

An Economic Response to Unsolicited Communication

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

August 29, 2005

Abstract

We investigate welfare effects of mechanisms designed to improve total communications value. Identifying first contact information asymmetry and negative externalities as the critical problems, we analyze best-in-class solutions from law, technology, and economics. Comparison leads to several useful conclusions. First, mechanisms designed to promote valuable communication can outperform those designed to block wasteful communication. The best mechanism can, on occasion, outperform even a “perfect filter.” Second, it is advantageous to shift focus from message content to senders’ private knowledge. Recipients can then use information revelation mechanisms to force people who knowingly misuse communication to drop out or incur higher costs. Third, private knowledge of message value and the choice to communicate might favor senders over receivers. These advantages disappear, however, under a take-it-or-leave-it offer by recipients to refuse non-conforming communications. Fourth, giving recipients rights in their own attention can improve willingness to signal their preferences, which facilitates efficient sender targeting. ¹

*tloder@umich.edu

[†]mva@bu.edu, marshall@mit.edu

[‡]rwash@umich.edu

¹This material is based upon work supported by the National Science Foundation under Grant No. 0114368.

Contents

1	Introduction	1
2	Prior Literature and Media Assumptions	3
2.1	Technological Solutions	3
2.2	Regulatory Solutions	4
2.3	Market Solutions	5
2.4	Media Class Assumptions	7
3	Modeling the Mechanisms	8
3.1	Comparison of Mechanisms	10
3.2	Baseline Case - Null Mechanism	11
3.3	Perfect Filter	12
3.4	Attention Bond Mechanism	13
4	Single Distribution Results	14
4.1	Social Welfare	15
5	Tax Comparison	17
5.1	Baseline	18
5.2	Tax	19
5.3	ABM	21
5.4	ABM Versus Tax	24
5.5	Additional Advantages	25
6	Why not Always Seize the Bond?	26
7	Conclusions	28

1 Introduction

Communications technology brings both blessing and curse. It extends our relationships and our influence, draws remote information closer, and alters our time via batching and asynchrony. It also helps total strangers reach us. Unsolicited contact can be desirable but need not be. If contact requires our attention and strangers ignore our opportunity costs, their choices can reduce our welfare. Perhaps the most well known and timely example is spam. The problem of unsolicited contact, however, is pervasive, affecting numerous media including text and instant messaging, telephone, and standard post.

Sender choices produce welfare losses through externalities and information asymmetry. “Call” externalities arise because senders internalize neither recipient gains nor losses from processing messages. The decision to send, which produces call externalities, is distinct from the network participation decision, which produces “access” externalities, the demand economies that scale in total contacts.

A two-sided information asymmetry compounds the communications problem. Before first contact, recipients do not know sender qualities, and typically, senders do not know recipient tastes. Without ability to target or need to address call externalities, senders bombard recipients to gain their attention.

Unlike senders, message recipients largely face a choice of whether to participate in a communications network, not a choice of whether to receive a specific message. This arrangement favors senders who are privileged to decide whether to communicate and what to convey.

Designing a mechanism to correct for call externalities leads to considering a Pigouvian tax. Assuming access to information on sender quality and recipient tastes, a central planner could force senders to internalize costs they impose on others. Naturally, we might also consider the Coase Theorem and grant, for example, a recipient property right in attention. Ideally, pricing attention rights should restore efficient transaction levels. Alone in this context, however, the latter solution is untenable.

Not only can high transactions costs cause market failure but, more importantly, traditional economic solutions that assume zero cost negotiations are misapplied. Negotiation requires communication, and unwanted communication is itself the problem. One way forward might be to put information in the market that promotes sender screening and recipient signaling, with minimal need to negotiate.

This paper describes and compares alternative mechanisms for raising the value of communication. We distinguish each mechanism by its characteristics: whether it is decentralized or imposed centrally, how it handles negotiation,

whether it supports side payments, whether it distinguishes first from subsequent contact, what information it provides decision makers, and when such information is revealed.

Our analysis considers three idealized mechanisms beyond the null intervention baseline: These are (1) an efficient Pigouvian tax that redistributes all proceeds, (2) a costless “perfect filter” that exhibits neither false positives nor false negatives, and (3) a new proposal, an “Attention Bond Mechanism” (ABM) designed to signal preferences while screening low value content.

The ABM facilitates bilateral distributed choice. While it affords a more limited veto than the perfect filter, it introduces both recipient private knowledge of preferences and sender private knowledge of message match to the sender’s decision to send. Its joint signaling and screening properties imply that social welfare under the ABM frequently, though not always, exceeds that of the Pigouvian tax and the perfect filter.

A substantial source of leverage in our analysis is the point of intervention. The tax affects all communication, regardless of local information. The perfect filter grants recipients a message level veto but provides little information to the market. In contrast, the ABM focuses on first contact among strangers. Subsequent contact can then depend on revealed information.

Comparison leads to several observations. First, mechanisms designed to promote valuable communication can outperform those designed merely to block wasteful communication. The best mechanism can, on occasion, outperform the ideal tax and the perfect filter. Second, 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. Third, although private knowledge of message value might favor senders over receivers, this advantage can disappear under a take-it-or-leave-it off by recipients to refuse non-conforming communications. Fourth, giving recipients rights in their own attention can motivate them to signal their preferences in ways that facilitate efficient sender targeting.

The remainder of this paper proceeds as follows. Section 2 states our assumptions on media properties and reviews previous work in this area. Section 3 develops a model of initial communications, and distills three theoretical mechanisms for dealing with unsolicited communication. Section 4 compares the ABM to the perfect filter and shows how screening and wealth transfers are possible. Sections ?? and ?? expand the screening model and relax assumptions. Then Section 5 introduces signaling and compares the ABM to an efficient tax. Following the analysis, Section 6 addresses implications. Finally, Section 7 concludes.

