An Economic Response to Unsolicited Communication

Thede Loder* Marshall Van Alstyne† Rick Wash‡

May 23, 2006

Abstract

If communication involves some transactions cost to both sender and recipient, what policy ensures that correct messages – those with positive social surplus get sent? Filters block messages that harm recipients but benefit senders by more than transactions costs. Taxes can block positive value messages, and allow harmful messages through. In contrast, we propose an "Attention Bond," allowing recipients to define a price that senders must risk to deliver the initial message.

The underlying problem is first-contact information asymmetry with negative externalities. Uninformed senders waste recipient attention through message pollution. Requiring attention bonds creates an attention market, effectively applying the Coase Theorem to price this scarce resource. In this market, screening mechanisms shift the burden of message classification from recipients to senders, who know message content. Price signals can also facilitate decentralized two-sided matching. In certain cases, this leads to greater welfare than use of even "perfect" filters. [1] [2]

*tloder@umich.edu
†mva@bu.edu, marshall@mit.edu
‡rwash@umich.edu

[1]This material is based upon work supported by the National Science Foundation under Grant No. 0114368.
[2]*Appearing in the B.E. Journals: Advances in Economic Analysis & Policy*

Contents

1 Introduction

Though the most well-known and timely example is email spam, the problem of unsolicited and unwanted contact is pervasive, affecting numerous media such as text and instant messaging, the telephone, and regular mail.[2] Information asymmetry combined with sender choices produces a form of message "pollution." Not bearing the negative externalities of mistargeting, senders rationally bombard recipients to gain their attention. If senders knew which recipients would like their messages, they could target more successfully. But before first-contact, recipients do not know sender qualities, and senders do not know recipient tastes.

Herein, we describe and compare alternative mechanisms for raising the value of first-contact communication. The mechanisms vary in their effects, largely due to differences among several characteristics: whether the mechanism is decentralized or imposed centrally, how it permits negotiation, whether it supports side payments, whether it can distinguish first from subsequent contact, what information it provides decision-makers, and when such information is revealed.

In addition to open access, our analysis considers three idealized mechanisms: (1) a costless "perfect filter" that exhibits neither false positives nor false negatives, (2) an efficient Pigouvian tax that redistributes all proceeds, and (3) a new proposal, an "Attention Bond Mechanism" (ABM) that allows a recipient to mandate a sender risk prior to accepting a message and permits side-payments.

With respect to filters, we ask 'what is the best that a filter can do?'. For the Pigouvian tax, assuming information on sender quality and recipient tastes is accessible, a central planner could force senders to internalize costs they impose on

[2]The communications media targeted by our analysis can most intuitively be described as *multi-party point-to-point media with cheap identities and imperfect information.* We consider the following five properties to be important:

Multi-Party Matching There are multiple participant types. Variation in participant utilities implies that senders prefer to reach some recipient types more than others. Likewise, recipients prefer to hear from some sender types more than others.

Point-to-Point Each point of entry and exit has a unique identifier. This implies that both broadcast and single messages can be treated as pairwise interaction.

Cheap Identities A sender can generate new identities at negligible cost. In contrast, the cost of acquiring an identity in use by another sender is presumed to be prohibitively high.

Low Transmission Costs Sending a message costs very little with open access, particularly when compared to the cost of correlating preferences with an identity.

Imperfect Information Senders do not know recipient types, unless revealed. Recipients do not know sender types, unless revealed.

others. Building on this insight, the third mechanism applies the Coase Theorem and grants property rights in attention. The application is somewhat more complex than in other situations; negotiation requires communication and unwanted communication is itself the problem. Thus we add signaling and screening, as a means to separate high quality from low, minimizing the need to negotiate.

Comparing these three mechanisms leads to several observations:

- It is advantageous to shift focus from the information in a message to the information known to the sender. Information revelation mechanisms can then be used to force people who knowingly misuse communication to incur higher costs than those who do not. Regardless of mechanism, higher costs encourage targeting and discourage low value messages.

- Filters and taxes make insufficient use of private information. The centralized Pigouvian tax affects all communication, first and subsequent, leading to broader distortion. The perfect filter grants recipients a message level veto but provides no preference information to the market. In contrast, giving recipients rights in their own attention motivates them to signal their preferences in ways that facilitate efficient targeting.

- Mechanisms designed to promote valuable communication can outperform those designed merely to block wasteful communication, especially given side payments. Conditions exist for which subsidizing recipients causes them to read valuable messages they would have discarded. Reversing these subsidies can cause senders to generate valuable messages they would never send.

- Although private knowledge of message content might favor senders over recipients, this advantage can disappear under a take-it-or-leave-it offer by recipients to reject non-conforming communications.

- The prospect of side payments comes at a cost. Recipients can strategically misrepresent their preferences in order to capture sender surplus.

The remainder of this paper proceeds as follows. Section 2 reviews previous work in this area and states our assumptions on applicable media. Section 3 develops a model of first-contact communications and distills three theoretical mechanisms for dealing with it. Section 4 provides an initial comparison of the mechanisms and shows how screening and wealth transfers affect welfare. Section 5 adds sender heterogeneity. Section 6 adds recipient heterogeneity and focuses on a flat tax. In Section 7 we address first vs subsequent contact and discuss why a recipient would not always seize a bond. Finally, Section 8 concludes.

2 Prior Literature

Approaches for dealing with unwanted communications fall roughly into three categories: technological, legal, and market-based. Throughout this review, *ex ante* intervention refers to the period before a recipient learns message content.

2.1 Technological Solutions

Technological solutions generally classify senders or classify content, and these techniques may be combined (Cranor and LaMacchia, 1998). Under blacklists, all messages are accepted unless the sender's identity appears on the list. Whitelists work in reverse; all messages are rejected unless the sender's identity appears on the list (Cranor and LaMacchia, 1998). But if senders can obtain cheap new identities, blacklists are generally ineffective; unscrupulous senders can simply take on new identities. Friedman and Resnick (2001) show that newcomers must "pay their dues" in any open society — one that does not charge per access — for precisely this reason. Similarly, if senders can easily "spoof" i.e. forge identities of others, then whitelists are generally ineffective. Whitelists alone have another drawback: without a way around a whitelist, senders are forced outside the medium, often at higher cost, to request permission to send.

Granted, various forms of authentication and digital signatures reduce "spoofing" by making identities "strong" (Tompkins and Handley, 2003). Multiple firms and standards bodies are working to ensure such technologies are prevalent. Strong identity will inevitably become part of any realistic solution, but is not in itself a complete solution.

With content-based filtering, the message itself provides clues used for blocking. Rule-based and Bayesian filters use deductive and probabilistic inference to classify spam *ex ante*. "Community filters" harvest the classification efforts – "flagging" – of a few participants *ex post* to remove spam from the inboxes of other participants *ex ante*. The main problem with content-based filters is classification error: "false positives" (good messages incorrectly classified as spam) and "false negatives" (unwanted messages that are accepted).

¿From a technology perspective, semantic content analysis has proven difficult. Computers barely understand natural language without intentional distortion, and the space of identifiable message permutations on single words is much larger than generally realized. For example, using only one-letter substitutions, misspellings, and spacing, one estimate[3] gives more than 6×10^{20} variations for a single six-letter

[3] `http://cockeyed.com/lessons/viagra/viagra.html` examples include v1agra, vi@gra, v:i:a:g:r:a, viagorea, viatgra, V-!.a-g*r-a, via6ra, and ViagrYa, just to name a few.

word, the anti-impotence drug Viagra. An adversary can also use filters themselves as a means to identify seemingly desirable messages, which can then be used for filter penetration (Graham-Cumming, 2004). More intuitively, the expressiveness of natural language lets senders use creative metaphors to carry multiple meanings in single statements.[4] Language plasticity permits an escalating arms race in which one side seeks better ways to block unwanted access and the other seeks better ways to gain it. Internet service providers report that spammers have responded to changes in filtering technology in as few as two hours (Libbey, 2004).

2.2 Regulatory Solutions

Various laws have been proposed or enacted to tax communication, ban inappropriate content (backed by fines or criminal penalties), force identification tags, or create marketing opt-out lists. 2004's CAN-SPAM Act[5] requires senders to provide valid subject lines, legitimate return addresses, and adult content labels. Initial results appear poor (Rainie and Fallows, 2004). In contrast, the 1991 Telephone Consumer Protection Act established national do-not-call and do-not-fax registries. By August 2005, these registries contained over 100 million phone numbers.

If one identifies "information overload" as overexploitation of scarce recipient attention, the right tax, even one that destroys tax revenues, can enhance welfare by forcing senders to "target" (self-limit) their messages (van Zandt, 2004). Shortcomings of taxes include jurisdiction, enforceability, and distortion if applied selectively. Taxing one type of packet, e.g. SMTP, can cause shifts to other types, e.g. HTTP.

Enforcement of laws concerning email and VoIP calls, for example, has proven challenging. Senders illegally acquire access to computers and turn them into "spam zombies", hiding their identity. Alternatively or in combination, they send messages (initiate calls) from across jurisdictional boundaries. Neither packet-switched medium currently supports authentication or audit trails as effectively as the circuit-switched and more regulated POTS (telephone). Problems of accurately tracing senders and packets have been so great that one legal scholar proposed a federally supported bounty on criminal spammers [cf. Lessig in (Bazeley, 2003)].

[4]Being 'tone deaf' to metaphor, filters passed one coauthor an invitation to "get a rod like a firehose."

[5]Controlling the Assault of Non-Solicited Pornography And Marketing, Public Law No. 108-187, passed Jan 1, 2004

2.3 Market Solutions

Given that recipient attention is a scarce resource, others propose means of pricing it. Proposals include transferable stamps, challenge-response, selling "interrupt rights," and using auctions.[6]

An experimental investigation of pricing recipient attention via email stamps found that charging does cause senders to be more selective and to send fewer messages (Kraut et al., 2003). In particular, variable rate usage charges reduced communication more than flat rate access charges. Interestingly, recipients did not see postage as a signal of value and the authors conclude that such systems show great promise but "need more work "(p. 206).

Focusing on call externalities, Hermalin and Katz (2004) consider pricing messages to senders and receivers, and find reason to charge both. Efficiency gains can then lead, in certain limited circumstances, to first-best welfare.

Challenge-Response (CR) systems achieve a crude form of negotiation and pricing. All use whitelists to limit costs to first-contact. Under CR, a recipient's proxy system temporarily escrows messages from unrecognized senders and directs those senders to pass a test in the form of a challenge or CAPTCHA.[7]. Computational challenges require proof of calculation effort, and consume computer rather than human time (Dwork and Naor, 1993).[8] If a sender passes, the system delivers the escrowed message and whitelists the sender. Otherwise, it destroys the message.

Both forms of CR resemble a tax on senders with certain differences. With local control, recipients can adjust the tax to match their opportunity costs. But, CR systems do not collect transferable utility nor facilitate side payments. Computational challenges are susceptible to attacks from harnessed "zombie" machines that expand the resources of illegitimate senders (Laurie and Clayton, 2004).

A more attractive mechanism for allocating attention, "selling interrupt rights," is outlined in Fahlman (2002), and presented more casually in Ayres and Nalebuff (2003). A protocol for implementing CR with side payments appears in Krishnamurthy (2004). For pricing valuable attention when recipients are boundedly rational, van Zandt (2004) points to efficient use of a Vickrey auction.

[6](Kraut et al., 2003; Dwork and Naor, 1993; Fahlman, 2002; Hermalin and Katz, 2004; van Zandt, 2004)

[7]Completely Automated Public Turing Test to Tell Computers and Humans Apart (von Ahn et al., 2003) CAPTCHAs are human cognition tests, such as pattern recognition

[8]CAPTCHAs themselves are frequently inverted. Senders embed messages in images that only a human can recognize, since recipients' proxy systems cannot themselves pass cognition tests.

3 Modeling the Mechanisms

To compare mechanisms, we start with a model of relevant media, and determine participant surplus and welfare under Open Access. We then explore differences in information, choices, and payoffs for each party based on interventions modeled as filters, taxes, and bonds.

Let N participants have access to a communications medium. Any participant i can send a message to any other participant j in the population; implying $N \times N$ possible sender-recipient exchanges. (For clarity, we index senders using i and male pronouns, and index recipients using j and female pronouns; since our analysis concerns first-contact information asymmetry among strangers, $i \neq j$.)

For each sender-recipient pair, the probability distribution of first-contact message values is $f_{i,j}(s, r)$, where s is sender and r is recipient value to i and j respectively, but only if j actually reads i's message. Otherwise, $s = r = 0$. The values of s and r are private information.

We define the rectangular region in the $\langle s, r \rangle$-plane, where $f_{i,j}(s, r)$ is positive, as the region bounded by $\underline{r}_{i,j}, \overline{r}_{i,j}, \underline{s}_{i,j}, \overline{s}_{i,j}$ (See Figure 1). $\underline{r}$ is the supremum of r such that $\forall r < \underline{r}$, $f_{i,j}(s, r) = 0$. The other bounds are similarly defined. We assume f is fixed for each i, j pair.

Sending a message costs the sender $c_s \geq 0$ regardless of whether the recipient actually reads it. If a sent message is read, the recipient bears cost $c_r \geq 0$. Expected per-message surplus SS for the sender and RS for the recipient result from integrating net message value over f. Both senders and recipients prefer positive surplus. Since senders do know s but recipients do not know r prior to receiving, observed values for SS will be non-negative but those for RS need not be.

Under open access, the choice of whether or not to send individual messages favors senders, who have *interim* rationality. They exercise transactions-level choice. Open access only grants recipients *ex ante* rationality. They exercise market-level choice to participate.

