

Information Congestion: open access in a two-sided market

Simon P. Anderson and André de Palma*

June 2003; this version April 2007

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, while 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 to receivers are moderate.

JEL CLASSIFICATION: D11, D60, L13.

KEYWORDS: information overload, congestion, advertising, common property resource, two-sided markets, junk mail, email, telemarketing, Do Not Call List, message pricing policy.

*Simon P. Anderson: Department of Economics, University of Virginia, PO Box 400182, Charlottesville VA 22904-4128, USA, sa9w@virginia.edu. André de Palma: 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 Rabah Amir, Maxim Engers, Joshua Gans and Bruno Jullien 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), the 2006 Economics of ICT Workshop (ENST, Paris), and the 2006 CRETE meeting in Rethymon.

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-

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

²The 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.

³The 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.

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

⁵Details 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 price of 39 cents for first class mail.⁷

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. 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 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

⁶Unsolicited 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.

⁷This 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>

of messages examined because the expected quality rises.

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, and, most importantly, by restricting the number of messages (advertising breaks on TV, for example: see Anderson and Coate, 2005, for a description of the business model of advertising-financed media). In this two-sided market context, the coordinating entity is responsible for getting both sides of the market (advertisers and viewers) on board its “platform” in numbers that maximize its profit. It faces a trade-off because more advertisers spoil the experience for the viewers and so reduce the viewer base, but advertisers are willing to pay more for more viewers. Our paper considers the economics of an open-access platform (the individual’s mailbox, for example), without the coordinating owner which attracts one side by limiting the other’s access. Without such rationing, consumer attention is a common property resource for advertisers. The attractiveness of the platform, and hence the desire for the receiver to “join” (supply attention) is given endogenously by the expected profile of messages sent. Joining is voluntary, without payment nor price inducement.

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), and our analysis is complementary to his. He is interested in targeted recipients of messages (his receivers examine a fixed number of messages) and his receivers have different worth to different senders. Therein lies the efficiency benefit in his model from a tax on messages. A small tax may be Pareto-improving 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

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

does not arise in our main model. Instead, we emphasize the interaction of the sending decisions with the examination decision. Examination is treated as exogenous in Van Zandt’s model, but endogenous in ours. Here, both sides of the market exert effort which depends on their anticipation of the other side’s actions. In this context, we analyze policies that alter transmission costs, or that give property or pricing rights to receivers, and compare with the monopoly platform 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 examines 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 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 analyze a single receiver and many senders. The surpluses to each party (conditional on a message being processed) depend only on the sender type, so that there is no “business stealing.” We make this strong assumption in order to focus clearly on competition among messages 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 can also be viewed as the product of the probability of buying (being interested in an advertised product) and the conditional surplus from buying. Although the surplus split may vary quite widely across products, the commonality

of the interest probability to both sender and receiver might (loosely) lead one to expect higher profits go along with higher consumer surplus.

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.

2.1 Information senders

A sender of type $\theta \in [0, 1]$ gets a profit $\pi(\theta) > 0$ conditional on its message being examined. We assume that this expected profit is independent of which other messages are examined. We define types by ranking senders from high to low by their conditional expected profits and assume that $\pi(\theta)$ is strictly decreasing and continuous. Note that θ is the fraction of senders below type θ .

There is a cost, $\gamma \in (\pi(1), \pi(0))$, for sending a message, which will imply that there are always some active senders and some senders “waiting in the wings” in equilibrium. 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.⁹ We assume that at most one message can be sent by each sender to the receiver.¹⁰ 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 is no congestion. The marginal sender type then attains its greatest value, θ^{\max} , which is

⁹Think, 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).

¹⁰All senders will transmit only one message if a second message is not profitable for the highest profit type, $\theta = 0$. In terms of the notation below, a first message generates profit $\pi(0) \frac{\phi}{n} - \gamma$. Two messages imply an examination probability of $\left(1 - \left(1 - \frac{\phi}{n}\right)^2\right)$ and so one message is preferred by sender 0 to two if $\pi(0) \frac{\phi}{n} - \gamma > \pi(0) \left(1 - \left(1 - \frac{\phi}{n}\right)^2\right) - 2\gamma$ or $\gamma > \pi(0) \frac{\phi}{n} \left(1 - \frac{\phi}{n}\right)$ where n solves $\pi(n) \frac{\phi}{n} = \gamma$. 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.

therefore also equal to the maximal number of senders, n^{\max} . This marginal type is determined from the break-even condition as $\pi(\theta^{\max}) = \gamma$; since all lower sender types transmit, the corresponding number of senders is $n^{\max} = \theta^{\max}$. This gives the vertical segment in Figure 1.

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

On the other hand, if $\phi < n^{\max}$, then there is congestion. Given that the receiver cannot tell a priori which messages contain which offers, she examines them at random. The likelihood that any given sent message is examined is then $\frac{\phi}{n}$, where n is the number of messages sent. The profitability of sending a message is $\pi(\theta)$ weighted by this examination probability, and so the marginal sender type, θ^* , is uniquely defined by

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

with $n = \theta^*$ (in the sequel we keep n and θ^* distinct, despite this equality, for contextual clarity). 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 slopes up.¹¹ A higher transmission cost, 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.

We assume that messages are independent: how and whether the receiver responds to any message does not depend on which other messages are received. This means the only competition between messages is

¹¹In summary, $N(\phi) = n^{\max}$ for $\phi \geq n^{\max} = \pi^{-1}(\gamma)$ and for $\phi < n^{\max}$ its inverse function is given from (1) as $\phi = \frac{n\gamma}{\pi(n)}$.

for the receiver's attention. Let $s(\theta) \geq 0$ be the expected surplus enjoyed by the receiver after examining a message of type θ . Her marginal benefit from opening a message is then just the *average* expected surplus over the configuration of messages received, $s_{av} = \frac{1}{\theta^*} \int_0^{\theta^*} s(\theta) d\theta$. The receiver may be constrained by the number of messages received, so that

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

In the sequel we consider $s(\theta)$ constant, increasing or decreasing. A decreasing relation 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 $\phi < n$. An increasing relation 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.

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. Thus, an equilibrium will be described by a pair (ϕ^e, n^e) such that $N(\phi^e) = n^e$ and $\Phi(n^e) = \phi^e$. This is simply where the sender transmission function and receiver examination function intersect.