2 Prior Literature and Media Assumptions

The numerous approaches for dealing with unwanted communications fall roughly into three categories: technology, law, and market based. This section reviews proposed solutions then generalizes properties from each into a set of assumptions regarding potential interventions. Throughout, *ex ante* intervention refers to the period before a recipient learns message content.

2.1 Technological Solutions

Technological solutions include several types of content filtering, whitelists, and blacklists (Cranor and LaMacchia, 1998). Rule-based and Bayesian filters use deductive and probabilistic inference to classify spam *ex ante*. Community filters harvest the classification efforts of a few participants *ex post* to remove spam from the inboxes of other participants *ex ante*. Regardless of method, messages judged as spam are typically discarded.

Whitelists distinguish non-spam based on prior knowledge of sender identity, and allow only messages from positively identified senders to pass. In contrast, messages from blacklisted senders are deleted or sidelined. Blacklists are generally ineffective due to the ease of obtaining cheap new identities. Whitelists block spam, but without a way around the whitelist censor, the sender is forced outside the medium, often at much higher cost, to request permission to send.

A promising and essential element of whitelists is authentication through digitally signed addresses (Tompkins and Handley, 2003). These limit “spoofing,” that is, deliberate misuse of third party addresses to gain access. Consistent with media class assumptions below, address verification will inevitably become part of any realistic solution.

Verification, though necessary, is not sufficient. The ease of obtaining new addresses, verifiable or not, allows unscrupulous senders to refresh their identities. Friedman and Resnick (2001) show that newcomers inevitably need to “pay their dues” in any open society — one that does not charge per access — for precisely this reason.

There are several reasons why classification has proven so intractable from a technology perspective. First, the space of identifiable message permutations is much larger than generally realized. Using only one-letter substitutions, misspellings, and spacing, one estimate gives more than 6×10^{20} variations for a single six letter word, the anti-impotence drug Viagra.² Second, the

²<http://cockeyed.com/lessons/viagra/viagra.html> examples include v1agra,

expressiveness of natural language lets senders use creative metaphors to bury obvious connotation well below their literal statements.³

Finally, technological classification suffers from serious problems of false positives and false negatives. Such imperfect filtering permits an arms race. Internet service providers report that spammers have responded to changes in filters in as few as two hours (Libbey, 2004). In fact, computer scientists have argued persuasively that the arms race is unwinnable (Graham-Cumming, 2004). Assume, for purposes of contradiction, that filtering technology exists capable of identifying unwanted communications based on any finite set of message attributes. Then, in the hands of an adversary, that same technology can identify communications in the complement of that set. If it is possible to embed an arbitrary message in the set of identifiable complements, then success of the filtering technology implies success of the filter penetration technology.

Regardless of method, *ex ante* filtering technologies create incentives to lie. They motivate senders of unwelcome messages to forge message attributes, offering red herrings as means to penetrate filters. In contrast, *ex post* verification mechanisms with a binding commitment create an incentive to tell the truth. Knowing that claims will be verified, sender willingness to incur a risk based on these claims signals their belief in accurate content.

2.2 Regulatory Solutions

The intent of legal approaches is to regulate behavior. Various laws have been proposed to tax communication, ban inappropriate content, force identification tags, create opt-out lists, and impose criminal penalties on illicit activity. One legal scholar has even proposed a federally supported bounty on criminal spammers [cf. Lessig in (Bazeley, 2003)].

Motivated by the success of the the Telephone Consumer Protection Act of 1991, which, after revision in 2003, established do-not-call and do-not-fax registry (maintained by the Federal Trade Commission), the CAN-SPAM Act requested a formal recommendation by the FTC regarding the establishment of a do-not-spam registry. The FTC recommended against the creation of the list. Other aspects of the CAN-SPAM act aimed at curbing spam took effect Jan. 1, 2004. Regrettably, the volume of spam is undiminished (Rainie and Fallows, 2004).

Although a detailed analysis of specific legislation is outside the scope of this paper, legislation alone is likely to fail for several reasons: high costs

vi@gra, v:i:a:g:r:a, viagorea, viatgra, V-!.a-g*r-a, via6ra, and ViagrYa, just to name a few.

³The filters for one coauthor passed through an invitation to “get a rod like a firehose.”

of enforcement, lack of incentive compatibility, issues of jurisdiction, and a legal one-size-fits-all definition of nuisance. In terms of measurable harm, laws banning certain forms of communication altogether simultaneously eliminate potential transactions resulting from more targeted communication. While these laws have specific exemptions for non profit, political action, and research groups, they provide no recourse to individuals for the nuisance costs such interruptions create.

One of the best regulatory analyses is provided by van Zandt (2004). This research identifies “information overload” as overexploitation of scarce recipient attention by senders who crowd out each other’s competing messages. Low transmission costs permit senders to increase chances of having even low value messages read by bombarding recipients. A single sender who internalized this externality would self-limit, preferring to send only his highest value messages. Van Zandt demonstrates that even a tax policy that destroys tax proceeds can enhance welfare when the value of forcing senders to target exceeds the cost of the tax. Rebating the tax improves welfare further.

Shortcomings of a tax include voluntary participation, misrepresentation, and inflexibility. Participants can voluntarily switch protocols or media if one format is disadvantaged by a tax surcharge and others not, since taxing all forms of bit transfer can impose high efficiency losses. Under a tax system, senders also retain the ability to lie about content *ex ante* in order to elevate interest. Such blunt instruments do not adjust to individual preferences, and while they do eliminate wasteful messages, they also eliminate valuable messages for recipients that are only weakly valuable to senders. In contrast, a first contact mechanism that distorts only a fraction of all communication, yields significant efficiency gains.