Figure 1-a represents possible message values for one sender-recipient pair (all i, j's have been removed to reduce clutter). The upper right Quadrant, labeled "i", origin at (c_s, c_r), shows where both SS and RS are positive and so lies above $r = c_r$ (horizontal dashed line) and to the right of $s = c_s$ (vertical dashed line). Messages in Quadrant i are always efficient. The upper left Quadrant ii represents messages the recipient would like to get, but senders choose not to offer without side payments. We can assume such messages are costly to produce or have negative consequences for senders. In contrast, the lower right Quadrant iv represents messages such as leaflets that the sender wants to distribute but recipients consider waste. These may represent efforts to persuade, offend, poll, or defraud. Quadrant iii exhibits negative surplus to both parties and would either be counter-factual or sent

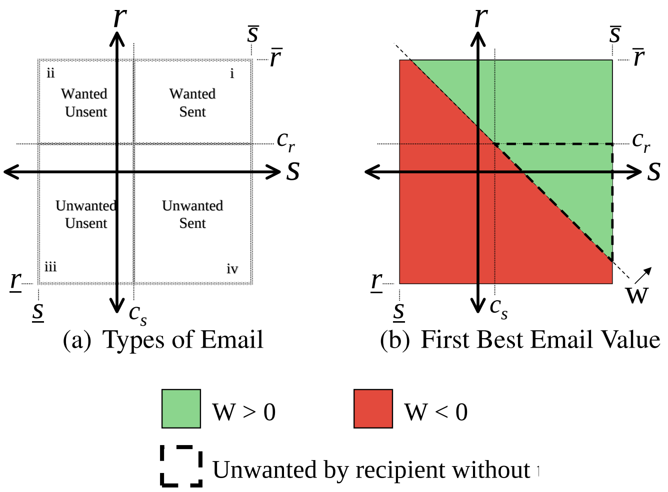


Figure 1: Distribution of Message Values

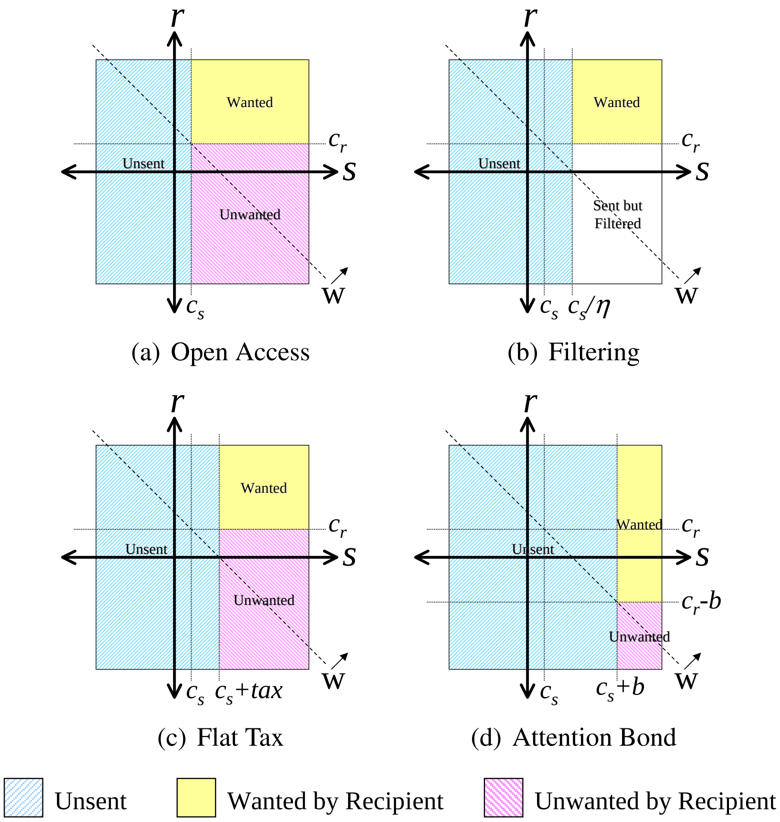


Figure 2: Comparison of mechanisms based on the distribution of message values.

by accident. Example messages for each Quadrant appear in Table 1.

ii	*Wanted / Unsent*	i	*Wanted / Sent*
	Custom news & Sales leads Credit scores Subscription Newsletter Analyst reports, Expert advice Personalized loan application		Sale notice for desired products Long lost friend, New colleague Favorable press inquiry News of weddings & reunions New project opportunities
iii	*Unwanted / Unsent*	iv	*Unwanted / Sent*
	Offensive news sent to wrong person Embarrassing personal revelation Accidentally telling boss *real* opinion		Unfocused ads, Propaganda campaigns Polling & Market research Phishing attacks, Viruses Offensive pornography, Hate mail

Table 1: Examples illustrate each region of message value. Quadrant order parallels that of Figure 1-a

A message's total contribution to welfare is the sum of sender and recipient surplus as shown by the diagonal line 'W' in Figure 1-b. Messages with positive net welfare lie to the northeast of this line (in green) and are desirable from the standpoint of a social planner. Messages to the southwest of this line (in red) are socially undesirable, even if recipients or senders individually find them desirable.

Note that the triangular region below c_r but above 'W' represents messages that are socially efficient, but which recipients prefer not to receive without side-payments. Filters convert this region to deadweight loss. An analogous region exists to the left of c_s above 'W' for messages that are costly to senders but attractive to recipients. These unsent messages are also deadweight loss. To the extent that a mechanism can facilitate transactions in these regions, welfare is improved. As we show below, certain mechanisms fare better than others at recovering surplus. Still, information asymmetry makes recovering first-best welfare difficult. With these assumptions and definitions, we turn now to analysis of how various interventions affect the open access baseline.

3.1 Open Access

Under open access to a medium, all sent messages are received. The sender's only cost is his marginal per-message cost c_s, and so he sends all messages where $s > c_s$. Thus the region of interest to senders is given by $c_s \leq s \leq \overline{s}_{i,j}$ and that of interest to recipients by $\underline{r}_{i,j} \leq r \leq \overline{r}_{i,j}$. Expected surplus and welfare are:

$$\text{SS}_0 = \int_{\underline{r}}^{\overline{r}} \int_{c_s}^{\overline{s}} f_{i,j}(s,r)\,(s - c_s)\ ds\ dr \tag{1}$$

$$\text{RS}_0 = \int_{\underline{r}}^{\overline{r}} \int_{c_s}^{\overline{s}} f_{i,j}(s,r)\,(r - c_r)\ ds\ dr \tag{2}$$

$$\text{W}_0 = \int_{\underline{r}}^{\overline{r}} \int_{c_s}^{\overline{s}} f_{i,j}(s,r)\,[(s - c_s) + (r - c_r)]\ ds\ dr \tag{3}$$

Figure 2-a shows the region of message values sent under Open Access, namely those messages to the right of c_s. But, only those above c_r are actually wanted by the recipient. These represent both Quadrants i and iv in Table 1. Taking the welfare line 'W' from Figure 1 and superimposing it over Figure 2-a, shows that part of the region labeled 'Unwanted' lies to the northeast of 'W' and is therefore welfare positive, but much of it is not.

3.2 Flat Tax

Under a flat tax, a sender pays amount t to a central authority for each message sent, and all sent messages are received. The central authority knows only the distribution of message values $f_{ij}(s,r)$ but not the private information associated with a specific message. Thus t must be decided globally. For now, taxes raise sender costs to $c_s + t$ but, to strengthen the case for taxes relative to the ABM, we rebate proceeds to society to avoid welfare losses. Expected welfare is thus:

$$\text{SS}_t = \int_{\underline{r}}^{\overline{r}} \int_{c_s+t}^{\overline{s}} f_{i,j}(s,r)\,(s - (c_s + t))\ ds\ dr \tag{4}$$

$$\text{RS}_t = \int_{\underline{r}}^{\overline{r}} \int_{c_s+t}^{\overline{s}} f_{i,j}(s,r)\,(r - c_r)\ ds\ dr \tag{5}$$

$$\text{W}_t = \int_{\underline{r}}^{\overline{r}} \int_{c_s+t}^{\overline{s}} f_{i,j}(s,r)\,[(s - c_s - t) + (r - c_r) + t]\ ds\ dr \tag{6}$$

Figure 2-c shows the messages sent under a flat tax. A positive tax has the effect of shifting the sender's decision threshold to the right, from the vertical line c_s under open access, to $c_s + t$ under the tax, thus roughly and globally correcting recipient's negative call externality.

The true effect of a flat tax, however, depends on the disposition of proceeds and the distribution of agent types. In the case of homogeneous senders and receivers, if t is returned to senders, it has no effect on sender behavior. If t is destroyed, total welfare is Pareto inferior to that under a transfer to recipients. But, if t is transfered to recipients, it represents a special case of the recipient-chosen ABM; thus we defer tax analysis to Section 4. Introducing agent heterogeneity produces novel tax results, which we analyze in Section 6.

3.3 Perfect Filter

The perfect filter is an idealized pure classification technology. Using perfect knowledge of recipient preferences, it blocks undesirable messages for which $r < c_r$. The perfect filter is costless to operate (free classification and free disposal) and makes no mistakes (neither false positives nor false negatives). It therefore discards all messages in Quadrant iv of Figure 1-a, removing negative recipient surplus in the range of $\underline{r} \leq r < c_r$ while preserving positive surplus $c_r \leq r \leq \overline{r}$.

The filter does not disclose recipient private information r, but by causing certain message to go undelivered, it alters the sender's decision to send. If $\eta \in [0, 1]$ represents the fraction of messages permitted to pass, then sender surplus is positive when $\eta s \geq c_s$. This implies a higher cost basis such that senders only send more valuable messages $s \geq \frac{c_s}{\eta}$. Together, these new constraints on delivered messages redefine surplus terms as follows:

$$\mathrm{SS}_{PF} = \int_{c_r}^{\overline{r}} \int_{\frac{c_s}{\eta}}^{\overline{s}} f_{i,j}(s,r)\,(s - c_s)\ ds\ dr + \int_{\underline{r}}^{c_r} \int_{\frac{c_s}{\eta}}^{\overline{s}} f_{i,j}(s,r)\,(-c_s)\ ds\ dr \tag{7}$$

$$\mathrm{RS}_{PF} = \int_{c_r}^{\overline{r}} \int_{\frac{c_s}{\eta}}^{\overline{s}} f_{i,j}(s,r)\,(r - c_r)\ ds\ dr \tag{8}$$

$$\mathrm{W}_{PF} = \mathrm{SS}_{PF} + \mathrm{RS}_{PF} \tag{9}$$

Figure 2-c shows message values realized under the Perfect Filter mechanism. The filter accepts messages only where $r \geq c_r$, indicated by the yellow region to the upper right. The vertical dashed line labeled $\frac{c_s}{\eta}$ bounds this region on the left and represents the minimum s that will increase net sender surplus. Note that in comparison to Open Access (in Figure 2-a), the "Wanted" region is smaller; recipients lose certain valuable messages since the sender never sends them. Further, in the clear "sent but filtered region" (below $r = c_r$), the sender incurs cost c_s but receives no value ($s = 0$). This region represents pure social loss.

3.4 Attention Bond Mechanism

Under the ABM, a recipient posts a take-it-or-leave-it offer in the form of a bond "price" ϕ, which is public information. Recipients must choose ϕ *ex ante*. Each sender must then post a bond with value $\geq \phi$ for each first-contact message he wants delivered. For any bond smaller than ϕ or no bond at all, a sent message is destroyed; incurring c_s without a gain of s. Since ϕ is public information, the sender never makes this mistake and we assume for now that he posts exactly ϕ.

All bonded messages are delivered creating immediate surplus $s - c_s$ and $r - c_r$ for senders and receivers respectively. Upon reading the message, however, the recipient may, at her sole discretion, seize the bond or release it. Seizing the bond effects a wealth transfer, increasing her surplus and reducing that of the sender by ϕ.

In expectation, the sender's realized surplus also depends on the probability of forfeiting bonds. Thus we define b as $b = p\phi$, with $0 \leq p \leq 1$, and interpret p as the seize probability. If the recipient always seizes the bond ($p = 1$), then $b = \phi$. For now we assume $p = 1$ and explore the effect of different seize policies in Sections 5 and 7.2. Side payments then alter expected welfare as follows:

$$\mathrm{SS}_b = \int_{\underline{r}}^{\overline{r}} \int_{c_s+b}^{\overline{s}} f_{i,j}(s,r)\,(s - c_s - b)\;ds\;dr \tag{10}$$

$$\mathrm{RS}_b = \int_{\underline{r}}^{\overline{r}} \int_{c_s+b}^{\overline{s}} f_{i,j}(s,r)\,(r - c_r + b)\;ds\;dr \tag{11}$$

$$\mathrm{W}_b = \int_{\underline{r}}^{\overline{r}} \int_{c_s+b}^{\overline{s}} f_{i,j}(s,r)\,(s - c_s + r - c_r)\;ds\;dr \tag{12}$$

Figure 2-d depicts the transactions realized under the ABM. As with taxes, the send threshold shifts right, in comparison to Open Access, to $s = c_s + b$. But unlike Open Access, taxes, and filters, the ABM's side payment b shifts the lower bound on the "Wanted" region below c_r. This changes the ratio of "Wanted" to "Unwanted" messages.

4 Homogeneous Senders and Recipients

In this section, we derive results using two simplifying assumptions: (1) participant types are homogeneous, and (2) message values are heterogeneous and uniformly distributed. We begin with a single i, j relationship and so drop subscript notation. The second assumption lets us replace $f_{i,j}(s,r)$ in the equations from the previous section with the constant $k = \frac{1}{(\overline{r}-\underline{r})(\overline{s}-\underline{s})}$. This can be loosely interpreted as the result of s and r being uncorrelated due to private information before agents have had a chance to communicate. It also greatly simplifies integral expressions and allows η to be endogenized as $\frac{\overline{r}-c_r}{\overline{r}-\underline{r}}$. While these assumptions do not match exact conditions in real world media, they allow us to derive insights into the functioning of each idealized mechanism. We examine correlation in later sections.

4.1 ABM vs. Perfect Filter

The recipient's first choice is to calculate her optimal bond.

Lemma 1 *The recipient's optimal bond* b^+*is*

$$b^+ = \frac{1}{2}\left((\overline{s} - c_s) - \left(\frac{\overline{r}+\underline{r}}{2} - c_r\right)\right)$$

Proof:
Evaluate the right hand side of Equation (11) to produce

$$E[RS] = k \cdot \frac{\overline{s} - c_s - b}{\overline{s} - \underline{s}}\left(\frac{\overline{r}+\underline{r}}{2} - c_r + b\right) \tag{13}$$

Apply first order conditions to Equation (13) with respect to b, noting that the second derivative $-2\left(\frac{1}{\overline{s}-\underline{s}}\right)$ ensures b^+ is a maximum. ■

Bond b^+represents the monopoly price for recipient attention (competition is briefly explored in Section 7.2). The optimal bond in this case is the expected sender surplus minus half the expected recipient surplus. If kept, the bond represents an individually targeted Pigouvian tax on senders, adjusted for the recipients' own message surplus. Recipients have internalized a sender's expected call externality.

Bond sign also has useful implications. Let sender and receiver expected values be $\tilde{s} \equiv \frac{\overline{s}-c_s}{2}$ and $\tilde{r} \equiv \frac{\overline{r}+\underline{r}}{2} - c_r$ with the difference explained by *interim* versus *ex ante* rationality. A positive bond (meaning $\tilde{s} > \frac{\tilde{r}}{2}$) implies low expected value, consistent with wasteful first-contact. A negative bond implies subsidy. A recipient could share her surplus if first-contact were valuable instead of wasteful.

Recipient surplus under the ABM is always at least that of Open Access, $\mathrm{RS}_{b^+} \geq \mathrm{RS}_0$. Since $b^+ = 0$ reproduces the baseline case, the added degree of freedom ensures the left hand side weakly dominates the right. More importantly, the ABM can actually perform better than even a perfect filter.