Suppose $s(\theta)$ is constant, so $s(\theta) = \bar{s}$. 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 some are. An increase in γ has no effect on the receiver examination function, but it shifts the sender transmission function left. Figure 3 shows, for three transmission costs, $\gamma_1 > \gamma_2 > \gamma_3$, the sender transmission functions and the corresponding equilibria.

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

For the highest transmission cost, γ_1 , so few messages are sent that the receiver examines them all and would examine even more if they were received. The marginal sender's decision neglects the positive net surplus to the receiver, so that total surplus would rise if the value of n increased. This could be achieved by subsidizing message transmission.

In the second case, with intermediate γ , receivers 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

$$\pi(\theta_2) = \gamma_2, \quad n_2 = \theta_2 = \phi_2^c, \quad \text{and} \quad C'(\phi_2^c) = s_{av}(\theta_2) \quad (3)$$

where in this section $s_{av}(\theta_2) = \bar{s}$. Subsidies on transmission costs 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. 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, $\gamma > \gamma_2$ as given in (3), all messages sent are examined in equilibrium 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 in equilibrium, 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. Van Zandt (2004) obtains a welfare-improving tax through a different mechanism. He allows targeting of different consumer types. Then a price 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.

3.2 Optimal examination: bored receivers and hyperactive senders

As we show below, the first best optimum is unattainable if the equilibrium has congestion (the same reasoning applies to the set-up of the next section). This is because 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 price 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

$$W(n) = \int_0^{\hat{\theta}} \pi(\theta) d\theta - n\gamma - C(n) + \int_0^{\hat{\theta}} s(\theta) d\theta, \quad (4)$$

where $\hat{\theta} = n$. The welfare derivative is

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

Clearly, the optimum equates the benefit to the marginal sender and net marginal cost to the receiver, i.e., $\pi(\hat{\theta}) - \gamma = C'(n) - s(\hat{\theta})$. This relation underscores the two biases in the congested equilibrium conditions. In equilibrium, marginal examination cost equals *average* surplus, and *probability-weighted* marginal sender profits equals transmission costs, i.e., $\frac{\phi}{n} \pi(\hat{\theta}) - \gamma = 0 = C'(\phi) - s_{av}(\hat{\theta})$.

If the equilibrium is congested (STF3 in Figure 3), the marginal cost is zero at the elbow, n_2^{\max} ($= \pi^{-1}(\gamma_3)$), 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 allocation can be attained by a tax on senders and a subsidy to the receiver.

On the other hand, suppose the equilibrium is uncongested and sender constrained in the sense that more messages would be examined if sent (STF1 in Figure 4). Then the marginal benefit in (5) is zero at n_1^{\max} ($= \pi^{-1}(\gamma_1)$) but the marginal cost is negative. Conversely, evaluating at n_2^{\max} ($= \pi^{-1}(\gamma_2)$), the marginal benefit is negative (a marginal sender would make negative profits), and the marginal cost is zero. In this case the optimum is on the 45 degree line between n_1^{\max} and n_2^{\max} , and a subsidy to senders suffices to attain it. Finally, if $\gamma = \gamma_2$, the derivative in (5) is zero at n_2^{\max} because both marginal benefits and marginal costs are zero. This means the full optimum is attained, albeit fortuitously. The next Proposition summarizes.

Proposition 2 *For $\gamma < \gamma_2$, the first best optimum involves $\phi = n \in (\pi^{-1}(\gamma_2), \pi^{-1}(\gamma))$, and can be attained with a tax on senders and a subsidy to the receiver. For $\gamma > \gamma_2$, the first best optimum involves $\phi = n \in (\pi^{-1}(\gamma), \pi^{-1}(\gamma_2))$ and can be attained with a subsidy on senders. If $\gamma = \gamma_2$, the market equilibrium is first best optimal.*

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 when there is congestion. 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.

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 mail).¹² 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 $a(\cdot)$ is an increasing function and h is an index such that 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

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

To see the effects of size, divide both sides of this equation by $a(h)$ to get $\pi(\hat{\theta}_h)\frac{\bar{\phi}}{n_h} = \frac{\gamma}{a(h)}$. 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.¹³

A rise in the transmission price 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 prices deter the least socially desirable senders.

The optimal transmission price trades off the numbers of households of each type. It would exceed marginal

¹²Targeting 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 households which are ranked by size, which is akin to vertical differentiation. Targeting across households that differ by relative tastes is akin to horizontal differentiation.

¹³We 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.

cost if relatively few households are uncongested and their surplus is small. Of course, a transmission price 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 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!)¹⁴ The larger households have a higher equilibrium congestion level, while smaller ones may face no congestion: a welfare improvement would divert message volume from large to small.

4 Receiver effort: variable surpluses

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

4.1 Decreasing receiver surplus

Since profits and receiver surpluses are only earned when receivers are interested in taking up the offers, decreasing surplus is more likely when conditional surpluses are aligned. In this case, the average benefit (s_{av}) decreases 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^* falls, and so ϕ rises. With a low enough number of messages sent, the constraint $\phi = n$ is reached. Thereafter, a lower n (lower 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 decreasing receiver surplus.

¹⁴A 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.

The interesting feature of this case is the beneficial effects of a small tax, τ , on transmission when there is congestion. Then, the number of messages examined rises because a higher transmission price crowds out senders with higher θ . This raises the average surplus from examination, causing the receiver to examine more messages. The receiver is clearly better off (by the envelope theorem) given that s_{av} has risen. Not all senders are better off since the marginal ones are now crowded out, but aggregate sender surplus may go up or down.¹⁵ However, the sum of sender surplus plus tax revenue must rise. To see this, first note that sender surplus plus tax revenue is $\Xi \equiv \int_0^{\hat{\theta}} \pi(\theta) \frac{\phi}{n} d\theta - \gamma n$, where γ is the real resource cost of transmission, with $\pi(\hat{\theta}) \frac{\phi}{n} = \gamma + \tau$ and $C'(\phi) = s_{av}(\hat{\theta})$. As noted above, ϕ rises with τ , and this effect alone improves the surplus, Ξ . So consider the effect of decreasing n . The surplus derivative with respect to n is

$$\frac{\partial \Xi}{\partial n} = \pi(\hat{\theta}) \frac{\phi}{n} - \int_0^{\hat{\theta}} \pi(\theta) \frac{\phi}{n^2} d\theta - \gamma$$

and we wish to show this is negative. It suffices to note that the first two terms can be written as $\frac{-\phi}{n^2} \int_0^{\hat{\theta}} [\pi(\theta) - \pi(\hat{\theta})] d\theta$ which is indeed negative since $\pi(\theta)$ is decreasing. The next Proposition summarizes.