2.3 Market Solutions

A few articles have explored market based mechanisms for pricing messages and allocating recipient attention (Kraut et al., 2003; Dwork and Naor, 1993; Fahlman, 2002; Hermalin and Katz, 2004; van Zandt, 2004). Suggested means include transferable stamps, challenge-response, selling “interrupt rights,” and using auctions. The essential idea is that raising sender costs shifts the burden of screening from recipients back to senders.

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. Using a property rights approach, our proposal takes a step toward allowing participants to internalize call externalities. Efficiency gains can then lead, in certain limited circumstances, to first-best welfare.

Challenge-Response (CR) systems with CAPTCHAs (Completely Automated Public Turing Test to Tell Computers and Humans Apart (von Ahn et al., 2003)) or computational challenges achieve a crude form of negotiation and pricing. All use whitelists. Under CR, messages from unrecognized senders are held temporarily while the senders receive an auto-response directing them to pass a test. CAPTCHAs are usually human cognition tests such as pattern recognition. Computational puzzles require proof of performing a costly computation, which consumes computer rather than human time (Dwork and Naor, 1993). If a sender passes the test, the system delivers the original message and whitelists the sender. Otherwise, it destroys the message.

Both forms of CR resemble a tax on senders with certain differences. Being decentralized, they can adjust to the opportunity costs of recipients. As a limited form of negotiation, they can distinguish between first and subsequent contact. But, relying on nontransferable utility, they cannot make side payments. They also return no information to the market. Computational challenges are also susceptible to harnessing “zombie” machines to expand the resources of illegitimate senders (Laurie and Clayton, 2004). CAPTCHAs are frequently inverted. Senders embed messages in images that only a human can recognize, since recipients’ agents cannot themselves pass Turing tests.

A more attractive mechanism for allocating attention, that of “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 limited recipient attention, van Zandt (2004) points to efficient use of a Vickrey auction. Each approach incorporates the assumption that attention itself is a valuable resource, and communications media such as email, the telephone, and instant messaging allow the sender to interrupt a recipient to divert his attention. This serves as a starting point for our economic analysis.

While technological and regulatory approaches limit unwanted communications by blocking or banning them, we build on Coase’s insight that granting property rights under a system that facilitates valuable exchange through side payments generally dominates most other solutions to externality problems. It can also dominate systems, such as filters, that grant only unilateral veto to either party. Recognizing this problem, our added contribution is to connect

Coase's solution to a solution for two-sided information asymmetry among senders and receivers, and to demonstrate how this improves welfare. Importantly, property rights under this proposal do 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).

2.4 Media Class Assumptions

To establish points of intervention, we distill several common properties from each of the above proposals. Communications phenomena underlying our analysis can most intuitively be described as *multi-party point-to-point media with cheap identities and imperfect information*. Message values vary by participant, and the number of participants is large. More precisely, we consider the following five properties.

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, 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.

In general, these properties lead to timing differences for each mechanism. Technology approaches limit call externalities by *ex ante* message classification and destruction. Regulatory approaches accomplish this by applying *ex ante* tariffs to screen low value messages (from a sender perspective). They also apply *ex post* criminalization and fines to penalize low value or offensive content (from a recipient perspective). Market methods apply *ex ante* screens to separate and price messages based on sender knowledge. In the case of the ABM, *ex post* verification is then performed to ensure compliance.

Media exhibiting these properties include Internet email, SMS (text messaging), IM (Instant Messaging), and POTS (Plain Old Telephone System). To a lesser extent it also applies to common carrier letters.

3 Modeling the Mechanisms

We now introduce a simple model of individual message value. This model allows us to derive the participation value to senders and receivers, the social welfare contribution of message exchange, and to compare the effects of the different mechanisms. Initially, we consider a single sender and a single recipient, later extending the model to heterogeneous sender and receiver types. For clarity of exposition, yet acknowledging the unintentional associations that may arise from the pages that follow, we use male pronouns for senders and female pronouns for recipients.

We assume the sender chooses to initiate communication, incurring a per-message delivery cost, c_s , which can be arbitrarily small. Although a sender pays for every message sent, he receives a private value s if and only if a recipient reads it.

If a message reaches a recipient's unprotected inbox, we assume that she will read or otherwise process it. Doing so yields her a private value r , at a per-message cost c_r . A recipient cannot know her own value r until she has actually read the message. If the message is blocked, the recipient avoids the processing cost c_r and does not realize value r .

We assume stochastic message values for senders and recipients. Given a specific sender and recipient pair, let s have range $s \in [\underline{s}, \bar{s}]$ and r have range $r \in [\underline{r}, \bar{r}]$. For simplicity, we assume the uniform distribution, although our results do not depend on this choice. Before first contact, r and s are private information, implying each is subject to misrepresentation. Initially, they are uncorrelated. Section 5 explores the correlated case. Minimum values may be arbitrarily low, and may be negative, representing offensive communication for a recipient. Senders require $\bar{s} - c_s > 0$; otherwise they send no messages.

Per-message welfare contribution is defined as the sum of sender and recipient surplus, or gross message value minus gross cost: $W = SS + RS = (s - c_s) + (r - c_r)$. Thus, a message yields positive welfare when $W > 0$.