Proposition 1 *The ABM creates greater recipient surplus than the Perfect Filter,* $\mathrm{RS}_{b+} \geq \mathrm{RS}_{PF}$*, if and only if the optimally chosen bond is such that*

$$(b^{+})^2 \geq \frac{1}{2}\eta\left(\overline{s} - \frac{c_s}{\eta}\right)(\overline{r} - c_r)$$

Proof:

Apply Lemma 1 to the inequality $\mathrm{RS}_{b+} \geq \mathrm{RS}_{PF}$, let $\eta = \frac{\overline{r}-c_r}{\overline{r}-\underline{r}}$, and use Equations 11 and 8 to produce the following expression:

$$\frac{(\overline{r} - \underline{r})\left(\overline{s} - c_s + \left(\frac{\overline{r}+\underline{r}}{2} - c_r\right)\right)^2}{(\overline{r} - c_r)\left(\overline{r}\left(\overline{s} - c_s\right) - \overline{s} \cdot c_r + \underline{r} \cdot c_s\right)} \geq 2 \tag{14}$$

Rearranging terms and substituting for η in the test condition produces an identical expression. The sequence of steps connecting test and result can then be traversed in either direction. ■

The test condition is the product of surplus terms for the sender and recipient, where the left hand side of the inequality represents the bonded surplus, and the right hand side is the perfect filter surplus. Recipient surplus for the ABM is an average value reflecting the difference in IR participation constraints — senders participate for any message value above $c_s + b$, and recipients participate for average value above c_r. In contrast, the perfect filter gives perfect veto power to both parties, as both reject all messages below their private costs. That there exist values for which Equation 14 is true illustrates the point that facilitating exchange can dominate giving veto power.

While Equation 14 is not always true, it is in several cases. The test condition is more easily satisfied as recipients block more messages ($\overline{r} \to c_r$) and as senders have more surplus to transfer ($\frac{d(\mathrm{SS})}{d\overline{s}} = 2\left[(\overline{s} - c_s) + \left(\frac{\overline{r}+\underline{r}}{2} - c_r\right)\right] - 2\eta\left(\overline{r} - c_r\right)$ grows quickly as $\overline{s}$ grows). For similar reasons, it is also true for extreme values of $\overline{r}$ and $\underline{r}$. In contrast, the test fails when sender and receiver surplus have the same magnitude but opposite signs. For a single distribution, the IR constraint may bind the party with negative surplus. Long-lost friends and telemarketers may draw from different distributions, however, implying participation if the sum across distributions is positive. We explore the screening of different sender types in Section 5.

The ABM can beat the perfect filter because it supports wealth transfers that allow it to recapture the welfare positive region below c_r in Figure 1. Senders can pay recipients to read their messages. Explained graphically, for the ABM to succeed, the sum of welfare-positive messages lost from the "Wanted" region of Figure 2-c plus welfare from the negative welfare region enabled by the ABM below $c_r - b$ in Figure 2-d must be less than the sum of positive welfare contribution of messages transacted under the ABM shown in Figure 2d (bounded by $r = c_r$, $s = c_s + b$, W and $\overline{s}$) plus the value of the reclaimed deadweight-loss of the "Sent but Filtered" region in Figure 2-c that occurs under the Perfect Filter. In other words, geometric gains must outweigh geometric losses.

4.2 Social Welfare under the ABM and Perfect Filter

Having considered recipient benefit, the question remains whether the ABM also improves total social welfare.

Proposition 2 *Assuming heterogeneous and uniformly distributed values, the total social welfare with a recipient-chosen bond is greater than that of a perfect filter* $\mathrm{W}_{b+} > \mathrm{W}_{PF}$ *if and only if the following is true:*

$$\left[(\overline{s} - c_s) + \left(\frac{\overline{r} + \underline{r}}{2} - c_r\right)\right]^2 \geq \eta \left(\overline{s} - \frac{c_s}{\eta}\right)\left[(\overline{r} - c_r) + \left(\overline{s} - \frac{c_s}{\eta}\right)\right] \qquad (15)$$

Proof: To prove, construct the inequality $\mathrm{SS}_{b+} + \mathrm{RS}_{b+} \geq \mathrm{SS}_{PF} + \mathrm{RS}_{PF}$ using Equations 7, 8, 10, 11, and Lemma 1, and evaluate:

$$k \cdot \int_{\underline{r}}^{\overline{r}} \int_{c_s + b^+}^{\overline{s}} \big((r - c_r + b^+) + (s - c_s - b^+)\big)\, ds\, dr \geq$$
$$k \cdot \int_{c_r}^{\overline{r}} \int_{\frac{c_s}{\eta}}^{\overline{s}} ((r - c_r) + (s - c_s))\, ds\, dr - k \cdot \int_{\underline{r}}^{c_r} \int_{\frac{c_s}{\eta}}^{\overline{s}} (c_s)\, ds\, dr$$

Simplify and rearrange terms to produce the test condition. Beginning with the test condition, it is straightforward to reverse the sequence of inferences. ■

Note that each parenthesis term in the test condition represents the per-message surplus for either the sender or recipient. The left hand side gives ABM surplus terms; the right gives terms for the Perfect Filter. Similar to Equation 14, Equation 15 is not always true, but it is in many situations like those for Proposition 1.

An alternative method of choosing the bond is to maximize total social welfare. Thus, a social planner chooses the following bond:

Lemma 2 *The social welfare maximizing bond size is*

$$b^\star = \left(c_r - \frac{\overline{r} + \underline{r}}{2}\right)$$

Proof: Find $\mathrm{W}_{b^\star}$ using Equation 12 and FOC's with respect to b. ■

Note that $b^\star$ has the magnitude of expected recipient surplus, but is of the opposite sign. It is not a function of sender surplus at all. Since senders and recipients are homogeneous, $t = b^\star$ is the optimal Pigouvian tax. If the tax is negative, implying expected recipient surplus is positive, then this is a tax imposed on recipients and redistributed to the senders. Given the goal of minimizing wasteful communication, taxing recipients appears counterintuitive. The result, however, is highly socially efficient. Senders are generating positive surplus, so transferring surplus to them encourages this behavior.

In contrast, if the tax $t = b^\star$ is positive, then it is a tax imposed on senders. Senders know only the expected value of r but have full information on s. In contrast, recipients have full information on neither. Choosing this bond, or requiring this tax, causes the party with the best information to internalize the full consequences of each decision to send.

Knowing the welfare maximizing bond, we can ask when $b^\star$ produces greater social welfare than the perfect filter.

Corollary 1 *Assuming heterogeneous and uniformly distributed values, the social welfare with a flat tax* $\mathrm{t} = b^\star$ *is greater than that of a perfect filter,* $\mathrm{W}_{tax} > \mathrm{W}_{PF}$*, if and only if the following is true:*

$$\left[(\overline{s} - c_s) + \left(\frac{\overline{r} + \underline{r}}{2} - c_r\right)\right]^2 \geq \eta\left(\overline{s} - \frac{c_s}{\eta}\right)\left[(\overline{r} - c_r) + \left(\overline{s} - \frac{c_s}{\eta}\right)\right]$$

Proof: Apply the method of Proposition 2, replacing Lemma 1 with Lemma 2. ■

The social planner's optimum differs from the recipient-chosen optimum, given in Proposition 2, by a small constant. As a weaker test, this holds true more broadly. One unmodeled complication is that a Pigouvian tax, collected *ex ante*, can introduce moral hazard; if senders were to receive a sure subsidy, they might generate less valuable messages.

4.3 Reverse Bonds & Reverse Signaling

We have assumed that the recipient, uninformed about the specific value of any yet-to-arrive message, chooses the size of b. An alternative possibility is for the informed *sender* to signal his interest in a potentially valuable communication. If he does so, the choice of b is no longer optimal from the recipient's perspective, but merely chosen to avoid negative expected recipient surplus. That is, a sender minimizes b such that $E[\mathrm{RS}_{ABM}] \geq 0$, where $E[\mathrm{RS}_{ABM}]$ is given by Equation 11.

Integration of Equation 11 yields $k \cdot \frac{\overline{r}-c_s-b}{\overline{s}-\underline{s}}\left(\frac{\overline{r}+\underline{r}}{2}-c_r+b\right)$ for which the roots are $b \in \left\{\overline{s}-c_s, -\frac{\overline{r}+\underline{r}}{2}+c_r\right\}$. Thus, the expected signal is either the maximum surplus to the sender or the expected value to the recipient, whichever is less. Given that the situation is of first-contact, where the recipient can predict only the expected value of received messages, one might expect a reverse bond of $\frac{\overline{r}+\underline{r}}{2}-c_r<0$ so that recipient expected value is ≥ 0.

This analysis extends further to bilateral initiative. The ABM permits reverse signaling from a recipient who could *solicit* messages of high net value to herself but low net value to senders — messages that would otherwise have gone unsent. Such messages were previously described and inhabit region *ii* of Figure 1-a.

Recovering deadweight loss leads to the following proposition:

Proposition 3 *The option to post a negative bond can increase recipient surplus and also strictly increases total welfare.*

Proof:

Similar to the previous case, the surplus-maximizing recipient-chosen signal is the minimum b such that $E[\mathrm{SS}_{ABM}] \geq 0$, where $E[\mathrm{SS}_{ABM}]$ is given by

$$k \cdot \int_{c_r+b}^{\overline{r}} \int_{\underline{s}}^{c_s} (s-c_s+b)\, ds\, dr$$

Integration yields $\frac{(\overline{s}-c_s)}{(\overline{s}-\underline{s})(\overline{r}-\underline{r})}(c_r-\overline{r}+b)\left(\frac{\underline{s}-c_s}{2}+b\right)\geq 0$ implying that $b \in \left\{\overline{r}-c_r, \frac{c_s-\underline{s}}{2}\right\}$. Assuming a sender will keep the bond, a recipient fronts an amount equal to her own surplus or the sender's average losses, whichever is less. This leaves senders at least as well off as under the Perfect Filter. ■

For any region in which $r > \frac{c_s-\underline{s}}{2}$, new communications take place under the Attention Bond Mechanism that would never have occurred with the perfect filter or open access. In graphical terms, this recaptures the region of uncaptured surplus in Figure 3, leading to a strict Pareto improvement.

This section yields two observations. First, we again see the value of wealth transfers. The ABM permits the rational exchange of socially valuable messages that are otherwise lost. With reverse signaling, welfare can come from messages a sender normally would not send but the recipient welcomes.

Second, recipient greed under the ABM can cause her to price too high, relative to the social optimum, displacing welfare positive transactions. In this respect, a government-issued tax would be better. A central tax, however, has other issues highlighted in Section 6.

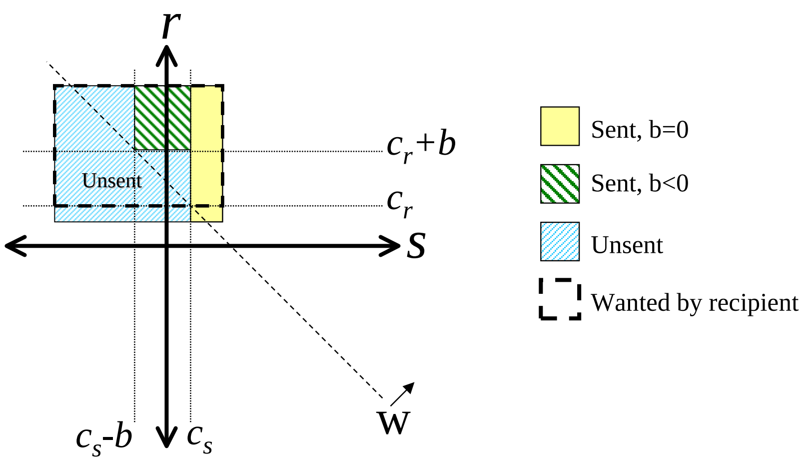


Figure 3: A negative bond permits recipients to buy valuable messages.

5 Heterogeneous Senders

Up to this point, we have studied only homogeneous senders and recipients. As lost acquaintances and spammers may represent different senders, we extend this model to handle multiple sender types. Let value distributions $V_1, \ldots, V_n$, have bounds $\overline{s}_i, \underline{s}_i, \overline{r}_i, \underline{r}_i$ with messages from sender type i being drawn from V_i. We assume also that a sender knows his type. Further, a recipient can process N message types. Otherwise, a recipient who can process only $M < N$ might efficiently use a Vickrey auction (van Zandt, 2004).

A recipient cannot know the type of a specific message *ex ante*. While she does know the relative likelihood α_i of an email coming from a given type i ($\sum_{i=1}^{n} \alpha_i = 1$), she can only choose a single bond size ϕ for use across all messages.

After reading a message, a recipient can choose to seize or release a bond. Let her decision represent a "seize policy." (Justification for different policies based on signaling recipient type is explored in Section 7.2.) With multiple messages, she may associate a different seize probability p_i with each sender type. So *ex ante*, bonds ϕ must be the same for all senders, but *ex post* p_i may differ by type. For multiple senders, a recipient learns the sender type from reading the message and can apply different bond seize policies $(p_1, \ldots, p_n)$ specific to the inconvenience of that sender type. The expected side payment is thus $b_i = p_i \cdot \phi$. Senders' total expected surplus is given by

$$\mathrm{SS}_{\phi,p_1,\ldots,p_n} = \sum_{i=1}^{n} \alpha_i k_i \cdot \int_{\underline{r}_i}^{\overline{r}_i} \int_{c_s+p_i\phi}^{\overline{s}_i} (s - c_s - p_i\phi)\, ds\, dr \tag{16}$$

Likewise, the total expected recipient surplus is

$$\mathrm{RS}_{\phi,p_1,\ldots,p_n} = \sum_{i=1}^{n} \alpha_i k_i \cdot \int_{\underline{r}_i}^{\overline{r}_i} \int_{c_s+p_i\phi}^{\overline{s}_i} (r - c_r + p_i\phi)\, ds\, dr \tag{17}$$

Multiple sender types alter the recipient optimal bond as follows.

Lemma 3 *Given n sender types with value distributions $V_1, \ldots, V_n$, the recipient-chosen optimal bond $\phi^+$is*

$$\phi_n^+ = \frac{\sum_{i=1}^{n} \frac{\alpha_i}{\overline{s}_i - \underline{s}_i} \cdot p_i \cdot \left(\overline{s}_i - c_s - \left(\frac{\overline{r}_i + \underline{r}_i}{2} - c_r\right)\right)}{\sum_{i=1}^{n} \frac{\alpha_i}{\overline{s}_i - \underline{s}_i} p_i^2}$$

Proof: Total recipient surplus is given by Equation (17). Apply first order conditions with respect to ϕ and solve to yield ϕ_n^+. ■

Note that this is a weighted average of bonds from the homogeneous case.

Lemma 4 *For any sender type i, given a bond ϕ, the optimal policy p_i^+ is*

$$p_i^+ = \frac{1}{2\phi}\left(\overline{s}_i - c_s - \left(\frac{\overline{r}_i + \underline{r}_i}{2} - c_r\right)\right)$$

Proof: We take $\mathrm{RS}_{\phi,p_1,\ldots,p_n}$ and calculate the first derivative with respect to p_i. This is simple, as p_i only appears in one term, so the remaining terms are irrelevant. Solving for p_i yields the above equation for p_i^+. ■

With the optimal policy and bond for any set of distributions (from Lemma 4), we can compare expected message costs across multiple sender types. As one might suspect, the greater the inconvenience to the recipient, the higher the cost to the sender.

For comparison, consider the special case of two distributions, the "good" communications $\mathbb{G}$ and the "bad" communications $\mathbb{B}$. Qualitatively, the difference between these two distributions is that the expected value of communications to the recipient for $\mathbb{G}$ is significantly higher than for $\mathbb{B}$: $\frac{\overline{r}_G + \underline{r}_G}{2} > \frac{\overline{r}_B + \underline{r}_B}{2}$.

