messages whence they have lowest social value.

We now add the receiver examination condition into the mix. Assume that $s_m(\theta)$ is increasing (so that profits and sender surplus are aligned) and let $s_f(\theta)=s_m(1-\theta)$ (symmetry). Now, the examination level is determined by the equality of marginal search cost with average surplus, i.e.,

$$C'(\phi)=\frac{\int_0^{\tilde{\theta}}s_f(\theta)\,d\theta}{n},$$

where again $n=M\tilde{\theta}$ by (10).

Now consider the social welfare function:

$$W\left(\tilde{\theta},\phi\right)=\Pi\left(\tilde{\theta},\phi,n\right)+\frac{\int_0^{\tilde{\theta}}s_f(\theta)\,d\theta}{n}\phi-C(\phi)$$

The derivative with respect to ϕ reveals a negative externality that remains at the equilibrium solution:

$$\frac{dW}{d\phi}=\frac{\partial\Pi\left(\tilde{\theta},\phi,n\right)}{\partial\phi}>0,$$

which arises since a greater examination rate by receivers would yield a higher profit to senders. Second,

$$\frac{dW}{d\tilde{\theta}}=\frac{\partial\Pi\left(\tilde{\theta},\phi,n\right)}{\partial n}M+\frac{1}{n}\left[s_f\left(\tilde{\theta}\right)-s_{av}\right]<0,$$

where the first term is the negative one identified earlier, and the second is also negative since the marginal surplus is lower than the average one. This second effect reflects the fact that discrimination raises average quality (to the receiver) of messages received and so induces more examination, but the marginal sender does not internalize this effect.

Proposition 9 *Senders devote insufficient effort to market research, and so blanket mailing is excessive. Excluding marginal types from mailing to their weaker group gives positive benefits to that group and to other senders.*

The model sketched above has the motivation for targeting as saving on transmission costs. Other variants of the model have different reasons for targeting, which may give different partitions of sender types. One such reason is price discrimination. A sender may wish to target different types with different offers – as, famously, does American Express with its card feature offers. The senders that gain most from differential offers would engage in more market research.[47]

The incentives to invest in targeting are likely to be socially insufficient. Untargeted messages hit all households, and so cause greater congestion externalities on senders. When recipients respond to expected quality of messages, congestion is exacerbated because untargeted messages are likely of lower expected benefit to the receiver. These two externalities suggest that not enough resources will be allocated to careful screening of appropriateness of receiver type. One factor that might mitigate this tendency for excess blanketing is that the demand for information is efficiently addressed to intermediaries. Specialized firms can deliver information at low cost due to significant returns to scale in gathering market research data that can be used by many clients. This could greatly reduce the potential inefficiencies. A careful modeling of pricing of providers of such information is needed here.

10 Conclusions

This paper has studied the economics of communication through unsolicited advertising. Receiver (consumer) attention is a scarce resource, but may be considered as "common property" by senders (advertisers), and so be over-utilized. Consumers also need to expend effort to process and absorb the content of unsolicited advertising. In determining how much effort to exert, they consider only the average benefit from the advertising sent (unlike standard markets in which the marginal agents on both sides determine the volume

[47] Hui and Png (2005) survey theoretical and empirical research on privacy, and note that information may impart direct externalities on market participants (for example, in the sending of unsolicited messages, which is our focus).

of transactions), and they ignore the surplus created on the other side of the market. Performance might be enhanced by restricting access to consumers by adjusting pricing or direct regulation.

One regulatory solution actually enacted is the DNC list for telemarketing. Despite widespread success, it does not account for sender surplus. Other proposals (e.g. Ayres and Nalebuff, 2003) would allow recipients to set their own rates: an undergraduate might accept e-mail with a one-cent stamp; a busy chief economist might demand three dollars. Personal pricing correlates access demand prices with nuisance, but gives all the power to the recipient. This could be a desirable property if producer surplus carries a low weight in the social welfare. Although it opens communication that is foreclosed when a recipient exercises her Do-Not-Call right, it can be more distortionary than the effective zero price under Do-Not-Call for those remaining in. It is possible that a monopoly intermediary (which is the standard business model considered in the two-sided markets literature[48]) would price more efficiently than allowing individuals to choose access prices. Although this market structure also has market power, it tends to under-price relative to individual choice because it neglects nuisance costs to consumers. Only when these costs are either very large or small is the individual pricing rule preferred on efficiency grounds. In its defence, it has more equitable distribution (compared to AT&T getting all the proceeds!), and it is tailored.