Proposition 3 *Suppose that receiver surplus, $s(\theta)$, is decreasing. For $\gamma > \gamma_2$ (given by (3)) all messages sent are examined. For $\gamma < \gamma_2$, only a fraction of the messages sent are examined and a small tax on transmission causes fewer messages to be sent and more to be examined. This raises both receiver welfare and the sum of sender surplus plus tax revenue. It thus raises total welfare.*

This Proposition highlights the possibility of an extra social benefit ensuing a transmission price rise by crowding out less attractive messages. The price rise thus leads to higher examination rates.¹⁶ 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

¹⁵ Aggregate sender surplus will go up if surplus is very high for low θ types and low for the marginal θ types. Eliminating some of the latter improve the chances of the former and they get more profit despite a higher transmission price.

¹⁶ 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.

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 . As was also true for the case of constant $s(\theta)$, the full optimum though has a higher examination rate.

4.2 Increasing receiver surplus

This case leads to an upward-sloping receiver examination function. Such an upward-sloping relation allows for multiple equilibria, as the following Proposition illustrates.

Proposition 4 *Let $s(\theta) = 0$ for $\theta \in [0, 1/2]$ and $s(\theta) = 1$ for $\theta \in (1/2, 1]$ with $C(\phi) = \frac{\phi^2}{2}$, and let $\pi(\theta) = k[1 - \theta]$ with $k > 0$. Then there exists an equilibrium with $n = \phi = 0$. For $\gamma/k > 1/8$ this is the unique equilibrium; for $\gamma/k < 1/8$ there are two other equilibria. The latter are both congested.*

More generally, the receiver examination function may join the constraint, $\phi \leq n$ and leave it, then join it again, etc. Figure 5 illustrates four such equilibria. The Figure is drawn with a positive vertical intercept for the REF. This arises if $C'(0) < s(0)$, meaning the receiver is interested if only the top profit sender were active. If instead $C'(0) > s(0)$, the REF hugs the horizontal axis until a sufficient number of senders transmit that the average surplus is high enough to make it worthwhile to start examining messages.

INSERT FIGURE 5. Equilibria with increasing receiver surplus.

The stable equilibria are the two where the REF cuts the STF from above. The zero-equilibrium in the Figure is unstable (true whenever the REF meets the STF from above), and the one at (ϕ^c, n^{\max}) is stable.¹⁷ 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.

¹⁷This logic also implies that the no examination/no send equilibrium is unstable for the constant and increasing sender benefits cases.

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. Hence, at any stable equilibrium, an increase in γ causes both examination and transmission to fall. 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). A lower transmission rate may also have drastic consequences: the middle two equilibria disappear and the message volume may jump up.

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 Proposition 4 above.

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. Hence, multiple equilibria are possible when receiver surplus, $s(\theta)$, is increasing and equilibria with higher levels of messages transmitted are Pareto superior. The equilibrium with the highest level of transmission is an obvious candidate for selection by dint of it being Pareto superior.¹⁸

The example of the next sub-section shows that the market may be closed down when it ought optimally to be functioning. It also emphasizes a further property of the example in Proposition 4, that the market solution can involve one set of messages while the optimum involves another set (in Proposition 4, the market outcome entails the low- θ messages while the social optimum may value more highly the high- θ ones).

4.3 Gresham's law of junk mail

When profits are negatively related with social surplus the wrong products may be emphasized because the market sorts out senders on the basis of profits. To illustrate, suppose that there are two different classes of products. Let π_i denote the sender benefit of each product in class i , and similarly let s_i denote the receiver benefit $i = 1, 2$. We assume that $\pi_1 > \pi_2$ and $s_1 < s_2$, so that the first class has higher sender benefit and

¹⁸The 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.

lower receiver benefit than the second class. There is a large enough number of independent products in each class. In equilibrium, only the high-profit senders survive, if any: low-profit senders are 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 active if $\pi_2 + s_2 > \pi_1 + s_1$. In equilibrium, the low-profit senders are chased from the market by the others, even though the social surplus associated with them is higher. The equilibrium then has the “wrong” message types sent as long as $C'(0) < s_1$ (meaning that at least some messages will be examined).

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 extreme form of this phenomenon arises when the receiver surplus on the high-profit messages is below marginal examination cost, i.e., $C'(0) \geq s_1$. Then the only equilibrium has no messages sent - the high-profit ones would 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.¹⁹ This suggests that the market failure is likely greater the bigger the inverse relation: 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 to 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^*

¹⁹ 16% of email address changes have been ascribed to excessive spam (<http://spam-filter-review.toptenreviews.com/spam-statistics.html>).

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. 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

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

where Π^* is the equilibrium value of expected sender surplus. 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 $B^{oo} > B^A$, which condition is given in the next Proposition.

Proposition 5 *Allowing opt-out is socially preferable to banning messages entirely. Opt-out is strictly better than allowing free access if and only if $[\Pi^* - n\gamma + S^* - C^*][1 - G(\hat{\omega})] < n \int_{\hat{\omega}}^{\bar{\omega}} \omega dG(\omega)$.*

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 for ω is below the privately-chosen one. This is because the individual does not account for the profit of senders at the margin. The first part of the Proposition 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.²⁰

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

²⁰The 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 is produced. 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.

(cf. (5))

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

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.²¹

The individual receiver's problem is to maximize

$$p(n)n - C(n) + \int_0^{\hat{\theta}} s(\theta) d\theta - n\omega,$$

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 $p(n) = \pi(\hat{\theta}) - \gamma$, 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 $\partial \left[\left[\pi(\hat{\theta}) - \gamma \right] n \right] / \partial n = \left[C'(n) - s(\hat{\theta}) + \omega \right]$.²² 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).