Proposition 4 *Assuming heterogeneous and uniformly distributed values for each sender type, if the difference in mean recipient values exceeds the difference in maximum sender surplus, there exists a separating equilibrium in which less desirable messages incur higher sender costs. Specifically,*

$$E[r_G] - E[r_B] > \overline{s}_G - \overline{s}_B$$

if and only if $c_s + p_G^+\phi^+ < c_s + p_B^+\phi^+$

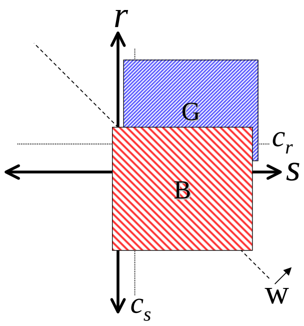
Proof: Consider the reverse direction. Apply Lemmas 3 and 4 to the total costs for each type $c_s + p_G^+\phi^+ < c_s + p_B^+\phi^+$ to produce

$$\left(\overline{s}_G - c_s - \left(\frac{\overline{r}_G + \underline{r}_G}{2} - c_r\right)\right) < \left(\overline{s}_B - c_s - \left(\frac{\overline{r}_B + \underline{r}_B}{2} - c_r\right)\right) \tag{18}$$

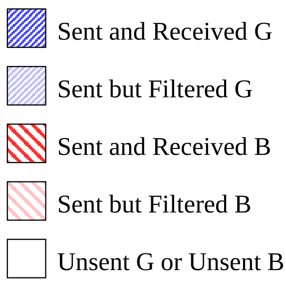
Rearrange to produce the claim

$$\frac{\overline{r}_G + \underline{r}_G}{2} - \frac{\overline{r}_B + \underline{r}_B}{2} > \overline{s}_G - \overline{s}_B. \tag{19}$$

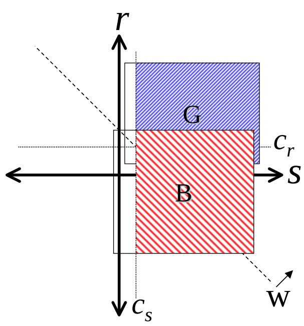
Alternatively, start with Equation 19 and rearrange to produce Equation 18. ■



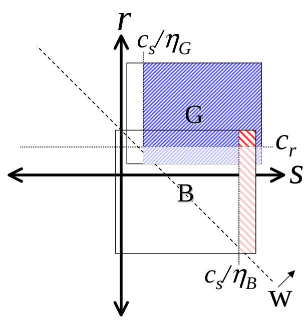
(a) Initial Distributions



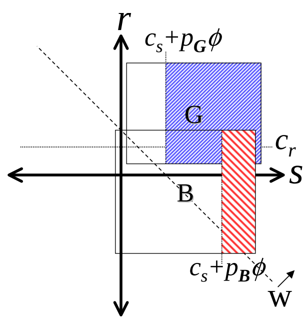
(b) Region Results



(c) Open Access



(d) Perfect Filter



(e) Attention Bond

Figure 4: Under screening, messages that are less valuable to recipients incur higher sending costs. Wealth transfers can also unblock welfare-positive communication.

It is easy to see that if maximum surplus is the same for both sender types, $\overline{s}_G = \overline{s}_B$, then the proposition holds trivially. The reason this is not always true is that self-interested recipients may seek to exploit high value senders, independent of their own average surplus.

The resulting Stiglitz (1975)-style screening helps resolve the adverse selection problem. Screening can happen only if costs differ by sender type. This is possible because *ex post* verification of type, which is known by senders *ex ante*, allows recipients to impose these differential costs via their collection policies. Justification for policy differences will become apparent after we introduce recipient heterogeneity.

Figure 4 graphically shows the effects of the various mechanisms applied to multiple distributions with high and low mean recipient values. The Attention Bond Mechanism (Figure 4-e) is differentially costly to senders whose extra bond costs are transferred as surplus to the recipient.

Note that, as we have modeled sender heterogeneity, it is impossible to totally order the set of sender types. Two types with different expected values for the recipient can still have the same cost if the senders surplus is sufficiently high.

This section shows how the ABM enables screening. Senders whose messages have less value to the recipient can expect higher costs. This also illustrates how expected bond size can adapt to individual recipient tastes. The *ex post* penalty that the recipient imposes on senders depends solely on their individual bond return policies. Additional reasons to release a bond appear in Section 7.

This section also shows that the strategy decision facing the recipient can be complicated. In this model, for n sender types, there are $n + 1$ decisions that the recipient has to make, all of which affect message traffic.

6 Heterogeneous Senders and Recipients

The technological complexity of the many-to-many ABM raises the question of how far a simple tax will go toward the social optimum. Section 4.2 shows that a socially optimal tax can yield greater social surplus than the ABM. Here, we analyze the effect of such a tax on communications when there are heterogeneous senders and recipients. We continue to focus on first-contact communications. This affords us simpler comparison. It also favors the apparent effectiveness of the tax, which would distort all communications, in contrast to the ABM, which distorts only first-contact.

As before, consider two types of senders, $\mathbb{G}$ and $\mathbb{B}$. If recipients either like or dislike messages from a given type, their preferences induce four possible recipient types. Universal recipients $\mathcal{U}$ value messages sent from both $\mathbb{G}$ and $\mathbb{B}$ senders. Type $\mathcal{G}$ recipients like only $\mathbb{G}$ messages; type $\mathcal{B}$ recipients like only $\mathbb{B}$ messages; and finally, there exist recipients who dislike all messages. Since participation is voluntary, recipients of the last type generally exit the market.[9]

[9] More generally, n sender types induce 2^n recipient types. An alternative model might consider a continuum of sender types then examine phase transitions. An assumed sort order, however, implies transitivity of recipient preferences that would often be violated when sorting over multiple

This creates expectations for three recipient types ($\mathcal{B}$, $\mathcal{G}$, $\mathcal{U}$) and two sender types ($\mathbb{B}$, $\mathbb{G}$). Let exogenous parameter γ represent the fraction of type $\mathbb{G}$ senders, with $(1-\gamma)$ representing the fraction of $\mathbb{B}$ senders. Among recipients, let ϵ_G be the fraction of $\mathcal{G}$ recipients and ϵ_B be the fraction $\mathcal{B}$ recipients. Thus $(1-\epsilon_G-\epsilon_B)$ is the fraction of $\mathcal{U}$ recipients.

For modeling simplicity, we assume homogeneous values for a given message type, and since participants know only expected values *ex ante*, we restrict attention to point masses $r_G = E\left[r_G^0\right]$ and $s_G = E\left[s_G^0\right]$ with parallel expressions for r_B and s_B. We allow correlated values for each successful match. So, for example, on receipt of a matched $\mathbb{G}$ message, recipient $\mathcal{G}$ realizes r_G and sender $\mathbb{G}$ realizes s_G. In contrast, on receipt of a mismatched $\mathbb{B}$ message, both recipient $\mathcal{G}$ and sender $\mathbb{B}$ realize 0. Regardless of match, all recipients pay cost c_r for each message read, and all senders pay cost c_s for each message sent. Figure 5 shows relative payoffs.

Senders can choose to send or not send to all recipients but cannot target a specific subgroup without more information. Given recipient heterogeneity, senders do not know s in advance and instead know only expected payoffs. Thus sender knowledge departs from the previous model. We assume that senders have no relevant capacity constraint (they participate under open access), so it is feasible to send to everyone. A sender capacity constraint is briefly considered in Section 7.2.

While positive correlation is unreasonable in some situations – such as with political messages, whose senders get free speech value even if recipients get none – it is common in others. For example, for marketing messages, senders receive value only when the recipient makes a purchase.

This setup now expands the model enough to analyze the three different scenarios of interest: the baseline open access (no modification) case, the ideal tax case, and the Attention Bond Mechanism.

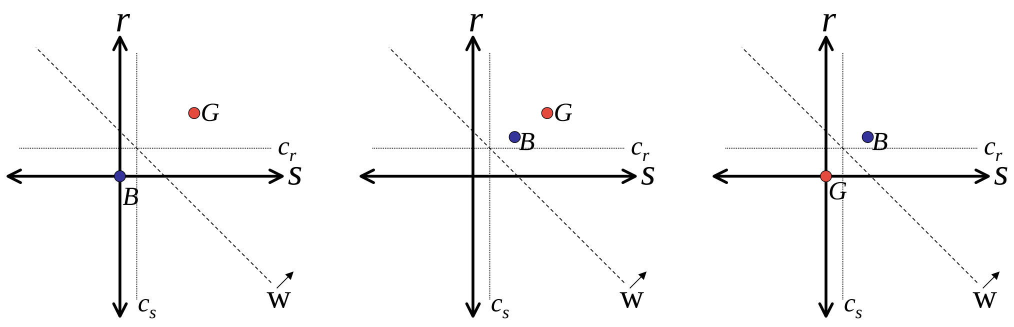
6.1 Open Access

Under open access, unable to distinguish among recipients before first-contact, the two sender types can expect the following surplus:

$$\mathrm{SS}_G^0 = (1-\epsilon_B)\cdot s_G - c_s \tag{20}$$

$$\mathrm{SS}_B^0 = (1-\epsilon_G)\cdot s_B - c_s \tag{21}$$

dimensions. Ordinal preferences of people who communicate on art, wine and file sharing are unlikely to match those of people who communicate on math, wine, and chemistry. Thus for n topics, we consider preferences over the 2^n subsets for the simplest informative case $n=2$.



(a) Recipient Type $\mathcal{G}$ (b) Recipient Type $\mathcal{U}$ (c) Recipient Type $\mathcal{B}$

Figure 5: Sender and receiver values by matching type. Note type $\mathcal{G}$ recipients have mass ϵ_G, $\mathcal{B}$ recipients have mass ϵ_B, and $\mathcal{U}$ recipients have mass $(1-\epsilon_G-\epsilon_B)$. Senders $\mathbb{G}$ and $\mathbb{B}$ have masses γ and $(1-\gamma)$.

In these expressions, the first term represents the fraction of matching recipients times the net value of each message. The second term represents the wasted cost of sending to everyone.

Parallel expressions hold for recipients. Each type receives positive value from matched messages only but bears costs from all received.

$$\mathrm{RS}^0_U = \gamma\cdot(r_G - c_r) + (1-\gamma)\cdot(r_B - c_r) \tag{22}$$

$$\mathrm{RS}^0_G = \gamma\cdot(r_G - c_r) + (1-\gamma)\cdot(-c_r) \tag{23}$$

$$\mathrm{RS}^0_B = \gamma\cdot(-c_r) + (1-\gamma)\cdot(r_B - c_r) \tag{24}$$

Here, the first term in each expression represents expected surplus from type $\mathbb{G}$ messages, and the second term represents the contribution from type $\mathbb{B}$ messages.

Combining these equations leads to total open access welfare.

$$\begin{aligned}
\mathrm{W}^0 &= (\gamma\mathrm{SS}_G + (1-\gamma)\mathrm{SS}_B) \\
&\quad + ((1-\epsilon_G-\epsilon_B)RS_U + \epsilon_G\mathrm{RS}_G + \epsilon_B\mathrm{RS}_B) \\
&= (\gamma(1-\epsilon_B)s_G + (1-\gamma)(1-\epsilon_G)s_B - c_s) \\
&\quad + (\gamma(1-\epsilon_G)r_G + (1-\gamma)(1-\epsilon_B)r_B - c_r) \\
&= \gamma(1-\epsilon_B)(s_G - c_s + r_G - c_r) \\
&\quad + (1-\gamma)(1-\epsilon_G)(s_B - c_s + r_B - c_r) \\
&\quad - \gamma\epsilon_B(c_s + c_r) - (1-\gamma)\epsilon_G(c_s + c_r)
\end{aligned} \tag{25}$$

In the final expression, the three principal terms represent (1) the total surplus of all $\mathbb{G}$ transactions, (2) the total surplus of all $\mathbb{B}$ transactions, and (3) the total waste, i.e. pollution, from misdirected communications. Waste results from two negative externalities, misdirected $\mathbb{G}$ messages and misdirected $\mathbb{B}$ messages.

6.2 Tax

To improve welfare, consider a government-issued tax t on all messages sent, geared to either compensate recipients for negative call externalities via side payments, or to cause one class of senders to stop sending.

Initially, imperfect information prevents senders from targeting first-contact messages. Senders could reduce the number of messages sent, but this is not rational since the marginal value of communications is linear in the number of recipients. Linearity occurs because the distribution of payoffs (i.e., recipient types ϵ_G and ϵ_B) is fixed, and marginal costs c_s are also fixed (and very nearly zero). Linearity implies the only possible maxima are corner solutions – either the sender sends to everyone or to no one. As such, the only tax that has any effect on welfare is one that eliminates one type of message, moving the solution from one corner to another. Transfer payments that do not alter send volume have no net effect on welfare.

To calculate the value of a welfare enhancing tax, assume *w.l.o.g.* that total expected value of type $\mathbb{G}$ communications exceeds that of type $\mathbb{B}$. That is:

$$\gamma(1-\epsilon_B)(s_G - c_s + r_G - c_r) > (1-\gamma)(1-\epsilon_G)(s_B - c_s + r_B - c_r) \tag{26}$$

This implies that if waste is large enough, social welfare improves when type $\mathbb{B}$ senders leave the market. This tax cannot be effective, however, unless it is also true that individual $\mathbb{B}$ senders have less transaction surplus than $\mathbb{G}$ senders.

$$(1-\epsilon_G)s_B - c_s < (1-\epsilon_B)s_G - c_s \tag{27}$$

If this were false, then a tax would either eliminate only type $\mathbb{G}$ senders, or eliminate all senders. As such, the optimal tax is

$$t = \mathrm{SS}_B^0 = (1-\epsilon_G)(s_B) - c_s \tag{28}$$

Any lower tax reduces sender income but fails to change the actions of any party. Any higher tax induces $\mathbb{B}$ senders to stop sending, but changes no other behavior unless it becomes so high that all senders leave.

This tax is Pigouvian. It is set at the level of call externalities for individual $\mathbb{B}$ senders. The optimal waste level is zero since the marginal message value to senders is constant and optima exist only as corner solutions. Both negative call externalities are also eliminated at the socially optimal level of transactions. These are mistargeted messages from $\mathbb{G}$ senders (to $\mathcal{B}$ recipients), and mistargeted messages from $\mathbb{B}$ senders (to $\mathcal{G}$ recipients). [10] No longer getting messages of any value, type $\mathcal{B}$ recipients leave the market.

[10] As the number of sender types N increases, the number of mismatch externalities grows as $O(N^2)$. These interactions imply that a single Pigouvian tax quickly has difficulty compensating for all possible call externalities.