Like email, people do not pay much attention to junk mail because the average message has too poor quality. Raising the postage rate on bulk mail may improve the allocation of resources through two sources. People recognize only the better offers will be sent, and therefore pay more attention. This mechanism elicits better mail. This surprising possibility may raise more revenue for the Postal Service because firms are prepared to pay more to mail – more messages will be opened if they cost more to send (because sending then signals it must be a worthwhile offer). Interestingly, the January 2006 rise in US Post Office rates has very high rate increases for bulk mail (8.8 to 13.1 cents at the lowest rate versus 37 to 39 cents for first class), in concurrence with this analysis. However, although a rise in the transmission rate will improve the welfare of some receivers who are currently in a congested state, it reduces the welfare of others who get fewer messages when they already desired more.

[48] See Anderson and Coate (2005) for the case of advertising-financed media.

As well as the tendency of senders to congest the receiver, the other dimension of market failure is the low receiver attention span. This can be more difficult to remedy, and indeed impossible with just a tax on messages. It is notable that monopoly power, either in the guise of the individual setting an access price or an intermediary doing so, will price out congestion. The monopoly intermediary has the added possibility (which we do not explore further here) of improving the allure of the message medium (more spending on programming in TV, say).

The attention span problem can also be eased when there are several media that differ by transmission cost. In equilibrium, different media are used by different firm types. Spam email is used by the lowest profit options, while telemarketing calls or notifications by FedEx have higher returns than bulk mail. The fact that the sender has paid a higher price provides content information before the message is examined, giving credence to the maxim "The Medium is the Message." Multiple media raise efficiency by inducing sorting of messages over media and giving more prominence to the more socially desirable messages selecting the more expensive media (and shaking out lower quality messages into cheaper media with more congestion).

The market solution also suffers from a tendency to blanket mailing without enough screening of where messages are sent. There are two positive externalities when a sender screens messages and so drops some addresses. Dropped addresses get a more suitable message profile, and there is less congestion for other senders. While indiscriminate mailing is likely to be more profitable than it is optimal, targeting raises the issue of how the information about consumers is collected and the concomitant worries about consumer privacy.[49] The modeling approach here can usefully be extended to these issues.

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 (2006) Shouting to be Heard in Competing for Attention: Sending Multiple Ads to Get the Message Through. Mimeo, University of Virginia.

[49] Hui and Png (2005) give a valuable survey on the Economics of Privacy; Hann, Hui, Lee, and Png (2005) introduce consumer avoidance (and marketers' consequent efforts to reach them) into an equilibrium marketing analysis.

[3] 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.

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

[5] Bagwell, Kyle (2006) The Economic Analysis of Advertising, in *Handbook of Industrial Organization*, III, forthcoming, Elsevier Science, North Holland.

[6] 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.

[7] Burdett, Kenneth and Kenneth L. judd (1983) Equilibrium Distributions of Sales and Advertising Prices. *Review of Economic Studies*, 44, 465-491.

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

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

[10] Engers, Maxim P. and Joshua Gans (1998): Why Referees are Not Paid (Enough), *American Economic Review*, 88, 1341-1349.

[11] 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.

[12] Hann, Il-Horn, Kai-Lung Hui, Sang-Yong T. Lee, and Ivan P. L. Png (2005) Consumer Privacy and Marketing Avoidance. Working Paper, Department of Information Systems, National University of Singapore.

[13] Hui, Kai-Lung, and Ivan P. L. Png (2005) The Economics of Privacy, in *Handbook of Information Systems and Economics*, Ed. Terry Hendershot, forthcoming, Elsevier Science, North Holland.