Figure 1 is a visual representation of the sender and recipient message value distribution. The vertical axis represents the value to the recipient; the horizontal axis is the message value to the sender. The rectangular region bounded by $\bar{s}, \underline{s}, \bar{r}, \underline{r}$ is the area of feasible message values. The horizontal line c_r represents the recipient processing costs. Any message having value

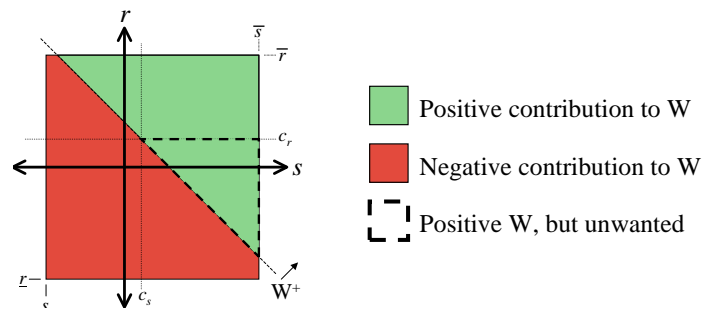


Figure 1: Message Value Distribution

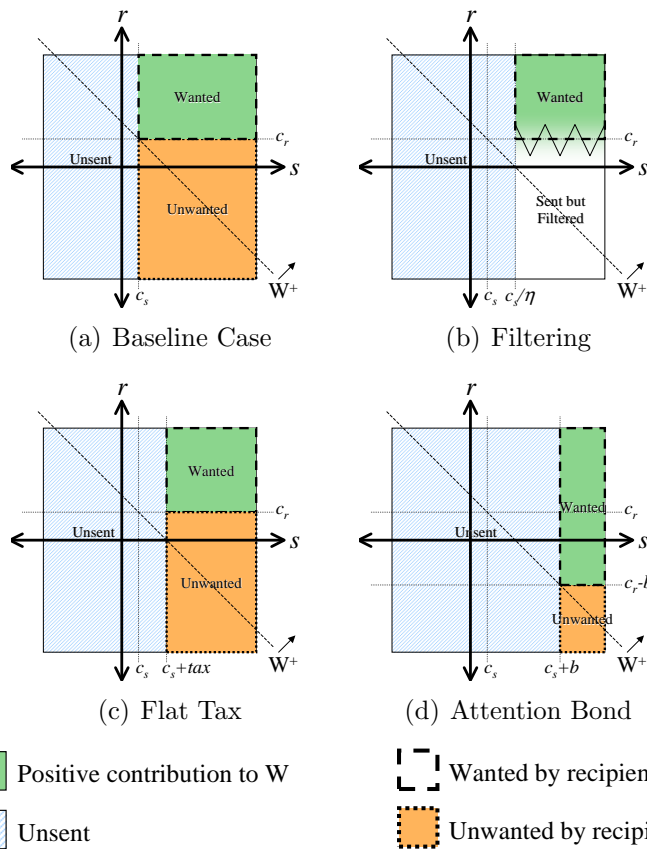


Figure 2: Comparison of mechanisms based on the distribution of message values.

r above this line is considered desirable by a recipient, since her net gain is positive. Similarly, the vertical line c_s represents the sender's cost to send, and any message to the right of c_s yields positive surplus. The diagonal line, labeled "W," is the "welfare line" based on the definition of welfare above. Messages in the region northeast of this line make a positive contribution to social welfare, and exchange of such messages is desirable from the standpoint of a social planner.

In Figure 2-a, the triangular unwanted region, below c_r , identifies messages that make a positive welfare contribution, yet which the recipient prefers not to receive. A similar triangular region, left of c_s , identifies messages that make a positive contribution to welfare but which the sender prefers not to send. To the extent that mechanisms facilitate the exchange of messages in either region, welfare is improved. As we shall show, some mechanisms fare better than others in this region, though information asymmetry makes it difficult to recover the welfare positive-messages perfectly.

A sender has interim individual rationality, implying he can decide whether or not to send an individual message based on his own self-interest. Having composed the message, he knows s , and sends whenever $s - c_s$ is positive. The recipient, in contrast, has only *ex ante* rationality. She can decide beforehand to participate or not, but not knowing r , she cannot choose individual messages without additional information.

3.1 Comparison of Mechanisms

In Section 2, we discussed three general classes of mechanism for communications media. To get a sense of how each approach compares, we have chosen three for analysis which we feel are readily understood, best represent their class, and are commonly cited as possible solutions in the literature. These three mechanisms are illustrated in Figure 2, showing the distribution of message values.

The mechanisms are a hypothetical "perfect filter" that represents an upper bound on the effectiveness of any filtering solution, a centrally imposed efficient Pigouvian tax that corrects for call externalities, and the Attention Bond Mechanism (ABM), a protocol that allows a recipient to inexpensively negotiate (make a take-it-or-leave-it offer) first contact communications and that facilitates bi-directional wealth transfers.

The tax affects all messages sent. The size of the tax is the only variable, and must be decided globally, since the central authority is assumed to lack access to recipient and sender private information. Raising the cost of communicating via a tax causes senders to send fewer messages and focus on those

which are net positive. The optimal tax depends on the size of externalities, driven by the distribution of sender and receiver types. For completeness, we show tax effects here based on the distribution of values but defer the main tax analysis until Section 5, which expands the distribution of types.

The perfect filter is an idealized representation of a pure classification solution. As a proxy, it uses local recipient preferences to reject messages considered undesirable. There is no cost to the recipient but also no negotiation or wealth transfer. Senders of rejected messages experience transmission costs but no benefit.

Under the ABM, recipients post a take-it-or-leave-it bond “price” for accepting first contact messages from strangers. Amount varies by recipient, and individual bond requests are public information that signals recipient expected value or opportunity cost. Messages associated with amounts less than that requested are rejected. A reluctance to post a bond screens senders whose private information suggests their messages have low expected value. In effect, the bond serves as a “transactions hostage” (Williamson, 1985) that reduces adverse selection. Such a simple contract avoids negotiation and reading costs (Rasmusen, 2001).

After learning message content, recipients may, at their sole discretion, release the bond and whitelist the sender. Seizing the bond effects a wealth transfer. Releasing the bond conveys further information about a recipient’s true type. Our initial analyses assume recipients keep first contact bonds. Section 6 shows why this frequently may not happen.