²¹The following argument shows the receiver does want to examine all messages for $s'(\theta) \leq 0$. Suppose instead there were congestion. Then, noting that $p(n) = \pi(\hat{\theta}) \frac{\phi}{n} - \gamma$ is the demand price of the marginal sender, the receiver's benefits would be (for $\phi < n$) $\left(\pi(\hat{\theta}) \frac{\phi}{n} - \gamma \right) n - C(\phi) + \int_0^{\hat{\theta}} \frac{\phi}{n} s(\theta) d\theta - n\omega$. The derivative of this expression with respect to $\hat{\theta}$ (recalling this is also synonymous with n) is $\pi'(\hat{\theta}) \phi - \gamma + \frac{\phi}{n} \left[s(\hat{\theta}) - \frac{1}{n} \int_0^{\hat{\theta}} s(\theta) d\theta \right] - \omega$. As long as $s(\cdot)$ is not increasing, the term in square brackets is negative, as are all the others. Thus the individual will never tolerate congestion, and will price it out. However, if $s'(\theta) > 0$, the receiver might price to encourage high θ types and not examine all messages.

²²The individual wants to examine all messages sent ex-post as long as $C'(n) \leq s_{av}$.

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. This intermediary is the platform in a two-sided market, in the parlance of the recent literature. The present context emphasizes the common property problem (and induced 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 beneficial 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 congestion.

Assume that the receiver examines at most $\bar{\phi}$ messages. This case arises when $s(\theta)$ is constant, so that $C'(\bar{\phi}) = \bar{s}$. 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 $p = \bar{\phi}\pi(\hat{\theta})/n$, where $n = \hat{\theta}$. With congestion, some mail is unexamined and $n > \bar{\phi}$. The constant marginal cost of delivering a message is γ .²³ The gatekeeper's profit is then

$$\begin{aligned} H(\hat{\theta}) &= \hat{\theta}(p - \gamma) \\ &= \bar{\phi}\pi(\hat{\theta}) - \hat{\theta}\gamma, \end{aligned}$$

which is decreasing in $\hat{\theta}$.²⁴ 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

²³This 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.

²⁴If 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.

indeed that all consumers are the same!); hence, $\pi(\hat{\theta}) = p$. Then the gatekeeper faces a simple monopoly problem of maximizing $H(\hat{\theta}) = (\pi(\hat{\theta}) - \gamma)\hat{\theta}$, which has a straightforward solution.²⁵ The first order condition to this problem yields $\frac{-\pi'(\hat{\theta})}{\pi(\hat{\theta}) - \gamma} = \frac{1}{\hat{\theta}}$. The LHS is increasing under the assumption that the function $\pi(\hat{\theta})$ is log-concave (while the RHS is decreasing), so there is a unique solution. Hence the monopoly gatekeeper either sets a price so high that $C'(n) < s_{av}(\hat{\theta})$ (the message-examination constraint is slack), or else $C'(n) = s_{av}(\hat{\theta})$ (the constraint holds with equality).

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).

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).

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.

²⁵When $s'(\theta) < 0$, 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.

The demand price for the monopoly intermediary is that of the marginal sender, namely $\pi(\hat{\theta})$ and the gatekeeper monopoly seeks to maximize $[\pi(\hat{\theta}) - \gamma]n$, 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, $C'(n) - s(\hat{\theta}) + \omega$. 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.²⁶ 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.²⁷ The monopoly ignores the receiver costs and so underprices (allows too much message volume) when these are high enough. With large enough (disregarded) costs, the extent of underpricing may be so great as to render the personalized price (overpricing) preferable again. The next Proposition is proved in the Appendix.

Proposition 6 *Suppose that $\pi(\theta)$ is log-concave and $s(\theta)$ is non-increasing, and suppose that $C'(n_m) + \omega < s(\hat{\theta}_m)$, where $n_m = \hat{\theta}_m$ denotes the unique solution for message volume to the monopoly gatekeeper's problem. Then a monopoly gatekeeper generates higher social surplus than personalized pricing if nuisance costs $\omega \in (\omega_1, \omega_2)$. Personalized pricing generates higher social surplus if $\omega < \omega_1$ or $\omega > \omega_3$. Here $\omega_1 = s(\hat{\theta}_m) - C'(n_m)$, $\omega_2 = \pi(\hat{\theta}_m) - \gamma + s(\hat{\theta}_m) - C'(n_m)$, and $\omega_3 = \pi(0) - \gamma + s(0) - C'(0)$.*

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 prices for contacting different individuals), while the personal solution is by construction individually tailored.

²⁶Low 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.

²⁷The 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.

7 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 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 prices: 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) 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 little interest. 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 price 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 price 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.

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).

Finally, we 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. Suppose that all junk mail is from credit card companies and all credit cards are perfectly homogeneous except for possibly their price.²⁸ If consumers open messages randomly and there is a cost to each additional message examined, then, as per the Diamond (1971) paradox, the only equilibrium is that all firms set the monopoly price.²⁹ 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 price, fewer messages are sent to vie for the fixed profit.

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

²⁸US households received just over 6 bn. credit card offers in 2005 (<http://core.synovate.comMAILVOL.asp>). (The response rate was 0.3%.)

²⁹This 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.

(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 APPENDIX

Proof of Proposition 4.

Proof. The STF is given by $\pi(n) \frac{\phi}{n} = \gamma$; for the specification $\pi(\theta) = k[1 - \theta]$ this means the (inverse) function is $\phi^S = \frac{\gamma}{k} \frac{n}{1-n}$, which is an increasing and strictly convex function on $[0, 1)$. The average receiver surplus is $s_{av}(n) = 0$ for $n \leq 1/2$, and $s_{av}(n) = 1 - \frac{1}{2n}$ for $n \geq 1/2$. Since $C'(\phi) = \phi$, the REF is given as $\phi^R = 0$ for $n \leq 1/2$, and as $\phi^R = 0 = 1 - \frac{1}{2n}$ for $n \geq 1/2$. This is a strictly concave function for $n \in (0, 1)$. For $\gamma/k > 1/8$ these functions cross only once, at $n = \phi = 0$. At $\gamma/k = 1/8$ the functions are tangent at $n = 2/3$. For $\gamma/k < 1/8$ there are two other solutions (in addition to the one at $n = \phi = 0$) and they involve one value of $n \in (\frac{1}{2}, \frac{2}{3})$ and the other with $n \in (\frac{2}{3}, 1)$: $\phi < n$ for both so they are both congested. ■

Proof of Proposition 6.