[14] McLuhan, Marshall (1964) *Understanding Media.*

[15] McLuhan, Marshall and Quentin Fiore (1967) *The Medium is the Massage*, Random House, 2000 reprint by Gingko.

[16] 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.

[17] 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.

[18] 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.

[19] Toffler, Alvin (1970). *Future Shock.* Random House.

[20] 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.

[21] Van Zandt, Timothy (2004): Information Overload in a Network of Targeted Communication, *RAND Journal of Economics*, 35(3), 542-560.

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

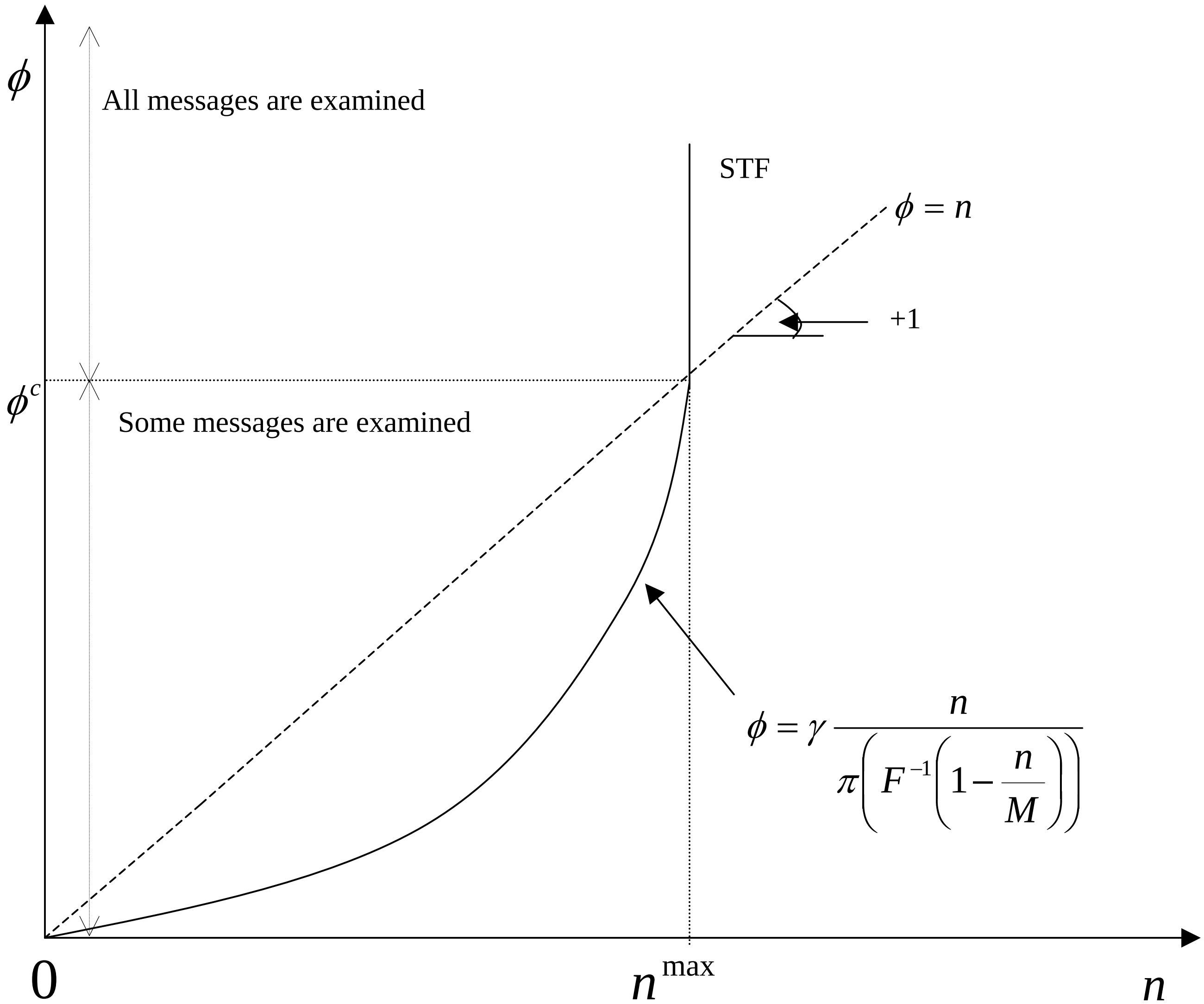

Figure 1. Sender Transmission Function N(ϕ)