3.2 Baseline Case - Null Mechanism

To derive results in the baseline case, we need no further assumptions. Due to his knowledge of expected message value s , a sender sends messages where he expects $s - c_s > 0$. This can be seen graphically as the “Unsent” region to the left of c_s in Figure 2-a. Sender surplus and recipient surplus are

$$SS_0 = k \cdot \int_{\underline{r}}^{\bar{r}} \int_{c_s}^{\bar{s}} (s - c_s) ds dr, \quad (1)$$

$$RS_0 = k \cdot \int_{\underline{r}}^{\bar{r}} \int_{c_s}^{\bar{s}} (r - c_r) ds dr \quad (2)$$

Both results have factor k as a function of the support parameters, where

$$k = \frac{1}{\bar{s} - \underline{s}} \cdot \frac{1}{\bar{r} - \underline{r}} \quad (3)$$

Evaluating Equations 1 and 2 yields $SS_0 = k \cdot \left(\frac{\bar{s}+c_s}{2} - c_s\right)$ and $RS_0 = k \cdot \left(\frac{\bar{r}+r}{2} - c_r\right)$, representing the mean value net of costs in both cases.

3.3 Perfect Filter

The perfect filter costlessly intercepts and deletes certain messages before they reach recipients. A message is deleted if and only if a fully informed recipient would not want the message based on message value. The perfect filter makes no mistakes in classification, generating neither false positives nor false negatives. Accordingly, the perfect filter provides interim individual rationality for the recipient, giving her the same balanced veto on participation in any single communication. Figure 2-b shows the rectangular surplus left by two-party veto – a sender only sends what he values, and a recipient only accepts what she values. The jagged line around c_r also illustrates an imperfect filter that does admit false positives and false negatives.

Since the rational recipient decision is to block messages of value $r < c_r$, the contribution to recipient surplus for such a message is zero. By our previous assumption, senders have no knowledge of recipient value - only their own expected value for accepted messages. For each message sent but not received, the sender incurs c_s , but gets zero payoff.

For the perfect filter, we allow the senders to adjust the expected value of s dynamically for use in decision making. This setup is consistent with observed spam sender behavior. A sender notices that only a fraction η of his messages reach recipients. The sender suffers a corresponding reduction in per-message value from the original expected s to $\eta \cdot s$. While at first a sender chooses to send messages when $s > c_s$, application of the perfect filter changes the sender threshold to $\eta \cdot s > c_s$. Equivalently, higher marginal costs per message delivered imply a choice of sending only when $s > \frac{c_s}{\eta}$. The filter can be defined endogenously as $\eta = \frac{\bar{r}-c_r}{\bar{r}-r}$.

Under these assumptions, total sender surplus for the perfect filter is

$$SS_{PF} = k \cdot \int_{c_r}^{\bar{r}} \int_{\frac{c_s}{\eta}}^{\bar{s}} (s) ds dr - k \cdot \int_{\underline{r}}^{\bar{r}} \int_{\frac{c_s}{\eta}}^{\bar{s}} (c_s) ds dr \quad (4)$$

Likewise, the total recipient surplus is

$$RS_{PF} = k \cdot \int_{c_r}^{\bar{r}} \int_{\frac{c_s}{\eta}}^{\bar{s}} (r - c_r) ds dr \quad (5)$$

In figure 2-b, the sender cost threshold shifts right, since $c_s/\eta > c_s$, reducing the space of viable messages. The white “Sent but Filtered” region is

now pure social loss since the sender pays for these undelivered messages, but neither party realizes positive value from them. Note also the reduction in the “Wanted” region, as compared to the baseline case (Figure 2-a). Recipients lose some contact that might have been valuable, since the sender never sends them.

3.4 Attention Bond Mechanism

Surplus under the ABM is depicted in Figure 2-d. Like the tax, the ABM alters the sender’s cost basis and thus his decision to send. This is $s - c_s - t$ in Figure 2-c versus $s - c_s - b$ in Figure 2-d. We assume quasi-linear utilities and that prior message values r and s remain independent of the tax or bond. The ABM subsumes the tax, however, in allowing direct sidepayments to specific recipients. The figure shows this as a lowering of the recipient’s cost basis, which increases her willingness to accept messages. The ABM also differs from the tax in allowing recipients to seize or release b according to their own preferences.

For later analyses, senders do not necessarily know recipient values or seize policies but can estimate seize rates across the market. If the starting bond price is ϕ and the expected seize rate is p , then the expected bond is $b = p\phi$. Initially, we assume $p = 1$ then define rational policies that justify bond release in later sections.

Recipients choosing first contact bonds know nothing about individual sender quality. The default starting bond ϕ cannot therefore be tailored to sender type except through *ex post* policy p . Recipients choose a single ϕ *ex ante* that becomes public information. In contrast to messages sent under a perfect filter, all messages sent with adequate bonds are actually received.

To determine sender and recipient surplus under these conditions, we again integrate over the joint payoffs for each party:

$$SS_{ABM} = k \cdot \int_{\underline{r}}^{\bar{r}} \int_{c_s+b}^{\bar{s}} (s - c_s - b) ds dr \quad (6)$$

and

$$RS_{ABM} = k \cdot \int_{\underline{r}}^{\bar{r}} \int_{c_s+b}^{\bar{s}} (r - c_r + b) ds dr \quad (7)$$

4 Single Distribution Results

For clarity, we begin analysis where the sender and recipient draw from a single distribution. The recipient's first choice is to calculate her optimal bond.