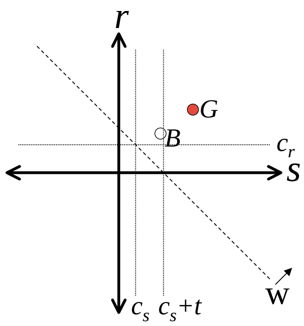


Figure 6: Tax Effect on R_U and the Senders

Relative to Open Access, each party now has surplus:

$$\mathrm{SS}_G^t = (1-\epsilon_B)\cdot(s_G - c_s - t) \tag{29}$$
$$\mathrm{SS}_B^t = 0 \tag{30}$$
$$\mathrm{RS}_U^t = \gamma\cdot(r_G - c_r) \tag{31}$$
$$\mathrm{RS}_G^t = \gamma\cdot(r_G - c_r) \tag{32}$$
$$\mathrm{RS}_B^t = 0 \tag{33}$$

Using tax revenues for social gain, all collections are added back into the social welfare equation. Total welfare under the tax scenario is therefore

$$\begin{aligned}\mathrm{W}^t &= \gamma\mathrm{SS}_G^t + (1-\epsilon_B-\epsilon_G)\mathrm{RS}_U^t + \epsilon_G\mathrm{RS}_G^t + \gamma(1-\epsilon_B)t \\ &= \gamma(1-\epsilon_B)(s_G - c_s + r_G - c_r)\end{aligned} \tag{34}$$

Comparing this welfare equation to W^0, it can be seen that the tax has two major effects. First, the tax eliminates all waste from misdirected messages (the last line of Equation 25). Second, it eliminates all benefits of type $\mathbb{B}$ transactions (the second-to-last line of Equation 25). The tax therefore is worthwhile for society when

$$\gamma\epsilon_B(c_s+c_r) + (1-\gamma)\epsilon_G(c_s+c_r) \geq (1-\gamma)(1-\epsilon_G)(s_B - c_s + r_B - c_r) \tag{35}$$

or when the waste from misdirected messages exceeds the benefits of type $\mathbb{B}$ transactions. The tax eliminates one entire class of senders in order to avoid the costs imposed on people who dislike their messages.

6.3 ABM

The Attention Bond Mechanism becomes more complicated to analyze in this setting. Internalizing the surplus of her matching sender, each recipient can rationally choose a different bond amount. Bonds then serve as signals of recipient type. Under a separating equilibrium, senders who can identify interested recipients can then avoid mistargeting.

Recipients, however, gain the ability to act strategically. They may post bond requests falsely signaling unrelated types in order to capture nonmatching bonds. They may even employ mixed strategies, posting bonds of different types with positive probabilities.

A recipient can always signal her true type. She can choose a “limit bond,” a bond request below the net surplus of a competing type, such that no other type prefers to pool. If all recipients signal truthfully, then all welfare positive messages are sent, and no welfare negative message are sent. This possibility leads to first-best welfare.

Not all pooling strategies are rational. A recipient who chooses to lie and pool with other recipient types cannot cause the sender who would normally communicate with that pool to exit the market. Her false signal has two consequences. It dilutes the number of legitimate matching transactions and it siphons off bonds. If a sender’s surplus falls below zero, his exit leads the recipient to prefer signaling truthfully. Open access surplus from her matched communications exceeds zero.

Senders still communicate with correctly matched recipients and those pooling with matched recipients provided that forfeiting misdirected bonds leave senders with non-negative surplus. This implies that equilibrium bonds of pooling recipients adjust to internalize not only the sender surplus, but also the dilution effects of untruthful recipients. Although truthful recipients could have chosen a limit bond, they prefer the pooling bond, even if diluted.

In contrast to recipients, senders under the ABM have little incentive to misrepresent their type as they gain nothing from mismatched messages and also lose their bonds. One exception occurs if the transaction surplus is so great that a sender can afford to follow an untruthful recipient into a nonmatching pool. If this happens, however, then the value of completed transactions exceeds the waste from mistargeting. A tax would have destroyed more value than the waste it avoided.

In general, the information signaling properties of the ABM lead to greater social welfare than that under any tax. We present this more formally below.

The number of equilibria possible with strategic behavior greatly complicates analysis and understanding of the ABM. To simplify, we divide the strategy space into the feasible equilibria of Table 2, then confirm stability and existence. This avoids overly complex characterizations of all possible optima under all possible values.

Below, we present lemmas that yield information about the behavior under the ABM. Proofs of these lemmas appear in the online Appendix.

Case	Pooled	Separate
1	$\{\}$	$\{\mathcal{U}\},\{\mathcal{G}\},\{\mathcal{B}\}$
2	$\{\mathcal{U},\mathcal{G}\}$	$\{\mathcal{B}\}$
3	$\{\mathcal{U},\mathcal{B}\}$	$\{\mathcal{G}\}$
4	$\{\mathcal{G},\mathcal{B}\}$	$\{\mathcal{U}\}$
5	$\{\mathcal{U},\mathcal{B},\mathcal{G}\}$	$\{\}$

Table 2: Feasible Equilibria

Lemma 5 *At least one sender type will have zero surplus under the ABM, but remain in the market.*

The intuition here is that rational recipients internalize as much of a sender's surplus as possible up to the constraint of keeping him in the market. If all sender types had surplus remaining, then recipients could post higher bonds.

Lemma 6 *Welfare under the ABM is at least that under open access, and all recipients that participate in open access remain in the market under the ABM.*

This is somewhat of a modeling detail, but is important. All recipients have the option of choosing a zero bond, replicating open access for themselves and their senders. Since they were in the market under open access voluntarily, they will never choose to leave when using the ABM, a result analogous to that of Section 4.

One subtlety, however, is that the ability to collect sender bonds can bring new participants into the market. Recipients who previously chose not to participate may enter. By the earlier logic, the pools that host these entrants cannot rationally require bonds that cause senders to exit, so these transfers alone leave total welfare unchanged. If these entrants bring their own positive call externalities or negative processing costs, total welfare can rise or fall.

Lemma 7 *Under the ABM, the sender with higher aggregate surplus is always preserved in the market, and always sends to all matching recipients.*

Self-interested recipients target the senders with more surplus to share. Recipients post bond requests as high as possible but cannot rationally post bonds that would drive the most valuable senders from the market, and so leave them with epsilon surplus. For the second half of the lemma, note that the matching recipient could not do better by switching, since mismatched senders have less aggregate surplus and mismatched message content provides no value, so they signal truthfully.

Lemma 8 *There exists a dominant strategy equilibrium where the ABM has first-best social welfare.*

Consider the case in which all recipient types choose distinct bond values, thereby uniquely identifying themselves to senders. When this happens, and we show in the Appendix that there exist values for which this is a dominant strategy, senders never mistarget their messages and everyone communicates with their matched type. Perfect matching and the absence of wasteful contact leads to a first-best, Pareto-optimal result.

Note, however, that this is not the sole case. There also exist distributions where one recipient can profit at the expense of another.

6.4 ABM Versus Tax

When the tax is worse for welfare than under open access, the ABM is better by transitivity. This occurs when $\mathbb{B}$ transaction value exceeds total waste or when the tax first drives $\mathbb{G}$ from the market. Hence, the primary remaining case arises when the costs of mistargeted messages exceed the lower value transaction.

Proposition 5 *Assuming heterogeneous senders and recipients, and homogeneous message value for each type, social welfare of a flat tax never exceeds social welfare under the ABM unless the $\mathcal{B}$ recipient is willing to forgo all type $\mathbb{B}$ messages.*

The proof is in the Appendix.

The intuition for why welfare under the ABM generally exceeds that under the tax follows from understanding the choices of each party: senders necessarily prefer to reach only matching recipients because unmatched recipients seize their bonds with no offsetting transaction benefit, a pure loss. Signaling choices of recipients then drive welfare.

If all recipients signal truthfully, the ABM yields first-best welfare and no mechanism can do better.

If any recipient signals untruthfully by choosing the bond of a mismatched type, she risks losing her native surplus, which includes both her message value and the bond from her matched sender type. A false signal indicates that (1) poached bond surplus exceeds her native surplus and (2) having pooled with a different type, the true bond signal of a recipient poacher will not be present in the market. The declared transaction type must then generate ample surplus implying that, for social efficiency, truthfully matched transactions should always complete.

Suppose that the matching sender for a recipient poacher chooses to follow that recipient by sending messages into the pool of the declared type. Such messages would represent spam to recipients who had truthfully declared their type, but would not be spam to the recipients who had lied. The decision to follow indicates that the sender (now a spammer) had sufficient surplus to lose bonds on truthful types, making up the value on transactions with untruthful types. A sender's decision to follow an untruthful recipient then implies that total transaction surplus, for types with no market signal, exceeds the externality waste they generate. Such transactions would have been destroyed by a tax, which would have reduced welfare.

To see that total value from these minority transactions exceeds total waste, note that honest recipient types must value the bond they receive from spammers more than the nuisance cost of processing spam. If not, they can choose a true type "limit bond" to force a separating equilibrium. Analogous to a limit price that keeps competitors from entering one's market, this limit bond prevents poachers from preferring an off-type bond to their native surplus. As a credible signal, a limit bond also creates no social inefficiency, as it merely returns surplus to senders.

This sections shows how, when using the ABM, the bond value functions as a signal to potential senders. This can facilitate improved matching of senders and recipients. However, it also shows there is strategic interaction between recipient types which complicate finding a strategy. Misrepresentation is not only possible, but new entrants may join with this intent.

6.5 Combining the ABM and Tax

As with the perfect filter, the tax could in theory be combined with the ABM. Such a mechanism might work by having recipients choose individual bond prices but having proceeds go to a charity or central government. If all sender and receiver values are correlated by type, and recipient values exceed transaction costs, first-best welfare might be possible.[11] Consider a budget balancing redistribution policy, like that of a Groves-Clarke mechanism, such that seized bonds would only be distributed to other players and not the recipient who decides whether a message is spam. Without direct bond payments, recipients lose their incentive to provide false signals regarding their own type (i.e. bond size) as well as for declaring spam (i.e. seize policy). This could give senders full information. If, however, recipients can still use filters to refuse messages below their private costs, or these values are not correlated as was the case in Section 4, this observation fails. Such combined mechanisms represent interesting avenues for further research.

[11] We thank an anonymous reviewer for this observation

7 Discussion and Extensions

This section examines the reasons for focusing on first versus subsequent contact, and considers why recipients might not seize all first-contact bonds. In the Appendix, interested readers may also find an analysis of increased information in the two cases where senders learn recipient values, and when both values are common knowledge. Additional information generally increases efficiency.

7.1 First Versus Subsequent Contact

This section resolves a participation puzzle: how can recipients post positive bond requests yet choose to participate in the medium under open access? Positive bonds imply negative expected message values, implying that recipients ought not participate.

As noted previously, the problem of first-contact is a two-sided information asymmetry among strangers. Senders do not yet know recipient tastes, and recipients do not yet know sender qualities. *Interim* rationality favors senders who put forward only net positive messages. Unfortunately for recipients, *ex ante* rationality implies that many first-contact messages can be pure waste.

To illustrate the recipient's dilemma when sender quality is unknown, let the expected value of the message stream for type $\mathbb{G}$ be positive, while the expected value for that of type $\mathbb{B}$ be negative. Further, let δ^t designate the discount rate for a message received at time t. Then the stream of expected value from each type is given by

$$E\left[r_G\right] = \sum_{i=0}^{\infty} \delta^t E\left[r_G^t\right] \geq 0 \qquad E\left[r_B\right] = \sum_{i=0}^{\infty} \delta^t E\left[r_B^t\right] \leq 0$$

Let $t = 0$ represent first-contact communications and, as before, let γ be the fraction of these messages from type $\mathbb{G}$ (and $1 - \gamma$ be the fraction from type $\mathbb{B}$). As with spam, the expected value of all first-contact messages will be wasteful if

$$\gamma E\left[r_G^0\right] + (1 - \gamma) E\left[r_B^0\right] \leq 0$$

Further, let α be the proportion of first-contact messages from the pool of all first and subsequent contacts. If reading a first-contact message allows a recipient to *ex post* verify sender type, she can stop challenging future messages from $\mathbb{G}$ via a whitelist (or, alternatively, change the bond amount). Since type $\mathbb{B}$ senders offer negative expected value, a recipient can rationally reject identifiable $\mathbb{B}$ content. So, its minimum value becomes not less than 0. Note that due to cheap identities, $\mathbb{B}$

senders can try to reenter the pool as first-contact strangers. Then, the total value of first and subsequent contact to a recipient is not less than

$$\alpha\left(\gamma E\left[r_G^0\right] + (1-\gamma)\, E\left[r_B^0\right]\right) + (1-\alpha))\left(\sum_{i=1}^{\infty} \delta^t E\left[r_G^t\right]\right) \tag{36}$$

Equation 36 is the minimum value for participation in the medium under open access. Although first-contact messages can generate negative value, implying a positive bond price, recipients can rationally participate so long as Equation 36 is positive.

We can interpret the legal, technical, and economic solutions for improving communications value within this highly simplified framework. Filtering solutions aim to increase communications value by rejecting type $\mathbb{B}$ messages, thereby increasing the fraction of good messages γ. Taxes and computational challenges seek to force type $\mathbb{B}$ senders to internalize costs, possibly driving them from the market. These costs reduce the proportion of first-contact messages α. If they differentially affect type $\mathbb{B}$ senders, they can shrink the total sender market and increase γ. In contrast, the ABM can decrease the fraction of type $\mathbb{B}$ messages, force $\mathbb{B}$ senders to internalize costs, and through wealth transfers increase the value of received messages. This potentially reduces α, and increases γ, but also increases $E\left[r_B^t\right]$ and $E\left[r_G^t\right]$.

One advantage of the ABM relative to taxes is that efficiency distortions only affect first-contact. Individual recipients can trivially distinguish between senders they know and senders they do not know with the assumption of strong identity. Distributed among market participants, this private knowledge is potentially observable but cannot be used by a government or central taxing authority without substantial loss of individual privacy. Thus, if α represents the proportion of first-contact messages, then the relative efficiency of a mechanism that distorts only first-contact communications is $\frac{1}{\alpha}$.

Further, as the number of participant types increases from 2 to N, the number of pairwise negative externality terms that result from misdirected messages rise as $O(N^2)$, while the number of message topic subsets that might interest any given recipient rises as $O(2^N)$. With linear marginal values, arising from fixed recipient probabilities and flat costs (or nearly so at large volumes), a tax can only alter the behaviors of senders whose marginal surplus is at or below the tax. Society's interest in restricting transactions below this cutoff hinges on the sort order of relative values. For each transaction, the sort order of sender surplus must be the same as the sort order of total surplus. This property may or may not hold, which makes determining the efficiency of a tax more difficult.

7.2 Why not Always Seize the Bond?

Since mistargeted communications provide no content benefit, recipients can rationally seize their bonds. For well-targeted communications, however, can a recipient justify releasing the bond? If not, the ABM exhibits one property of the tax in discouraging socially valuable messages that have little value to senders.

At issue is first-contact moral hazard. A recipient can seize the sunk cost bond then whitelist matching senders. Further, a sender threat not to send is not credible if $\sum_{i=0}^{\infty} E[s] - b \geq c_s$. Senders for whom further contact $\sum_{i=1}^{\infty} E[s] - b < c_s$ can trivially discover they have been whitelisted by not attaching a bond and noting the absence of a subsequent bond challenge.

We note that a clear reason to release the bond may be obscured by having simplified our analysis such that values are realized upon reading a message rather than after a full round of exchange. A sender might recover a bond, or at least negotiate, if he could charge on the return message generated by first-contact. Culture and social pressure could also constrain misbehavior among people who knew each other outside the communications medium. The following discussion offers four reasons why recipients might not otherwise seize bonds from matching senders based on signaling, reputation, competition, and repeated play.