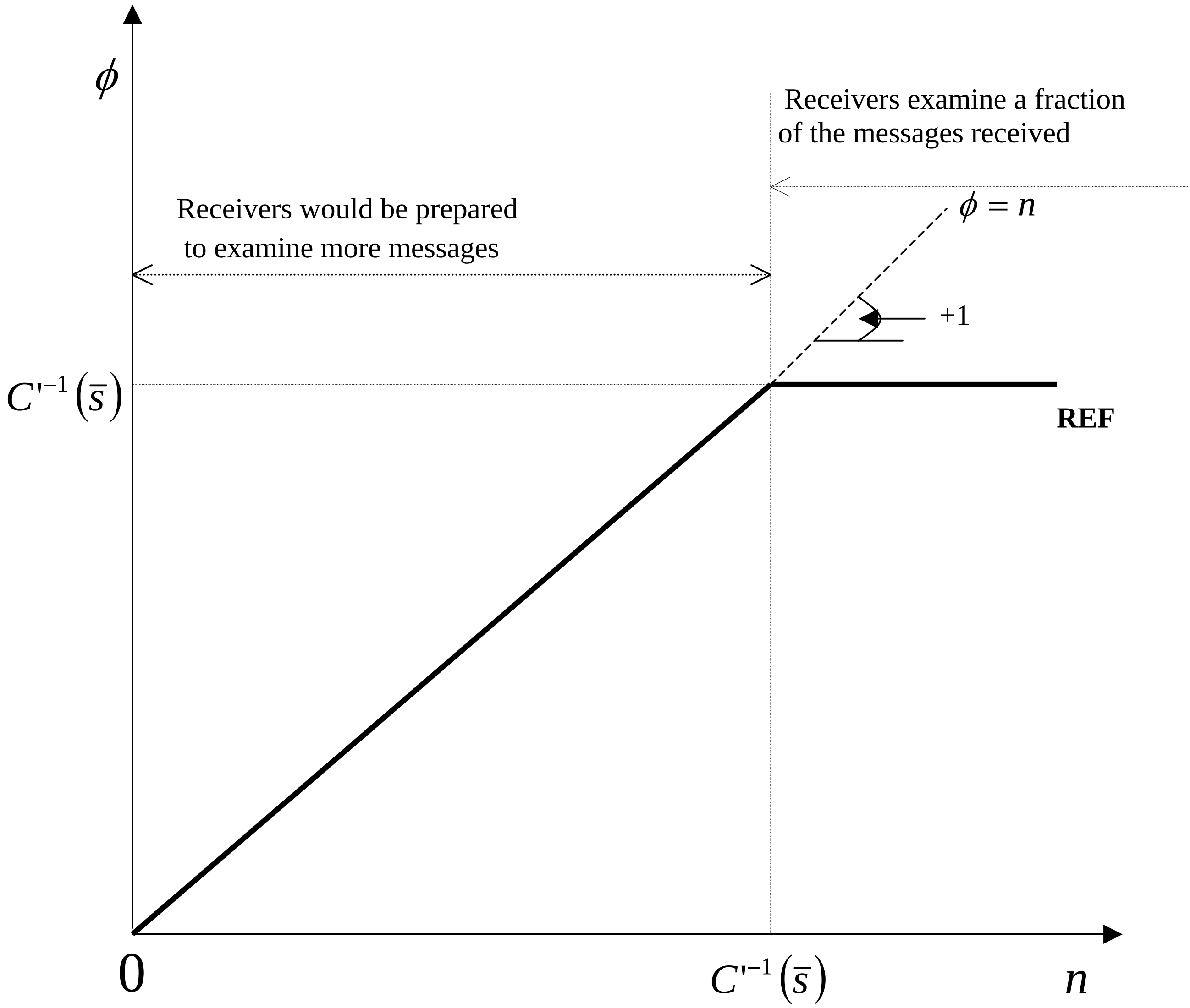

Figure 2. Receiver Examination Function Φ(n) when