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

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

Proof: Evaluate the right hand side of Equation (7) to produce

$$\frac{\bar{s} - c_s - b}{\bar{s} - \underline{s}} \left(\frac{\bar{r} + r}{2} - c_r + b \right) \quad (8)$$

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

This represents the monopoly price for recipient attention (competition is briefly explored in Section 6). The optimal bond is the expected sender surplus minus half the recipient surplus. If kept, the bond represents an individually targeted Pigouvian tax on senders, adjusted for 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 $\tilde{s} \equiv \frac{\bar{s} - c_s}{2}$ and $\tilde{r} \equiv \frac{\bar{r} + r}{2} - c_r$ with the difference explained by *interim* versus *ex ante* rationality. A positive bond $\tilde{s} > \frac{\tilde{r}}{2}$ 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 under the baseline case, $RS_{b^+} \geq RS_0$. Since $b^+ = 0$ reproduces the baseline case, the added degree of freedom ensures the left hand side weakly dominates the right.

More importantly, for certain distributions, the Attention Bond Mechanism can actually perform better than a perfect filter.

Proposition 1 *Using the ABM creates greater recipient surplus than using a perfect filter, $RS_{b^+} \geq RS_{PF}$, if and only if the optimally chosen bond is such that*

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

Proof: Apply Lemma 1 and $\eta = \frac{\bar{r} - c_r}{\bar{r} - r}$ to the inequality $RS_{b^+} \geq RS_{PF}$, using Equations 7 and 5 to produce the following expression:

$$\frac{(\bar{r} - \underline{r}) \left(\bar{s} - c_s + \left(\frac{\bar{r} + \underline{r}}{2} - c_r \right) \right)^2}{(\bar{r} - c_r) (\bar{r} (\bar{s} - c_s) - \bar{s} \cdot c_r + \underline{r} \cdot c_s)} \geq 2 \quad (9)$$

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 receiver, where the left hand side of the inequality represents the bonded surplus, and the right hand side is the perfect filter surplus. Reader surplus for the ABM is an average value reflecting the difference in IR constraints — senders participate for any message value above c_s , 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 cost. That there exist values for which Equation 9 is true illustrates the point that facilitating exchange can dominate giving veto power.

While Equation 9 is not always true, it is in several cases. The test condition is more easily satisfied as recipients block more messages ($\bar{r} \rightarrow c_r$) and as senders have more surplus to transfer ($\frac{d(SS)}{d\bar{s}} = 2 \left[(\bar{s} - c_s) + \left(\frac{\bar{r} + \underline{r}}{2} - c_r \right) \right] - 2\eta(\bar{r} - c_r)$ grows quickly as \bar{s} grows). For similar reasons it is also true for extreme values of \bar{r} and \underline{r} , a fact that we explore further in Section ???. 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. Lost friends and telemarketers may draw from different distributions, however, implying participation if the sum across distributions is positive. We also explore the screening of different sender types in Section ???.

4.1 Social Welfare

Having considered recipient benefit from using the ABM, the question remains whether its use improves total social welfare.

Proposition 2 *The total social welfare with a recipient-chosen bond is greater than that of a perfect filter $W_{b^+} > W_{PF}$ if and only if the following is true:*

$$\left[(\bar{s} - c_s) + \left(\frac{\bar{r} + \underline{r}}{2} - c_r \right) \right]^2 \geq 3\eta \left(\bar{s} - \frac{c_s}{\eta} \right) \left[(\bar{r} - c_r) + \left(\bar{s} - \frac{c_s}{\eta} \right) \right] \quad (10)$$

Proof: To prove, construct the inequality $SS_{b^+} + RS_{b^+} \geq SS_{PF} + RS_{PF}$ using Equations 4, 5, 6, 7, Lemma 1, and evaluate:

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

$$k \cdot \int_{c_r}^{\bar{r}} \int_{\frac{c_s}{\eta}}^{\bar{s}} ((r - c_r) + (s - c_s)) ds dr - k \cdot \int_{\underline{r}}^{\bar{r}} \int_{\frac{c_s}{\eta}}^{\bar{s}} (c_s) ds dr$$

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 surplus for either senders or receivers. The left hand side gives ABM surplus terms; the right gives terms for the perfect filter. Similar to Equation 9, this 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*

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

Proof: To find W_b , use Equations 6 and 7 and FOC's with respect to b . ■

Note that b^* represents total recipient surplus, and is unrelated to sender surplus. In effect, b^* is a Pigouvian *recipient* tax. Given the goal of minimizing wasteful communication, taxing recipients appears counterintuitive. The result, however, is highly socially efficient. Senders know the expected value of r but have full information on s . In contrast, recipients have full information on neither. The reverse Pigouvian bond 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^* produces greater social welfare than the perfect filter.

Corollary 1 *The total social welfare with a socially optimal bond is greater than that of a perfect filter, $W_{b^*} > W_{PF}$, if and only if the following is true:*

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

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. The b^+ result, however, is surprisingly close to the b^* result. Proximity is valuable since a reverse Pigouvian tax, collected *ex ante*, could introduce moral hazard in message quality.

5 Tax Comparison

The technological complexity of the many-to-many ABM raises the question of how far a simple tax will go toward the social optimum. Here, we analyze the effect of such a tax on communications. Our previous model is capable of modeling a flat tax on the distribution of message values, but does not capture recipient heterogeneity or the ability to drive unwanted senders from the market. This limitation rests on two related assumptions: a sender knows his private value s before sending, and values realized by senders and receivers are independent.

Here we expand the model that allows us to better analyze effects of heterogeneity on both sides of the market and also market foreclosure. We continue to focus on first contact communications. This affords us a simpler comparison with the ABM. It also favors the apparent effectiveness of the tax, which would distort all communications, in contrast to the ABM, which distorts only first contact.

For purposes of matching, 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 value messages sent from both \mathbb{G} and \mathbb{B} senders. These are type \mathcal{U} recipients. 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 avoid this market.

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.