Proof. Note first that $\omega_1 < \omega_2 < \omega_3$: $\omega_1 < \omega_2$ because $\pi(n_m) > \gamma$ under monopoly, and $\omega_2 < \omega_3$ since $\pi(\cdot)$ and $s(\cdot)$ are decreasing while $C'(\cdot)$ is increasing.

Define the solution to the monopoly gatekeeper's problem as $n_m (= \hat{\theta}_m) = \arg \max_n [\pi(n) - \gamma]n$: the solution is uniquely determined since $\pi(\theta)$ is log-concave, which implies the objective function is quasi-concave. It is also plainly independent of ω . Furthermore, the assumption $C'(n_m) + \omega < s(\hat{\theta}_m)$ ensures that there is no message congestion since the receiver examines all messages sent whenever $C'(n_m) < s(\hat{\theta}_m)$.

The solution to the personal pricing problem (given no congestion, which has been already shown) is $n_p = \arg \max_n \left\{ \left[\pi(\hat{\theta}) - \gamma \right] n - C(n) - n\omega + \int_0^{\hat{\theta}} s(\theta) d\theta \right\}$ with $\hat{\theta} = n$. This problem is also quasi-concave, with a solution that is continuous and decreasing in ω (and strictly decreasing as long as the solution is positive).

The social welfare function, as long as there is no congestion, is $W(n) = \int_0^{\hat{\theta}} \pi(\theta) d\theta - n\gamma - C(n) - n\omega + \int_0^{\hat{\theta}} s(\theta) d\theta$, which is a strictly concave function of $\hat{\theta}$ (for $s(\cdot)$ non-increasing, as assumed). Its derivative is

$$\frac{dW(n)}{dn} = \pi(\hat{\theta}) - \gamma - C'(n) - \omega + s(\hat{\theta}). \quad (6)$$

First consider $\omega < \omega_1$. Since $\omega < s(\hat{\theta}_m) - C'(n_m)$, for such values, then $n_p > n_m$. Given the concavity of $W(\cdot)$, it therefore suffices to show that $\frac{dW(n_p)}{dn} > 0$ in order to prove that $W(n_p) > W(n_m)$. Since n_p solves $[\pi(n_p) - \gamma] + \pi'(n_p)n_p - C'(n_p) - \omega + s(\hat{\theta}_p) = 0$, the welfare derivative (6) reduces to $\frac{dW(n_p)}{dn} = -\pi'(n_p)n_p$, which is positive, as desired.

For $\omega_2 > \omega > \omega_1$, then $n_p < n_m$. Since $W(\cdot)$ is concave, it suffices to show that $\frac{dW(n_m)}{dn} > 0$ in order to prove that $W(n_m) > W(n_p)$ for $\omega < \omega_2$. The derivative (6) is then $\frac{dW(n_m)}{dn} = \pi(\hat{\theta}_m) - \gamma - C'(n_m) - \omega + s(\hat{\theta}_m)$. Since $\omega < \omega_2 = \pi(\hat{\theta}_m) - \gamma + s(\hat{\theta}_m) - C'(n_m)$, this is clearly positive, as desired.

For $\omega > \omega_3$, $n_p = 0 < n_m$. The private pricing solution here is to price out all messages because the receiver's benefit derivative is $[\pi(n_p) - \gamma] + \pi'(n_p)n_p - C'(n_p) - \omega + s(\hat{\theta}_p)$, but since $\omega > \omega_3 = \pi(0) - \gamma + s(0) - C'(0)$, this derivative is negative at $n_p = 0$, and, by quasi-concavity of the benefit function, the solution is therefore zero. We therefore wish to show that $\frac{dW(0)}{dn} < 0$ to prove that $W(n_p) > W(n_m)$ by concavity of $W(\cdot)$. The derivative (6) is $\frac{dW(0)}{dn} = \pi(0) - \gamma - C'(0) - \omega + s(0)$, but this is negative, as desired (and equals $\omega_3 - \omega$), for $\omega_3 < \omega$. Q.E.D.

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 REGIS RENAULT (1999): Pricing, Product Diversity and Search Costs: a Bertrand-Chamberlin-Diamond Model. *RAND Journal of Economics*, 30, 719-735.
- [3] AYRES, IAN AND BARRY NALEBUFF (2003) Want to call me? Pay me! *Wall Street Journal*, October 8th, 2003.
- [4] BAGWELL, KYLE (2006) The Economic Analysis of Advertising, in *Handbook of Industrial Organization*, III, forthcoming, Elsevier Science, North Holland.
- [5] 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.
- [6] BURDETT, KENNETH AND KENNETH L. JUDD (1983) Equilibrium Distributions of Sales and Advertising Prices. *Review of Economic Studies*, 44, 465-491.
- [7] DIAMOND, PETER A. (1971): A Model of Price Adjustment. *Journal of Economic Theory*, 3, 156-168.
- [8] DIXIT, AVINASH K. AND VICTOR D. NORMAN (1978): Advertising and Welfare. *Bell Journal of Economics*. 9, 1-17.
- [9] ENGERS, MAXIM P. AND JOSHUA GANS (1998): Why Referees are Not Paid (Enough), *American Economic Review*, 88, 1341-1349.
- [10] 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.
- [11] 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.
- [12] 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.

- [13] 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.
- [14] TOFFLER, ALVIN (1970). *Future Shock*. Random House.
- [15] 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.
- [16] VAN ZANDT, TIMOTHY (2004): Information Overload in a Network of Targeted Communication, *RAND Journal of Economics*, 35(3), 542-560.
- [17] WILLMORE, C. (1999) A Penny for your Thoughts: E-mail and the Under-Valuation of Expert Time. *Working Paper*, University of Warwick.