s(θ) is constant

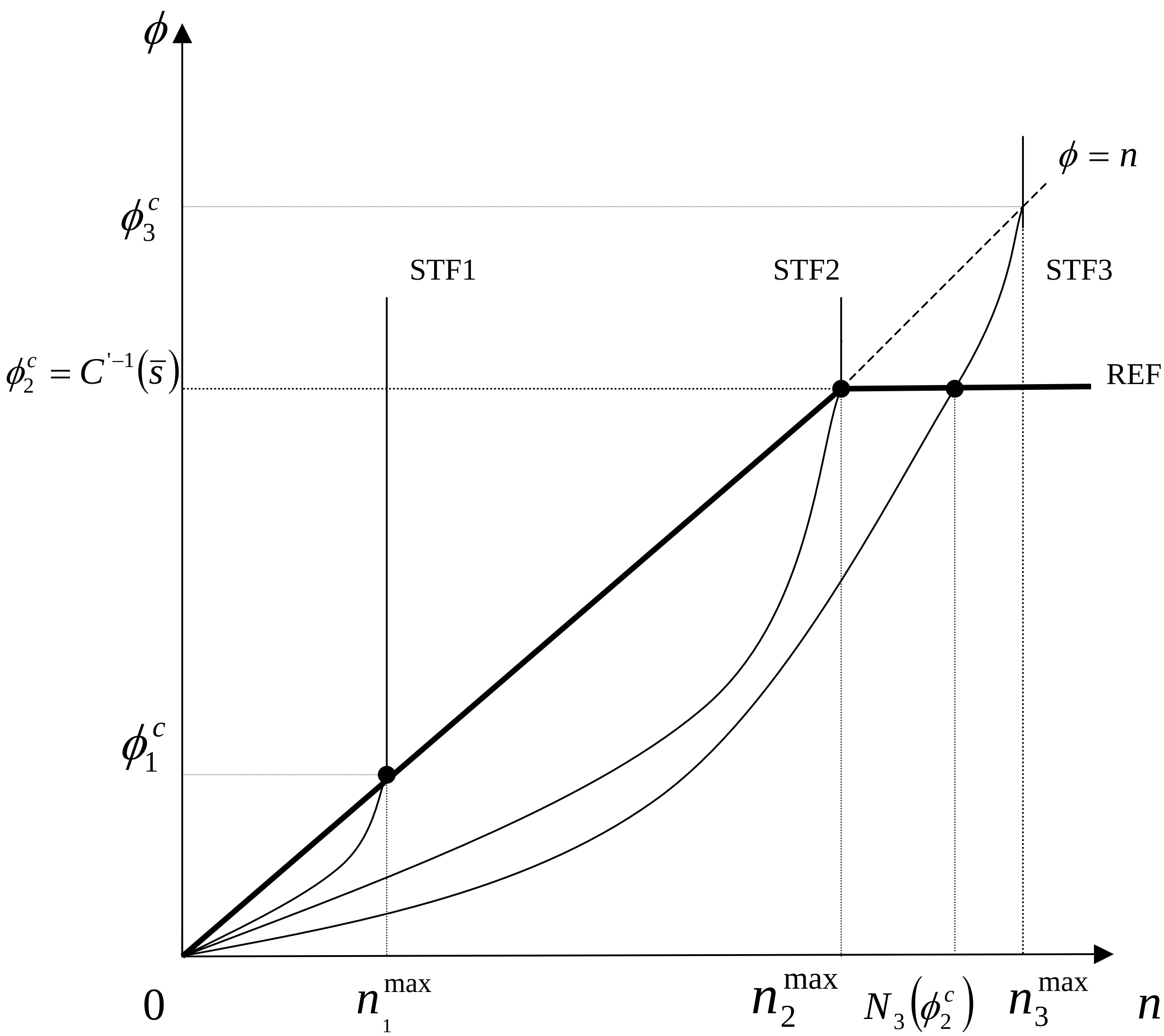

Figure 3. Equilibrium with constant surplus