Correlated values arise from a successful match. Consider first type \mathcal{G} recipients. On receipt of a \mathbb{G} message, \mathcal{G} realizes r_G and \mathbb{G} realizes s_G . In contrast, if a \mathcal{G} recipient opens a \mathbb{B} message, then both parties realize 0. Given

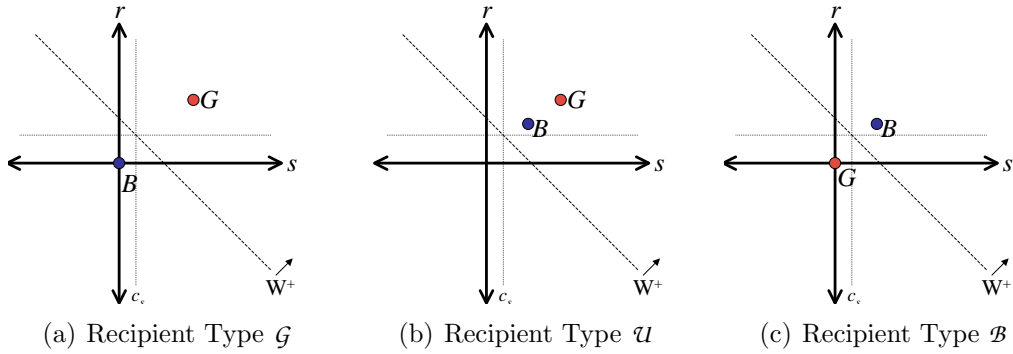


Figure 3: Sender and receiver values by matching type. Note type \mathcal{G} recipients have mass ϵ_G , \mathcal{B} recipients have mass ϵ_B , and \mathcal{U} recipients have mass $(1 - \epsilon_G - \epsilon_B)$. Senders \mathbb{G} and \mathbb{B} have masses γ and $(1 - \gamma)$.

that participants know only expected values *ex ante*, we restrict attention to $r_G = E[r_G^0]$ and $s_G = E[s_G^0]$ with parallel expressions for other types. 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 3 shows relative payoffs.

Note that sender knowledge departs from the previous model. 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 *ex ante* and instead know only expected payoffs. We assume that senders have no relevant capacity constraint (they participate in the baseline), so it is feasible to send to everyone. A sender capacity constraint is briefly considered in Section 6.

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 (no modification) case, the ideal tax case, and the Attention Bond Mechanism.

5.1 Baseline

To illustrate the nature of the problem, we start with baseline communications where no mechanism is imposed to improve value. Unable to distinguish among recipients before first contact, the two sender types can expect the following surplus:

$$\text{SS}_G^0 = (1 - \epsilon_B) \cdot s_G - c_s \quad (11)$$

$$\text{SS}_B^0 = (1 - \epsilon_G) \cdot s_B - c_s \quad (12)$$

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 will receive only positive value from matched messages but will incur costs for all messages they read.

$$\text{RS}_U^0 = \gamma \cdot (r_G - c_r) + (1 - \gamma) \cdot (r_B - c_r) \quad (13)$$

$$\text{RS}_G^0 = \gamma \cdot (r_G - c_r) + (1 - \gamma) \cdot (-c_r) \quad (14)$$

$$\text{RS}_B^0 = \gamma \cdot (-c_r) + (1 - \gamma) \cdot (r_B - c_r) \quad (15)$$

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

Combining these equations leads to total baseline welfare.

$$\begin{aligned} W^0 &= (\gamma \text{SS}_G + (1 - \gamma) \text{SS}_B) \\ &\quad + ((1 - \epsilon_G - \epsilon_B) \text{RS}_U + \epsilon_G \text{RS}_G + \epsilon_B \text{RS}_B) \\ &= (\gamma(1 - \epsilon_B) s_G + (1 - \gamma)(1 - \epsilon_G) s_B - c_s) \\ &\quad + (\gamma(1 - \epsilon_G) r_G + (1 - \gamma)(1 - \epsilon_B) r_B - c_r) \quad (16) \\ &= \gamma(1 - \epsilon_B)(s_G - c_s + r_G - c_r) \\ &\quad + (1 - \gamma)(1 - \epsilon_G)(s_B - c_s + r_B - c_r) \\ &\quad - \gamma \epsilon_B (c_s + c_r) - (1 - \gamma) \epsilon_G (c_s + c_r) \end{aligned}$$

In the final expression, the three principle 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 from misdirected communications. Waste results from two negative externalities, misdirected \mathbb{G} messages and misdirected \mathbb{B} messages.

5.2 Tax

To improve welfare, consider a government-issued tax t on all messages sent. The theory behind this tax is that it can either compensate recipients for negative call externalities via sidepayments or 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{B} communications is lower than that of type \mathbb{G} . That is,

$$\gamma(1 - \epsilon_B)(s_G - c_s + r_G - c_r) > (1 - \gamma)(1 - \epsilon_G)(s_B - c_s + r_B - c_r) \quad (17)$$

This implies that social welfare improves if type \mathbb{B} senders leave the market. This tax cannot be effective, however, unless it is also true that

$$(1 - \epsilon_G)s_B - c_s < (1 - \epsilon_B)s_G - c_s \quad (18)$$

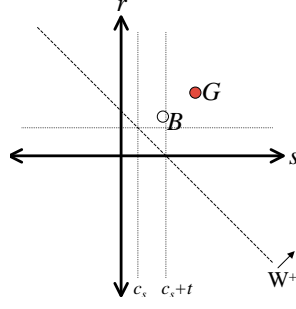
Individual \mathbb{B} senders have less transaction surplus than \mathbb{G} senders. If this were false, then a tax would either eliminate only type \mathbb{G} senders, or eliminate all senders. As such, the optimal tax is

$$t = SS_B^0 = (1 - \epsilon_G)(s_B) - c_s \quad (19)$$

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 a Pigouvian tax. It is set at the level of positive call externalities for individual 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)⁴. In terms of waste, we are primarily concerned with the second externality, though the first one also disappears. No longer getting messages of any value, type \mathcal{B} recipients also leave the market.