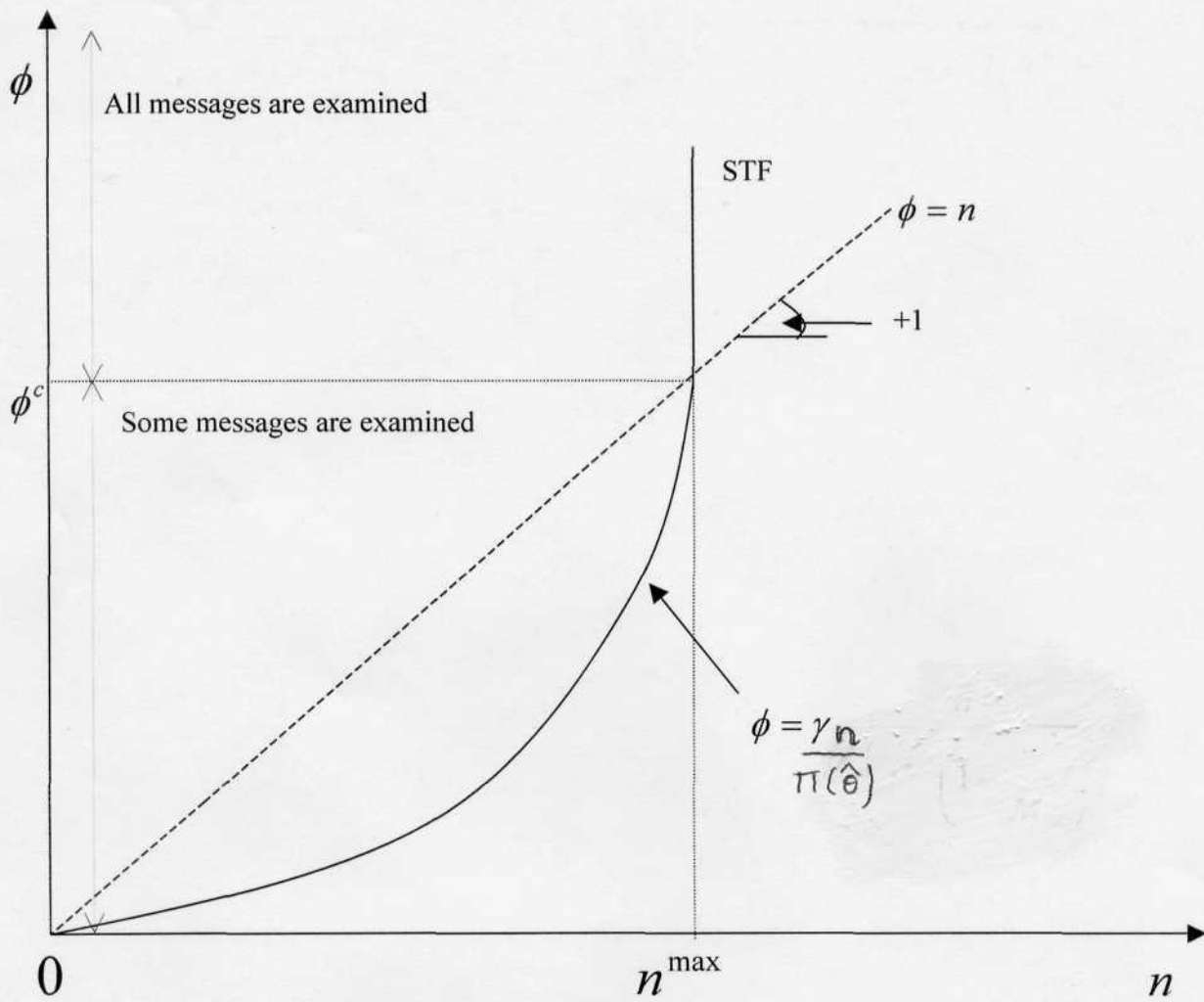


Figure 1. Sender Transmission Function $N(\phi)$

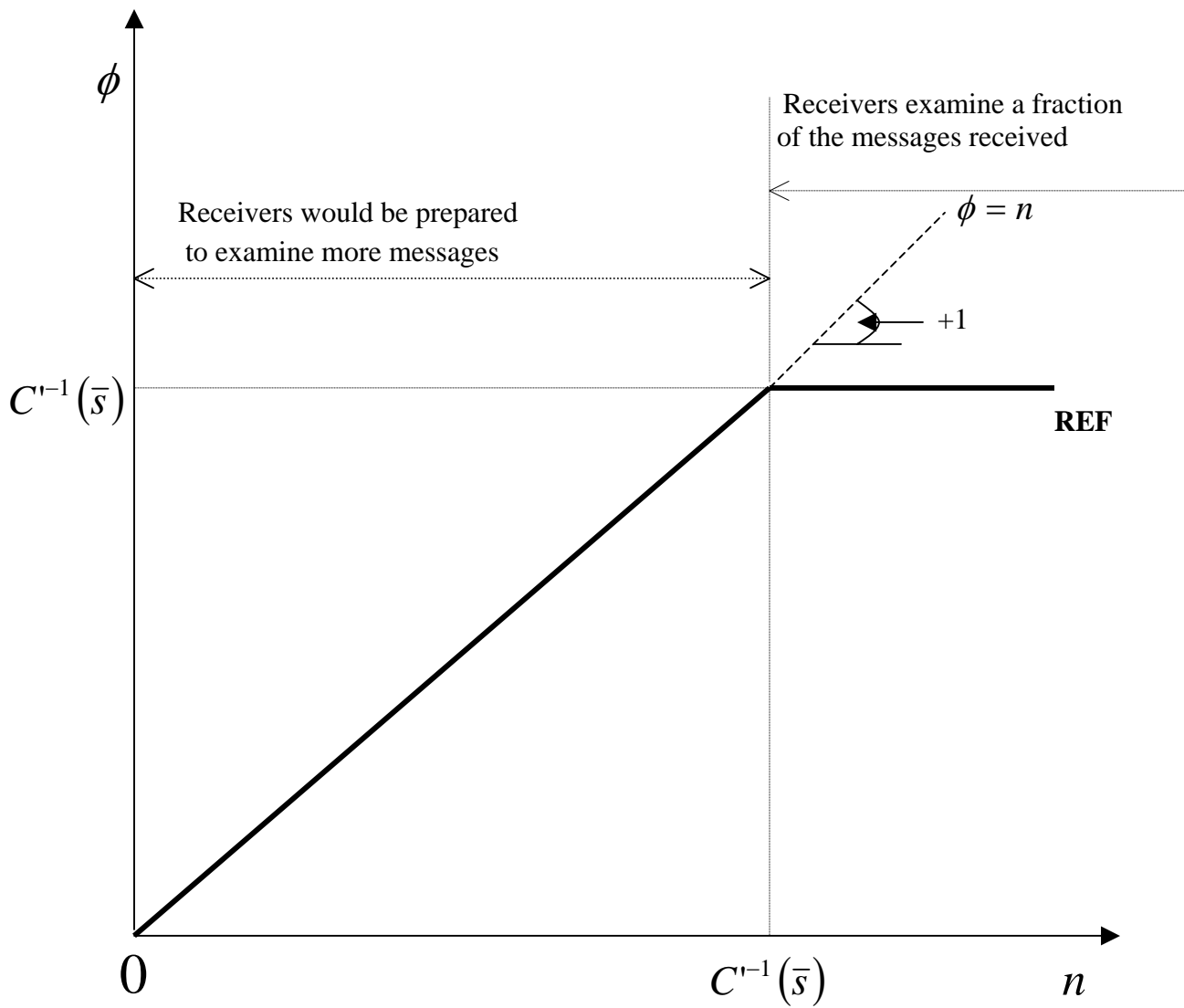


Figure 2. Receiver