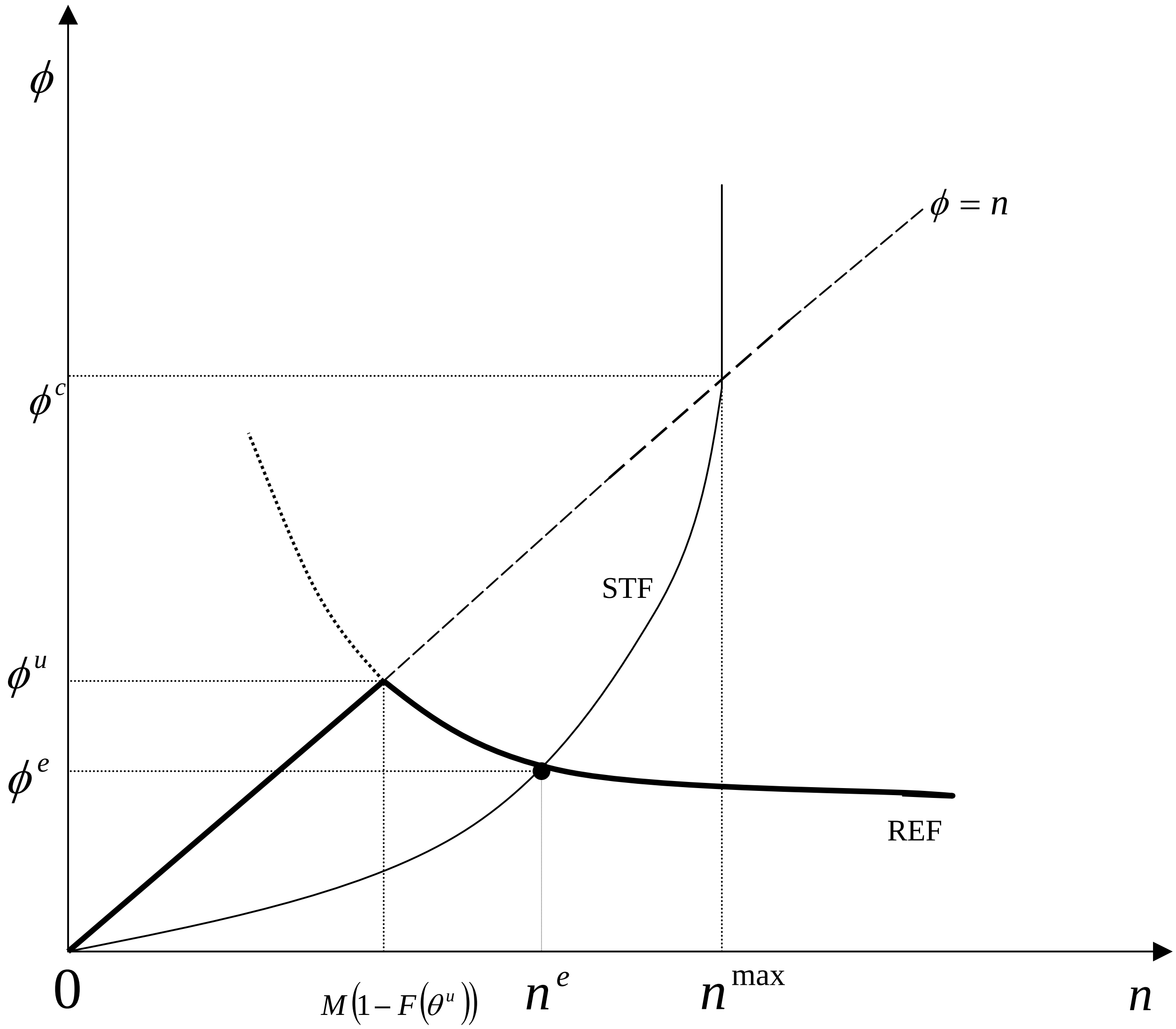

Figure 4. Equilibrium with increasing surplus