The first potential reason is that releasing the bond can credibly signal a match. If pooling creates ambiguity regarding recipient type, and she must distinguish herself to realize gains from a match, then releasing the bond represents a costly investment. Willingness to bond at a specific level signals sender type, resolving half of the two-sided information asymmetry, while refunding the bond is a credible signal that resolves the other half. Note that any refund amount is possible that separates the matching recipient from non-matching types in her pool. In a transactions cost sense (Williamson, 1985), b_G represents a hostage offered by a type $\mathbb{G}$ sender to identify a superior opportunity, while the refund offered by the $\mathcal{G}$ recipient represents a relationship-specific investment to realize these opportunity gains.

A second reason is to invest in reputation so as to generate good transactions with senders whose own profits are marginal. Consider a class of valuable senders who do not participate under open access; for them, matching transactions value is either too small or too risky. In the case of email, for example, spamming the entire market may not be cost-effective advertising, or else more legitimate firms might choose it.

A sender who posts a bond extends credit in anticipation of other gain. As total sender surplus rises, the more marginal senders can afford to transact. A reputation for being a good creditor – gained by refunding all or part of a bond – signals to marginal senders of like type that targeting this recipient can leave the sender with positive surplus. Refunds are rational so long as recipient choices are public information and expected new transactions value exceeds one time bond value. Investing in reputation is not novel. Analogous arguments have been made for signaling higher-quality products (Rogerson, 1983), and less risky contracts Rasmusen (2001). In fact, "reputation systems" are common at many popular auction sites.

A third reason for a recipient to forgo bonds is competition for scarce *sender* attention. If senders are resource-constrained and limited to N communications in a market with $N + M$ recipients (and both $N \geq 1, M \geq 1$), then first-contact moral hazard implies that, other things being equal, senders prefer recipients with lower posted bonds. That is, senders prefer $\mathcal{G}_1$ to $\mathcal{G}_2$ whenever $E[s^G] - b_1 > E[s^G] - b_2$ with $b_1 < b_2$. As noted in Section 4, a bond can represent a monopoly price on recipient attention. If recipients are substitutes, however, the presence of M uncompensated recipients implies Bertrand price competition. Lower bonds, even to the point of subsidy, are possible as recipients bid away their surplus. As illustration, one might consider the bargaining power implications for Helen of Troy in a population of all men versus those of Casanova in a population of all women. If senders were male in both cases, the recipient with a monopoly on her own attention in one market, could be a price taker in the competitive market for attention in the other.

A final reason to forgo the bond is the presence of transaction costs in the context of round-robin exchange and repeated play. Senders and recipients may commonly reverse roles as the monopsonist in one period becomes the monopolist in the next. Having solved the matching problem, each could reciprocally post take-it-or-leave-it bond demands, seeking to appropriate the other's surplus.[12] If, however, the challenge-response process of bond demand and collection introduces even modest transaction costs, then both parties can rationally decide to forgo these added costs after having identified each other as a match. Coordinating on the joint use of whitelists avoids destroying surplus.

8 Conclusions

This article explores mechanisms designed to improve the value of communication in multi-party, point-to-point media where there exist call externalities and two-sided information asymmetry. Many common media exhibit such properties, and message pollution ("spam") is particularly evident in email under open access.

Under certain circumstances, we find that systems designed to screen unsolicited contact from senders of unknown quality can produce greater social welfare if based on market rather than technological or regulatory principles alone. Whereas

[12]For further analysis of take-it-or-leave-it offers see the Appendix

filters and taxes may block the flow of welfare-negative messages, the market-based approach can perform comparably well, while at the same time permitting exchange of welfare-positive messages that would not otherwise occur. Combined solutions can do better still.

This paper makes two contributions. First, recognizing the nature of the communications problem, it implements Coase's insight that granting property rights under a system to promote valuable exchange generally dominates other solutions to externality problems. The Coase Theorem, however, depends on low negotiation costs and here unwanted communication is itself the problem. The second contribution is to then add signaling and screening mechanisms, minimizing the need to negotiate. Results show how a solution to the two-sided information asymmetry problem improves welfare considering such factors as whether a mechanism permits or prevents side payments, supports decentralized or centralized decision making, distinguishes first from subsequent contact, and discloses or hides recipient preferences and sender qualities.

By permitting wealth transfers, the ABM can recover surplus lost under pure blocking technologies such as filters. Unlike taxes, variable bonds tailor to individual preferences, enabling recipients to adjust penalties in order to receive more messages they want and fewer they do not. Since recipients decide the penalty after receiving a message, the ABM helps screen senders by imposing higher expected costs on lower valued messages. The ABM's recipient-chosen bond size functions as a signal to senders, which improves potential matching. Posting a bond then serves as an implied match. This shifts the burden of message classification to senders who have better private knowledge of message content. Finally, returning a bond can function as a recipient signal of intent to continue a relationship.

Limitations include recipient ability to misrepresent type, thus exploiting information asymmetry, and the possibility of recipients overcharging senders, which might lead to losing welfare-positive transactions. The strategic decisions facing recipients may be complicated, both in setting the bond (the value is a signal, so what do they want to signal?) and in returning the bond (the ideal would be to return at a rate that only punishes unwanted senders). Empirical estimates for message value distributions, beyond those modeled here, will also provide assurance of which approach balances welfare tradeoffs most successfully.

To implement the ABM – and give consumers property rights in their own attention – fortunately does not require the force of law. A basis for achieving equivalent rights is generally obtainable, especially for the class of media in question, through technology (Lessig, 2000). Moreover, the ABM's binding commitment and *ex post* verification of message content allow it to avoid the classification problem that real content-based filters, rather than our idealized perfect one, cannot. The ABM creates a market for attention that can benefit both senders and recipients.

Acknowledgements

We are very grateful for the assistance of Mark Benerofe in revising and clarifying the ideas presented here. We also appreciate the feedback on early versions of this paper from Bob Marinier, Terrance Kelly, Dan Silverman, Jussi Keppo, Peter Honeyman, Paul Laskowski, and Lori Cranor. Particular thanks goes to Jeff Mackie-Mason for a careful analysis of our methods and results. We appreciate constructive feedback from Andrei Hagiu, Lones Smith, the Associate Editor, and three anonymous reviewers.

References

Abadi, Martin; Birrell, Andrew; Burrows, Mike; Dabek, Frank and Wobber, Ted, "Bankable Postage for Network Services." In "Advances in Computing Science ASIAN 2003," Springer-Verlag, 2003, number 2896 in Lecture Notes in Computer Science, pp. 72–90.

Akerlof, George A, "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics*, *84*(3), pp. 488–500, 1970.

Ayres, Ian and Nalebuff, Barry, Why Not? How to Use Everyday Ingenuity to Solve Problems Big and Small. Harvard Business School Press, 2003.

Bazeley, M, "New Weapon for Spam: Bounty." In "San Jose Mercury News; April 26," 2003.

Cranor, Lorrie Faith and LaMacchia, Brian A., "Spam!" *Communications of the ACM*, *41*(8), pp. 74–83, 1998, URL `http://doi.acm.org/10.1145/280324.280336`.

Dwork, Cynthia and Naor, Moni, "Pricing via Processing or Combatting Junk Mail." In "Advances in Cryptology – CRYPTO 1992," Springer-Verlag, 1993, number 740 in Lecture Notes in Computer Science, pp. 139–147.

Fahlman, Scott E, "Selling Interrupt Rights: A Way to Control Unwanted E-Mail and Telephone Calls." *IBM Systems Journal*, *41*(4), pp. 759–766, 2002.

Friedman, E and Resnick, Paul, "The Social Cost of Cheap Pseudonyms." *Journal of Economics and Management Strategy*, *10*(2), pp. 173–199, 2001.

Glassman; Manasse; Abadi; Gauthier and Sobalvarro, "Escrow services are financially viable for sub penny transactions: The Millicent Protocol for Inexpensive Electronic Commerce." In "Fourth International WWW Conference," 1995.

Graham-Cumming, John, "Beating Bayesian filters." In "MIT Spam Conference," 2004.

Hann, Il-Horn; Hui, Kai-Lung; Png, I.P.L. and Lee, Sang-Yong Tom, "Direct Marketing: Privacy and Competition.", 2004, mimeo: University of Southern California.

Hermalin, Benjamin E. and Katz, Michael L., "Sender or Receiver: Who Should Pay to Exchange an Electronic Message?" *RAND Journal of Economics*, *35*(3), pp. 423–447, 2004.

Kraut, R; Sunder, Shyam; Morris, J; Cronin, M and Filer, D, “Markets for Attention: Will Postage for Email Help?” In “ACM Conference on CSCW,” 2003, pp. 206–215.

Krishnamurthy, Balachander, “SHRED: Spam Harassment Reduction via Economic Disincentives.”, 2004, URL `http://www.research.att.com/~bala/papers/shred-ext.pdf`, Working Paper, AT&T Research.

Laurie, Ben and Clayton, Richard, “Proof-of-Work Proves Not to Work.” In “Workshop on Economics and Information Security,” 2004, URL `http://www.dtc.umn.edu/weis2004/clayton.pdf`.

Lessig, Lawrence, Code and Other Laws of Cyberspace. Basic Books, 2000.

Libbey, Miles, “Learning from 2003: Spamming Trends and Key Insights.” 2004, presentation at the 2004 MIT Spam Conference. `http://spamconference.org/talks2004.html`.

Myerson, R and Satterthwaite, M, “Efficient Mechanisms for Bilateral Trading.” *Journal of Economic Theory*, *29*, pp. 265–281, 1983.

Pantel, Patrick and Lin, Dekang, “SpamCop– A Spam Classification & Organization Program.” In “Proceedings of AAAI-98 Workshop on Learning for Text Categorization,” 1998.

Rainie, Lee and Fallows, Deborah, “The CAN-SPAM Act has not helped most email users.” 2004, `http://www.pewinternet.org/report_display.asp?r=116`.

Rasmusen, Eric, “Explaining Incomplete Contracts as the Result of Contract-Reading Costs.” *Advances in Economic Analysis and Policy*, *1*(1), 2001.

Rogerson, William P, “Reputation and Product Quality.” *Bell Journal of Economics*, *14*(2), pp. 508–516, 1983.

Sahami, Mehran; Dumais, Susan; Heckerman, David and Horvitz, Eric, “A Bayesian Approach to Filtering Junk E-Mail.” In “Proceedings of AAAI-98 Workshop on Learning for Text Categorization,” 1998.

Shapiro, Carl, “Consumer Information, Product Quality, and Seller Reputation.” *The Bell Journal of Economics*, *13*(1), pp. 20–35, 1982.

Spence, Michael, “Job Market Signaling.” *The Quarterly Journal of Economics*, *87*(3), pp. 355–374, 1973.

Stiglitz, Joseph, "The Theory of Screening, Education, and the Distribution of Income." *American Economic Review*, *65*(3), pp. 283–300, 1975.

Tompkins, Trevor and Handley, Dan, "Giving E-mail back to the users: Using digital signatures to solve the spam problem." *First Monday*, *8*(9), 2003, `http://firstmonday.org/issues/issue8_9/tompkins/index.html`.

van Zandt, Timothy, "Information Overload in a Network of Targeted Communication." *RAND Journal of Economics*, *35*(3), pp. 542–560, 2004.

von Ahn, L.; Blum, M.; Hopper, N. and Langford, J., "CAPTCHA: Using hard AI problems for security." In "Proceedings of EuroCRYPT," 2003, URL `citeseer.ist.psu.edu/vonahn03captcha.html`.

Williamson, O, The Economic Institutions of Capitalism. Free Press, 1985.

An Economic Response to Unsolicited Communication
Appendices

Thede Loder[*] Marshall Van Alstyne[†] Rick Wash[‡]

February 26, 2006

A Increased Information

There exist two ways to increase participant information: the sender can know r, and the recipient can know s. Given the sequence of communication, we first analyze informed senders, then add informed recipients. In the latter, private values become common knowledge. In both models, increasing information increases efficiency relative to a perfect filter.

Note that perfect correlation of sender and receiver values (assuming invertibility) is a subset of the common knowledge case. As before, the recipient need not know sender value *ex ante* but she can infer it *ex post*. This follows from the fact that knowing either private value and knowing the correlation function is sufficient to compute the other value.

In the analysis that follows, we make one additional assumption, that the recipient can credibly commit *ex ante* to a policy. Since information is known, commitment can be enforced via external reputation systems or a trusted third party. It does not matter how it is enforced as long as the recipient has the proper incentives to execute the policy to which she committed. As in (Rasmusen, 2001) precommitment and reputations improve welfare. We also relax our assumptions on the distribution of communications values. The following mechanisms work regardless of the magnitude of value, the density function

[*] tloder@umich.edu
[†] mvanalst@umich.edu
[‡] rwash@umich.edu

on value, and the functional form on correlated value (assuming invertibility), and require only finite maximum and minimum values.

A.1 Informed Senders

Proposition 1 *If recipients can commit* ex ante *to a policy and the sender* ex ante *knows the recipient's value, the recipient surplus is at least as great under the ABM as under the perfect filter, always:*

$$\mathrm{RS}_{ABM} \geq \mathrm{RS}_{PF}$$

Proof: 1 We offer an existence proof that a Nash equilibrium choices in filtering strategies can do no better than those in ABM strategies. Consider a recipient request for a bond of size $b = c_r - \underline{r}$ coupled with the recipient's *ex ante* commitment to a policy of refunding $\rho(r)$, where

$$\rho(r) = \begin{cases} r - \underline{r} & \text{if } r < c_r \\ c_r - \underline{r} & \text{otherwise.} \end{cases}$$

First consider the case in which $r < c_r$. This represents communications that are unwanted by the recipient and would be filtered by the perfect filter. Here, the recipient keeps $b - \rho(r) = c_r - \underline{r} - (r - \underline{r}) = c_r - r$ of the bond. Adding this to the value of the email leaves her with surplus $r - c_r + c_r - r = 0$. Participation is Individually Rational for the recipient and provides the same surplus as the perfect filter. The sender, however, receives surplus $s - c_s - b + \rho(r) = s - c_s - (c_r - \underline{r}) + (r - \underline{r}) = s - c_s + r - c_r$. Whenever social surplus is positive, the sender rationally sends.

Consider now $r \geq c_r$. Here the recipient returns the entire bond ($\rho(r) = b$) and both parties keep their normal baseline surplus. This reflects the same surplus for both parties as under the perfect filter.

The strategy for the sender depends on knowing both r and s, while the recipient strategy depends only on knowing r. Since this strategy works regardless of sender strategy (it is a dominant strategy), any recipient chosen strategy must do at least this well. This strategy recreates the perfect filter strategy from the recipient viewpoint, but leaves the sender with the surplus of the communications that would have been filtered. This equivalence was not possible before because the private value of r did not allow senders to compute their own best response. ■

A.2 Common Information

Here we consider what happens when both the sender and recipient know each other's value. The sender knows the value to the recipient *ex ante*, but the recipient does not learn the sender's value until she learns her own private r after receiving and incurring costs c_r.