Examination Function $\Phi(n)$ when $s(\theta)$ is constant

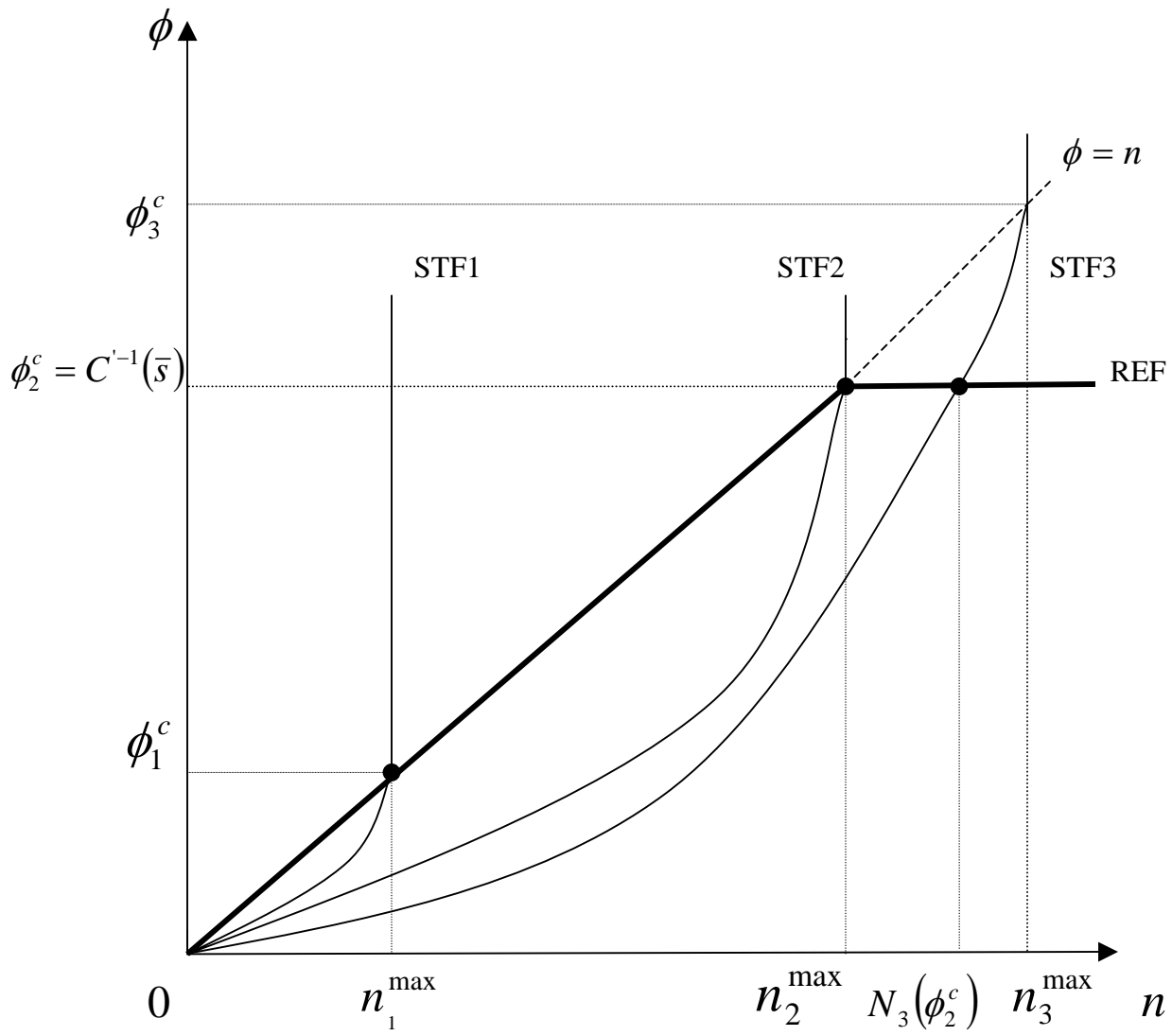


Figure 3. Equilibrium with constant surplus