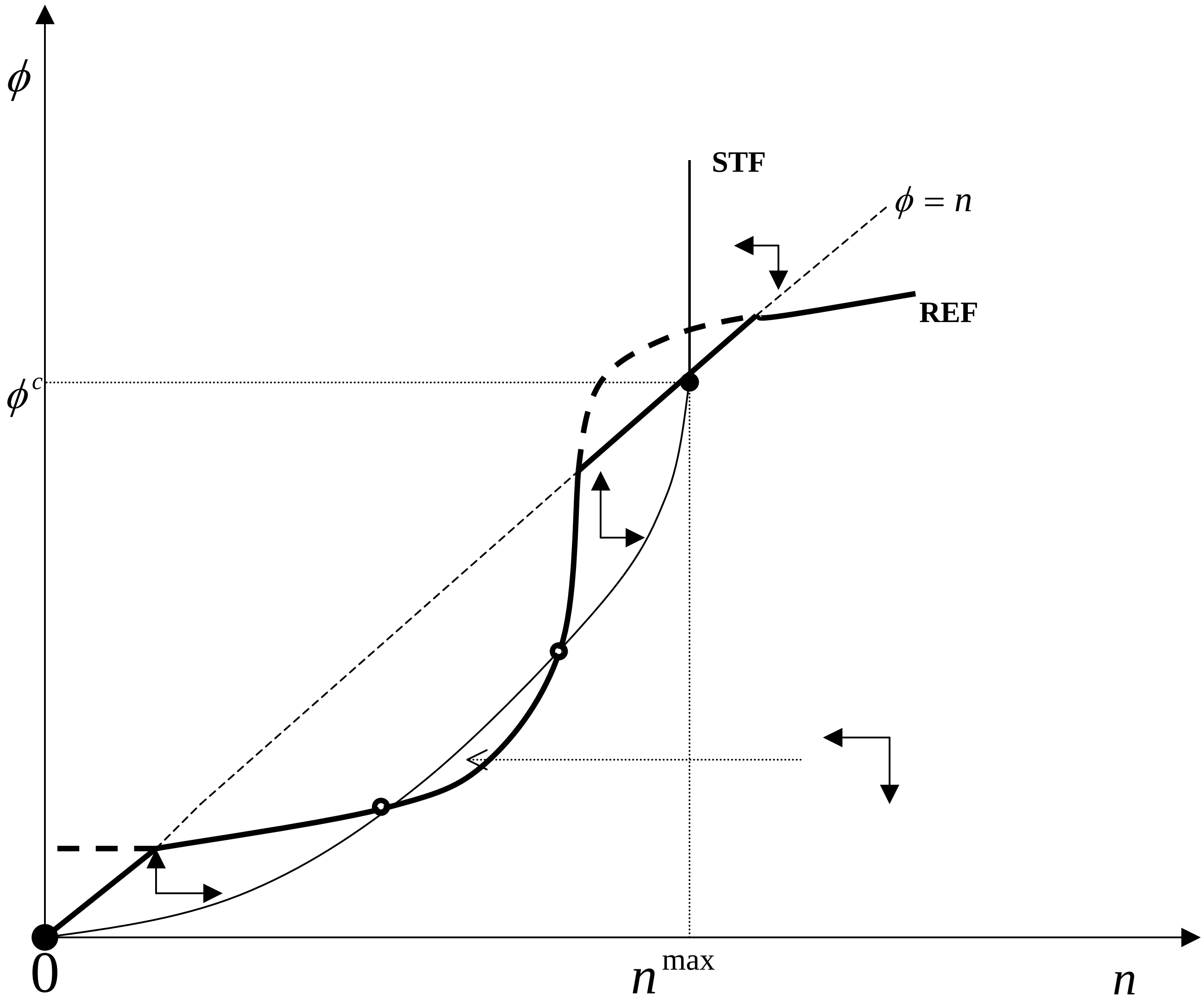

Figure 5. Equilibrium with decreasing surplus.