Proposition 2 *If the recipient can commit* ex ante *to a policy, senders know* ex ante *the value to the recipient, and the recipient learns or can infer* s *after learning* r*, then recipient surplus under the ABM is at least as great as that under the perfect filter:* $\mathrm{RS}_{ABM} \geq \mathrm{RS}_{PF}$*. In addition, maximum social surplus is achieved, with all surplus going to the recipient, meaning* $\mathrm{W}_{ABM} = \max W$*,* $\mathrm{SS}_{ABM} = 0$*, and* $\mathrm{RS}_{ABM} = \mathrm{W}_{ABM}$*.*

Proof: 2 Consider the following mechanism. The recipient requests a bond of size $b = \max\{\overline{s} - c_s, c_r - \underline{r}\}$. She commits *ex ante* to refund $\rho(s,r)$ to the sender, where

$$\rho(s,r) = b - \max\{s - c_s - \epsilon, c_r - r\}$$

and where $\epsilon \geq 0$. The sender's strategy is to send if and only if $s - c_s - b + \rho(s,r) \geq 0$.

This mechanism is Budget Balanced, Individually Rational for both parties, Efficient, and is a dominant strategy equilibrium. Since this mechanism uses no outside funds, it is trivially Budget Balanced. Next we show that it is Individually Rational.

The bond is added to the recipient's net value, but the refund is then removed. This yields her total surplus:

$$\begin{aligned}
\mathrm{RS}_{ABM} &= r - c_r + b - \rho(r,s) \\
&= r - c_r + b - (b - \max\{s - c_s - \epsilon, c_r - r\}) \\
&= r - c_r + \max\{s - c_s - \epsilon, c_r - r\} \\
&= \max\{r - c_r + s - c_s - \epsilon, 0\}
\end{aligned}$$

Since $\mathrm{RS}_{ABM} \geq 0$, it is *ex ante* Individually Rational for the recipient to participate. As before, the sender's interim Individual Rationality allows him to choose per-message whether to send.

Next we show efficiency. First we must determine sender willingness to send messages. The sender loses the value of the bond b and receives his net value plus the refund from the recipient:

$$
\begin{aligned}
\mathrm{SS}_{ABM} &= s - c_s - b + \rho(s, r) \\
&= s - c_s - b + b - \max\{s - c_s - \epsilon, c_r - r\} \\
&= s - c_s - \max\{s - c_s - \epsilon, c_r - r\} \\
&= \min\{(s - c_s) - (s - c_s - \epsilon), (s - c_s) - (c_r - r)\} \\
&= \min\{\epsilon, s - c_s + r - c_r\}
\end{aligned}
$$

Since $\epsilon \geq 0$, the sender will choose *not* to send messages only when $s - c_s + r - c_r < 0$. This is precisely the condition for positive welfare. Therefore, all positive welfare messages are sent, yielding the maximum possible total surplus.

Finally, we show that this mechanism is a dominant strategy equilibrium. Let $\epsilon \to 0$. This provides the recipient with all available surplus, and leaves zero surplus for the sender while still remaining Individually Rational. Since the sender has interim Individual Rationality, it is impossible for any mechanism to provide the recipient with greater surplus. Therefore, it is a dominant strategy for the recipient. Given this dominant strategy, the best response for the sender is to participate, as at worst he is indifferent between participation and nonparticipation. ■

Achieving maximum surplus is consistent with the Myerson and Satterthwaite (1983) claim on the impossibility of designing a mechanism for which Individual Rationality, Budget Balance, and Efficiency all hold when there also exists the possibility of inefficient trade, here $s + r < c_s + c_r$. Their theorem hinges on the fact that private information introduces inefficiency. The mechanism must provide both participation incentive and cover information rents. Because full information eliminates information rents, we can achieve all three criteria simultaneously. As a special case, consider the situation in which the recipient's value is any invertible function of the sender's value ($r = f(s)$).

Since communication itself is the negotiation problem, the ABM takes advantage of the fact that senders initiate contact. It avoids negotiation and instead has the recipient implement a take-it-or-leave-it contact policy that shifts power to the recipient. In this case, it is possible both for these transactions to be efficient and for the recipient to take all surplus from the sender.

Two other conditions can lead to first-best levels of welfare. One is a recipient capacity constraint. If recipients are boundedly rational and can process no more than m messages, sender and recipient values are private information, and senders want at most one unit of attention, then as van Zandt (2004) observes, scarce attention is efficiently allocated by a Vickrey auction. The communications network delivers messages with the m highest positive bids, transferring the value of the $(m+1)^{st}$ highest bid or 0 if senders place $\leq m$ bids. Because of the practical complexity of the first-best mechanism, van Zandt develops a second best alternative in the form of a tax. This reduces "information overload" caused by over-exploitation of scarce attention.

Section 6 in the main paper illustrates a second mechanism that can produce first-best welfare in the context of heterogeneous senders and recipients. In certain cases, a separating equilibrium is possible that supports perfect matching of senders and recipients. The choice of bond serves as both screen for unwanted contact and signal of recipient type.

B Proof of Lemma 5

The proof has two steps. We show that at least one sender always remains in the market, and of those that remain, at least one has zero surplus. We show this in reverse order.

Proof: Consider the choice of a bond for a recipient. We proceed with a proof by contradiction. Assume that a recipient pool chooses a bond that leaves all its matching senders with positive surplus. This is not incentive compatible because the recipient could increase the bond by ϵ, receive higher surplus, and not change the behavior of any sender in the pool. Therefore at least one sender will have zero surplus.

Next assume that the recipient chooses a bond higher than maximum expected surplus of any sender. This will cause all senders to stop sending to this recipient type causing the recipient to receive zero surplus. Reducing the bond enough so that at least one sender sends to this type increases that recipient type's surplus. Therefore, at least one sender will remain in the market. ■

C Proof of Lemma 6

Proof: To establish a contradiction, assume that a bond choice grants any recipient *less* surplus than under Open Access. Then, choosing a bond value of 0 reproduces surplus of the baseline for that recipient type and all matching senders. Note that any equilibrium choice but 0 by other recipient types improves sender targeting, strictly increasing welfare for recipients who choose 0 bonds. Thus recipient surplus can be equal to or greater than under Open Access, but it cannot be less. ■

D Proof of Lemma 7

Proof: Consider the recipient choice of the bond size, as this is the first strategic move. If a recipient chooses a pooling bond that causes her to stop receiving matched messages, then the value she receives from mismatched sender bonds must exceed the maximum value of receiving matching bonds and messages. Shown is the proof for $\mathcal{G}$ type recipients. For type $\mathcal{B}$ recipients, the proof is similar.

$$\gamma(s_G - c_s + r_G - c_r) \leq (1-\gamma)(s_B - c_s - c_r) \tag{1}$$

If this is true, then

$$\gamma(s_G - c_S) \leq \gamma(s_G - c_s + r_G - c_r) \leq (1-\gamma)(s_B - c_s - c_r) \leq (1-\gamma)(s_B - c_s) \tag{2}$$

so

$$\gamma(s_G - c_S) \leq (1-\gamma)(s_B - c_s) \tag{3}$$

So, if a recipient is willing to forgo receipt of messages she likes, then the other sender type has greater aggregate surplus. Or, the contrapositive of this is that if a sender has greater aggregate surplus, then no recipient who likes those messages is willing to forgo them. ■

E Proof of Lemma 8

Proof: If all three recipient types have separate bond values, then senders will be able to distinguish recipients by type. In this situation, senders prefer to send to only those recipients who receive value from their messages, and thus no losses are created due to mistargeting. At the same time, all messages that are valuable communications are sent. This leads to the maximum social surplus. This corresponds to Case 1 in Table 1.

If all three recipient types choose separate bonds, then senders will be able to distinguish recipients by type. In this situation, senders prefer to send to only those recipients who value their messages, and thus no losses are created as a result of mistargeting. At the same time, all messages that are valuable communications are sent. This leads to the maximum social surplus, corresponding to Case 1 above.

First, we define conditions that must be true for type $\mathcal{G}$ recipients to separate. If type $\mathbb{G}$ senders are willing to send to a pool including the type $\mathcal{G}$

recipients, then the constraint for type $\mathcal{G}$ recipients to separate is

$$(1-\gamma)c_r \geq ((1-\epsilon_G)s_B - c_s) - \gamma(s_G - c_s) \tag{4}$$

However, if the $\mathbb{G}$ senders forgo any value from type $\mathcal{G}$ recipients, then the constraint is

$$c_r \geq (1-\epsilon_G)s_B - c_s \tag{5}$$

Next, we have similar constraints for type $\mathcal{B}$ recipients:

$$\gamma c_r \geq ((1-\epsilon_B)s_G - c_s) - (1-\gamma)(s_B - c_s) \tag{6}$$
$$c_r \geq (1-\epsilon_B)s_G - c_s \tag{7}$$

Finally, we need to determine when type $\mathcal{U}$ recipients will separate. They have two logical strategies: pool with the recipients who collect a higher bond, forgoing any value from the other sender type, or collect a bond δ below the smaller bond, uniquely identifying themselves and receiving *all* messages. If the type $\mathbb{G}$ senders have a higher surplus, then the constraint for being unique is

$$\gamma(s_G - s_B) \leq (1-\gamma)(r_B - c_r + s_B - c_s) - \delta \tag{8}$$

Likewise, if type $\mathbb{B}$ senders have a higher surplus, then the constraint is

$$(1-\gamma)(s_B - s_G) \leq \gamma(r_G - c_r + s_G - c_s) - \delta \tag{9}$$

If all six of these inequalities hold, then no recipient will have an incentive to pool, and the three recipient types will all have distinct bond values. This situation produces a dominant strategy equilibrium.

Finally, we provide example values that show the intersection of these inequalities is non-empty:

$$\begin{aligned} c_s = c_r &= \frac{3}{2} \\ s_G = s_B &= 2 \\ r_G = r_B &= 3 \\ \gamma &= \frac{1}{2} \\ \epsilon_G = \epsilon_B &= \frac{1}{3} \end{aligned}$$

Note that if we let $c_s = c_r = 1$, then this first-best property of the ABM still holds, but a flat tax actually harms welfare. ■

Case	Pooled	Separate
1	$\{\}$	$\{U\}, \{G\}, \{B\}$
2	$\{U, G\}$	$\{B\}$
3	$\{U, B\}$	$\{G\}$
4	$\{G, B\}$	$\{U\}$
5	$\{U, B, G\}$	$\{\}$

Table 1: Feasible Equilibria

F Proof of Proposition 5

What follows is an analysis and comparison with a tax for each of the five equilibrium cases of the ABM.

$\mathcal{U}$, $\mathcal{G}$, and $\mathcal{B}$ are all separate By Lemma 8 above, the ABM is a first-best solution in this case. Therefore, it is at least as good as any other solution, including the flat tax.

$\mathcal{U}$ and $\mathcal{G}$ pool, $\mathcal{B}$ is separate For this situation, we must consider the behavior of the senders. A type $\mathbb{G}$ sender will always just send to the pooled $\mathcal{U}$ and $\mathcal{G}$ recipients, as they are exactly who he wants to reach. However, a type $\mathbb{B}$ sender has a choice. He can send only to $\mathcal{B}$ recipients, giving up any value from the $\mathcal{U}$ type recipients. Or he can send to everyone, incurring the costs of mistargeting the $\mathcal{G}$ types. Note that these costs also include the cost of paying a bond to the $\mathcal{G}$ types, who have no incentive to do anything other than collect the bond for mistargeted messages.

Subcase A: First let us consider the case where type $\mathbb{B}$ senders only send to type $\mathcal{B}$ recipients, and do not send to the pool. In this case, no costs arise from mistargeted messages. As such, regardless of bond values and collection policies, the total welfare for the system will be

$$\begin{aligned}\mathrm{W}^{\mathrm{ABM}} = {} & \gamma(1-\epsilon_B)(s_G - c_s + r_G - c_r) \\ & + (1-\gamma)(\epsilon_B)(s_B - c_s + r_B - c_r)\end{aligned} \tag{10}$$

which is strictly greater than W^t, the welfare from a tax.

Subcase B: Now, we must consider the case where type $\mathbb{B}$ senders send to everyone. In this case, the total welfare of the system will be

$$\begin{aligned} \mathrm{W}^{\mathrm{ABM}} = {} & \gamma(1-\epsilon_B)(s_G - c_s + r_G - c_r) \\ & + (1-\gamma)(1-\epsilon_G)(s_B - c_s + r_B - c_r) \\ & - (1-\gamma)(\epsilon_G)(c_s + c_r) \end{aligned} \tag{11}$$

$$\begin{aligned} \mathrm{W}^{\mathrm{ABM}} = {} & \gamma(1-\epsilon_B)(s_G - c_s + r_G - c_r) \\ & + (1-\gamma)\epsilon_B(s_B - c_s + r_B - c_r) \\ & + (1-\gamma)(1-\epsilon_G-\epsilon_B)(s_B - c_s + r_B - c_r) - (1-\gamma)(\epsilon_G)(c_s + c_r) \end{aligned} \tag{12}$$

In this equation, the first line represents all transactions between type $\mathbb{G}$ senders and the pool. This is the surplus under a tax. The second line represents all transactions between type $\mathbb{B}$ senders and type $\mathcal{B}$ recipients. This is positive by assumption. The final line represents the transactions between type $\mathbb{B}$ senders and the pool, including both messages to type $\mathcal{U}$ recipients and the mistargeted messages to type $\mathcal{G}$ recipients.

Considering bonds paid and received, the final line can be divided into three parts, the surpluses for the pooled recipients and the spamming sender, where $b - c_s$ is the bond charged by the pool:

$\mathcal{U}$ Recipients	$(1-\gamma)(1-\epsilon_G-\epsilon_B)(b - c_s + r_B - c_r)$
$\mathcal{G}$ Recipients	$(1-\gamma)\epsilon_G(b - c_s - c_r)$
$\mathcal{B}$ Senders	$(1-\gamma)\left[(1-\epsilon_G-\epsilon_B)s_B - (1-\epsilon_B)b\right]$

All three of these equations must be positive by Individual Rationality. The second equation shows the $\mathcal{G}$ recipient bearing all mistargeting costs in the form of reduced bonds and reading costs she pays directly. If the bond b is not enough to cover those costs, then $\mathcal{G}$ recipients could lower their bond by δ to distinguish themselves from $\mathcal{U}$ recipients. Therefore, the surplus generated by having these transactions in the market is positive, and the ABM is the more socially beneficial mechanism in this situation.

$\mathcal{U}$ and $\mathcal{B}$ pool, $\mathcal{G}$ is separate This situation is very similar to the previous case. It too has two subcases, that in which type $\mathbb{G}$ senders send to only $\mathcal{G}$ recipients, and the other in which they send to everyone.