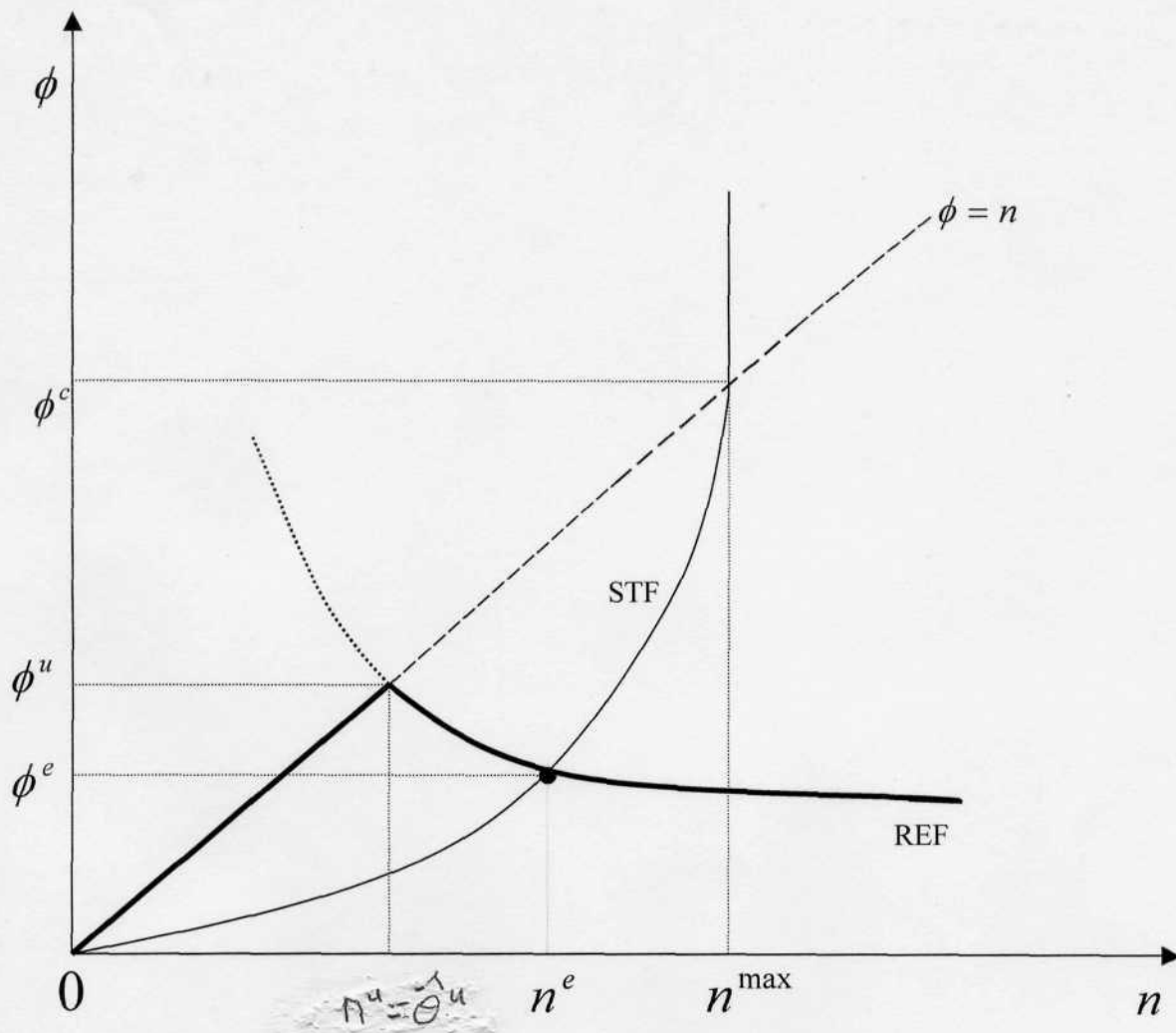


Figure 4. Equilibrium with increasing surplus

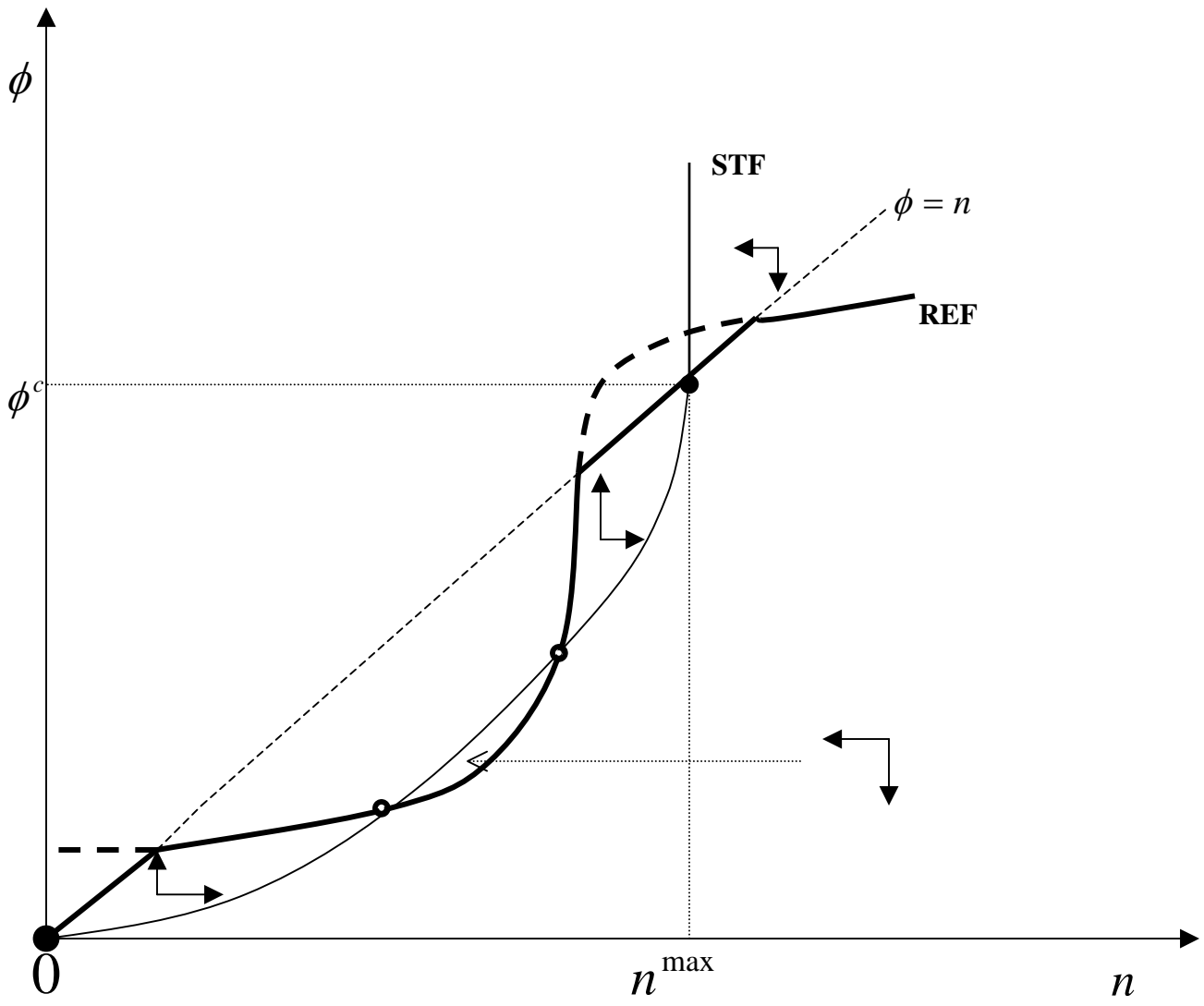


Figure 5. Equilibrium with decreasing surplus.