Subcase A: When the $\mathbb{G}$ senders do not send to the pool, then the total social welfare is

$$\begin{aligned} W &= \gamma\epsilon_G(s_G - c_s + r_G - c_r) + \\ &\quad (1-\gamma)(1-\epsilon_G)(s_B - c_s + r_B - c_r) \\ &= \gamma(1-\epsilon_B)(s_G - c_s + r_G - c_r) + \\ &\quad (1-\gamma)(1-\epsilon_G)(s_B - c_s + r_B - c_r) + \\ &\quad -\gamma(1-\epsilon_G-\epsilon_B)(s_G - c_s + r_G - c_r) \end{aligned}$$

The universals pool only if

$$\begin{aligned} (1-\gamma)(s_B - c_s + r_B - c_r) &\geq \gamma(s_G - c_s + r_G - c_r) + (1-\gamma)(s_G - c_s + r_B - c_r) \\ &\geq \gamma(s_G - c_s + r_G - c_r) \end{aligned}$$

Multiplying both sides by $(1-\epsilon_B-\epsilon_G)$, we get

$$\begin{aligned} \gamma(1-\epsilon_B-\epsilon_G)(s_G - c_s + r_G - c_r) &\leq (1-\gamma)(1-\epsilon_B-\epsilon_G)(s_B - c_s + r_B - c_r) \\ &\leq (1-\gamma)(1-\epsilon_B)(s_B - c_s + r_B - c_r) \end{aligned}$$

which implies that the welfare from the ABM is greater than that of the tax.

Subcase B: The welfare in this subcase is

$$\begin{aligned} W &= \gamma(1-\epsilon_B)(s_G - c_s + r_G - c_r) + \\ &\quad (1-\gamma)(1-\epsilon_G)(s_B - c_s + r_B - c_r) + \\ &\quad -\gamma\epsilon_B(c_s + c_r) \\ &= (1-\gamma)(1-\epsilon_G)(s_B - c_s + r_B - c_r) + \\ &\quad \gamma\epsilon_G(s_G - c_s + r_G - c_r) + \\ &\quad \gamma(1-\epsilon_G-\epsilon_B)(s_G - c_s + r_G - c_r) - \gamma\epsilon_B(c_s + c_r) \end{aligned}$$

As before, this last line represents the total surplus from type $\mathbb{G}$ senders when sending to the pool, and can be divided to the surplus retained by each party involved (given a bond $b - c_s$):

$\mathcal{U}$ Recipients	$\gamma(1-\epsilon_G-\epsilon_B)(b - c_s + r_G - c_s)$
$\mathcal{B}$ Recipients	$\gamma\epsilon_B(b - c_s - c_r)$
$\mathcal{G}$ Senders	$\gamma\left[(1-\epsilon_B-\epsilon_G)s_G - (1-\epsilon_G)b\right]$

All three of these surpluses must be non-negative by Individual Rationality. In addition to this, type $\mathcal{U}$ recipients had to have chosen to be in the pool:

$$\gamma(b - c_s + r_G - c_s) + (1-\gamma)(b - c_s + r_B - c_s) \geq \gamma(s_G - c_s + r_G - c_s) \quad (13)$$

An extended proof follows.

$$(1 - \epsilon_B - \epsilon_G)s_B - (1 - \epsilon_G)b \geq 0 \quad (14)$$

Senders afford to spam

$$(1 - \epsilon_G)s_G - \epsilon_B \geq (1 - \epsilon_G)b \quad (15)$$

$$b \geq c_s + c_r \quad (16)$$

since $\mathcal{B}$ recipients are willing to pool

$$s_G - c_s \geq b - c_S \quad (17)$$

$\mathbb{G}$ senders can afford to send to the pool

$$s_G \geq c_s + c_r \quad (18)$$

$$(1 - \epsilon_G)s_G - \epsilon_B(c_s + c_r) \geq (1 - \epsilon_G)s_G - \epsilon_B s_G \quad (19)$$

$$(1 - \epsilon_G)s_G - (1 - \epsilon_B)(c_s + c_r) \geq (1 - \epsilon_G)b \quad (20)$$

from Eq. 15 and Eq. 19

$$(1 - \epsilon_G)s_G - (1 - \epsilon_G)b \geq (1 - \epsilon_B)(c_s + c_r) \quad (21)$$

$$\gamma(b - c_s + r_G - c_s) + (1-\gamma)(b - c_s + r_B - c_s) \geq \gamma(s_G - c_s + r_G - c_s) \quad (22)$$

since $\mathcal{U}$ is willing to be in pool

$$(1-\gamma)(b - c_s + r_B - c_r) \geq \gamma(s_G - b) \quad (23)$$

$$(1-\gamma)(1 - \epsilon_G)(b - c_s + r_B - c_r) \geq \gamma(1 - \epsilon_G)(s_G - b) \quad (24)$$

$$\gamma(1 - \epsilon_G)(s_G - b) \geq \gamma(1 - \epsilon_B)(c_s + c_r) \quad (25)$$

from Eq. 21

$$(1-\gamma)(1 - \epsilon_G)(b - c_s + r_B - c_r) \geq \gamma(1 - \epsilon_B)(c_s + c_r) \quad (26)$$

$$s_B \geq b \quad (27)$$

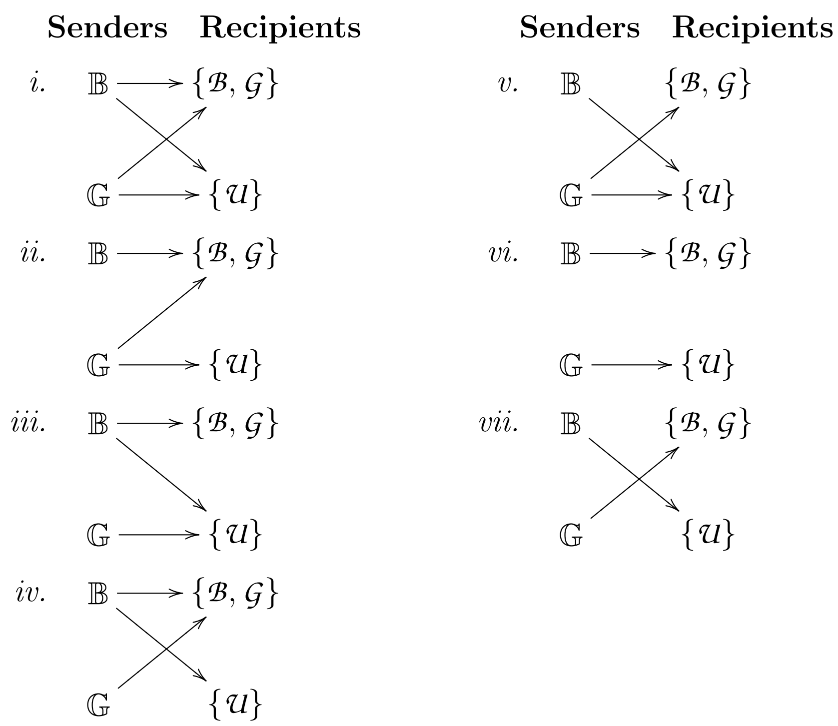


Figure 1: Subcases where $\mathcal{G}$ and $\mathcal{B}$ recipients pool together

$\mathbb{B}$ senders can afford it

$$(1-\gamma)(1-\epsilon_G)(s_B - c_s + r_B - c_r) \geq \gamma(1-\epsilon_B)(c_s + c_r) \tag{28}$$

This last line shows that the benefits of the ABM over the tax, and allowing type $\mathbb{B}$ to remain in the market, exceed any costs it imposes due to mistargeting.

$\mathcal{G}$ and $\mathcal{B}$ pool, $\mathcal{U}$ is separate This situation can be divided into five distinct subcases. These are graphically depicted in Figure 1. The first subcase involves all senders sending to all recipients, but universal recipients have a bond value distinct from the other recipient types. The second and fourth subcases both involve the situation where $\mathcal{B}$ and $\mathcal{G}$ recipients pool together on a bond low enough that both sender types send to them. The third and fifth subcases have the pool on a higher bond such that only one type of sender sends to the pool.

Subcase *i* This situation cannot happen. To see this, ask what (distinct) bond values the pool and type $\mathcal{U}$ recipients will have.

Lemma 1 *There cannot exist two distinct bond values such that all sender types send their messages to recipients with those values.*

Proof: A quick proof by contradiction will show this. Assume that there exist two distinct bonds such that both senders send and pay both bond values. At least one party on the lower bond has an incentive to increase her bond value to that of the higher bond, since that will not cause any sender to stop sending. Therefore, in equilibrium, two distinct bond values are not stable. ■

In this subcase, all sender types send to two distinct pools of recipients. In order to have two different pools, there must be two distinct bond values. By the lemma above, this cannot happen.

Subcase *ii* In this subcase, type $\mathcal{U}$ recipients have chosen to receive only type $\mathbb{B}$ messages (giving up the bond and message from type $\mathbb{G}$ senders), but type $\mathcal{B}$ recipients have chosen to give up the messages they like to get the bonds from type $\mathbb{G}$ senders. These choices are inconsistent.

The type $\mathcal{U}$ recipient's choice here is

$$(1-\gamma)(s_B-c_s+r_B-c_r) \geq \gamma((1-\epsilon_B)s_G-c_s+r_G-c_r)+(1-\gamma)((1-\epsilon_B)s_G-c_s+r_B-c_r) \tag{29}$$

The type $\mathcal{B}$ recipient's choice here is

$$\gamma\left(\frac{\epsilon_G}{\epsilon_G+\epsilon_B}s_G-c_s-c_r\right)+(1-\gamma)\left(\frac{\epsilon_G}{\epsilon_G+\epsilon_B}s_G-c_s+r_B-c_r\right) \geq (1-\gamma)(s_B-c_s+r_B-c_r) \tag{30}$$

These two equations together form a contradiction, so this situation cannot happen.

Subcase *iii* In this subcase, there exist times when a flat tax provides greater welfare than the ABM. To see this, let the fraction of type $\mathcal{U}$ recipients $(1-\epsilon_G-\epsilon_B)$ become arbitrarily small. However, this is consistent with the proposition, since in this case type $\mathcal{G}$ recipients are willing to not receive any message they like.

Subcase *iv* This situation is very similar to Subcase *ii*. There is an inherent contradiction in the choices of type $\mathcal{U}$ and type $\mathcal{G}$ recipients.

The type $\mathcal{U}$ recipient's choice here is

$$\gamma(s_G - c_s + r_G - c_r) \geq (1-\gamma)((1-\epsilon_G)s_B - c_s + r_B - c_r) + \gamma((1-\epsilon_G)s_B - c_s + r_G - c_r) \tag{31}$$

The type $\mathcal{G}$ recipient's choice here is

$$(1-\gamma)\left(\frac{\epsilon_B}{\epsilon_G + \epsilon_B}s_B - c_s - c_r\right) + \gamma\left(\frac{\epsilon_B}{\epsilon_G + \epsilon_B}s_B - c_s + r_G - c_r\right) \geq \gamma(s_G - c_s + r_G - c_r) \tag{32}$$

These two equations together form a contradiction. Therefore this situation cannot happen.

Subcase *v* In this subcase, as in Subcase *iii*, there exist times when a flat tax provides greater welfare than the ABM. However, this is consistent with the proposition, since in this case type $\mathcal{B}$ recipients are willing to not receive matching messages.

Subcases *vi* and *vii* These two subcases cannot happen. By comparing the rationality constraints for this case, it can be seen that it is irrational for universals to choose to receive messages from only one sender type when the recipients who like only that type are willing to forgo those same messages.

$\mathcal{U}$, $\mathcal{G}$, and $\mathcal{B}$ all pool This can again be divided into three distinct subcases, graphically depicted in Figure 2. Subcase *i* is the situation where all three recipients choose a common bond size, and that bond size is low enough that both senders are willing to send to everyone in the pool. Subcase *ii* is the situation where all recipients pool on a high bond, such that only type B senders can afford to send to the pool. Finally, Subcase *iii* is the analogous situation, but only type $\mathbb{G}$ senders can afford to send to the pool.

Subcase *i* clearly has welfare equal to W^0, the welfare in the baseline situation. This is because all possible transactions occur, both helpful and harmful. However, in this case, welfare under the ABM is greater than welfare under a flat tax because this case will only occur when a flat tax lowers welfare.

Let us assume that $(1-\epsilon_G)s_B - c_s \leq (1-\epsilon_B)s_G - c_s$. This implies that a flat tax would target and eliminate type $\mathbb{B}$ senders. In order for type $\mathcal{G}$ recipients to rationally choose this case, their surplus must be greater than that which they would get by choosing a high bond such that only type $\mathbb{G}$ senders could

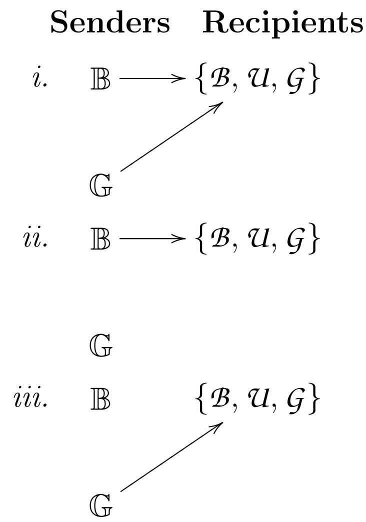


Figure 2: Subcases where all recipients pool together

reach them:

$$\gamma\left((1-\epsilon_G)s_B - c_s + r_G - c_r\right) + (1-\gamma)\left((1-\epsilon_G)s_B - c_s - c_r\right) > \gamma(s_G - c_s + r_G - c_r) \tag{33}$$

This equation and the assumption above together imply that

$$(1-\gamma)(1-\epsilon_G)(s_B - c_s + r_B - c_r) > ((1-\gamma)\epsilon_G + \gamma\epsilon_B)(c_s + c_r) \tag{34}$$

which implies that the tax harms welfare.

If we reverse the assumption, then the tax would target type $\mathbb{G}$ senders. Through an analogous derivation (starting with the $\mathcal{B}$ recipient's rationality constraint) a similar result (tax harms welfare) can be shown.

The intuition here is that in order for everyone to rationally choose a bond that keeps the type $\mathbb{B}$ senders in the market, the surplus gained from them must be enough to cover the costs of mistargeting caused by this.

Subcase *ii* cannot happen in equilibrium. The intuition here is that since type $\mathbb{G}$ senders cause more surplus by their transactions than type $\mathbb{B}$ senders do (by definition) and all recipients split the sender's surplus, at least one recipient type will prefer to set a bond such that type $\mathbb{G}$ senders can still receive them, since they have more surplus overall.

Subcase *iii* is one case where welfare from a flat tax clearly exceeds welfare from the ABM. In this case, recipients choose a bond equal to the flat tax, causing the type $\mathbb{B}$ senders to leave the market. However, unlike the tax, type $\mathcal{B}$ recipients remain in the market in order to extract some surplus in the form of bond payments. By staying in the market, they cause some mistargeting costs that do not exist under a tax, leaving lower total welfare.

However, note that in this situation, type $\mathcal{B}$ recipients are willing to forgo all messages they like in order to collect bonds from type $\mathbb{G}$ senders. This is consistent with the proposition.

References

Myerson, R and Satterthwaite, M, "Efficient Mechanisms for Bilateral Trading." *Journal of Economic Theory*, *29*, pp. 265–281, 1983.

Rasmusen, Eric, "Explaining Incomplete Contracts as the Result of Contract-Reading Costs." *Advances in Economic Analysis and Policy*, *1*(1), 2001.

van Zandt, Timothy, "Information Overload in a Network of Targeted Communication." *RAND Journal of Economics*, *35*(3), pp. 542–560, 2004.