⁴As the number of sender types n increases, the number of such externalities grows as $O(n^2)$. These interactions imply that a single Pigouvian tax quickly has difficulty compensating for all possible call externalities.


 Figure 4: Tax Effect on R_U and the Senders

Each party now has a new surplus:

$$SS_G^t = (1 - \epsilon_B) \cdot (s_G - c_s - t) \quad (20)$$

$$SS_B^t = 0 \quad (21)$$

$$RS_U^t = \gamma \cdot (r_G - c_r) \quad (22)$$

$$RS_G^t = \gamma \cdot (r_G - c_r) \quad (23)$$

$$RS_B^t = 0 \quad (24)$$

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

$$\begin{aligned} W^t &= \gamma SS_G^t + (1 - \epsilon_B - \epsilon_G) RS_U^t + \epsilon_G RS_G^t + \gamma(1 - \epsilon_B)t \\ &= \gamma(1 - \epsilon_B)(s_G - c_s + r_G - c_r) \end{aligned} \quad (25)$$

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 16). Second, it also eliminates all benefits of type \mathbb{B} transactions (the second-to-last line of Equation 16). The tax therefore is worthwhile for society when

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

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 save the costs imposed on people who dislike their messages.

5.3 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 aggregate surplus of a competing type, such that no other type prefers to pool. If all recipients signal truthfully, then all value adding messages will be sent, and no value destroying message will be 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. Baseline 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. Recipients have internalized senders’ call externality surplus and, in general, senders gain nothing by communicating with a pool that has signaled a nonmatching type. 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 saved.

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 feasible equilibria, then confirm stability and existence. This avoids overly complex characterizations of all possible optima under all possible values. The feasible equilibria appear in Table 5.3.

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 1: Feasible Equilibria

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

Lemma 3 *At least one sender type will have zero surplus under the ABM, but remain in the market.*

The intuition here is that rational recipients will internalize as much of a sender’s surplus (his positive call externality) as possible up to the constraint of keeping him in the market. If all sender types had surplus remaining, then recipients could collect more.

It is also impossible for any recipient to capture surplus generated by transactions outside of her own pool. Given that a recipient pool internalizes all of a matching sender’s surplus, diverting wealth would cause sender surplus from those outside transactions to fall below zero, violating individual rationality for that sender.

Lemma 4 *Welfare under the ABM is at least that under the baseline, and all recipients that participate in the baseline 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 the baseline situation for themselves and their senders. Since they were in the market in the baseline voluntarily, they will never choose to leave when using the ABM. In a result analogous to that of Section 4, welfare under the ABM cannot be lower than under the baseline.

One subtlety, however, is that the ability to collect sender bonds can bring new agents to 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 can leave total welfare unchanged. If these entrants bring their own positive call externalities or negative processing costs, total welfare can rise or fall.

Lemma 5 *Under the ABM, the sender with the higher aggregate surplus is always preserved in the market, and always sends to all matching recipients.*

Self-interested recipients post bond requests as high as possible. It follows that the senders with more surplus are the ones targeted by recipients, who also cannot rationally drive them from the market. For the second half of the lemma, consider recipient choice. By forgoing matched messages, they lose both message value and the aggregate surplus of a matching sender in order to poach surplus of a mismatched sender. Forgoing matched messages cannot be rational for recipients who prefer messages from the sender with higher surplus.

Lemma 6 *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 will show that there exists 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.

5.4 ABM Versus Tax

Now we compare the ABM to the tax. When the tax is worse for welfare than the baseline, 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 value of the lower type transaction.

Proposition 3 *Social welfare of a flat tax never exceeds social welfare under the ABM unless the \mathbb{B} recipient is willing to forgo all type \mathbb{B} messages.*

The proof is in the Appendix.

Intuition for why welfare under the ABM generally exceeds that under the tax follows from the an understanding of the choices of each party. Senders necessarily prefer to reach only their matched recipient; reaching an unmatched recipient forfeits the bond 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 system can do better.

If any recipient signals untruthfully by choosing the bond of a different 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) parasitic bond surplus from declaring a different recipient type exceeds her native surplus and (2) having pooled with a different type, the true bond signal of a recipient parasite will not be present in the market. The declared type must then generate ample surplus implying that, for social efficiency, truthfully matched transactions should always complete.

Suppose that the matching sender for a parasitic recipient 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 reduce 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 potentially choose a true type "limit bond" to force a separating equilibrium. Analogous to a limit price which keeps competitors from entering one's market, a limit bond prevents potential parasites from preferring an off-type bond to their native surplus. As a credible signal, a limit bond also creates no social investment inefficiency, as it merely returns surplus to senders.

5.5 Additional Advantages

A key advantage of the ABM is that efficiency distortions only affect first contact. Individual recipients can trivially distinguish between senders they know and senders they do not know. Distributed among individuals in the marketplace, this private knowledge is potentially observable but cannot be used by a government or central authority without substantial loss of individual privacy. Thus, if α represents the proportion of first contact and $(1 - \alpha)$ the proportion of subsequent contact communications, then the relative efficiency of a mechanism that distorts only first contact communications is $\frac{1}{\alpha}$.

Further, as the number of market types increases, the number of pairwise

externality contributions from misdirected messages rise as $O(n^2)$. With linear marginal values, arising from fixed recipient probabilities and near zero costs, a tax can only alter the behaviors of senders whose marginal surplus is at or below the tax. Further still, society's interest in restricting transactions below this cutoff depends on the sort order of sender surplus paralleling the sort order of total transaction value. This property may or may not hold.

In effect, welfare under the ABM can exceed that under the tax for reasons beyond the internalization of call externalities. Its use reveals more information. As a screen, the bond forces senders to warrant their messages as being of the correct type. As a signal, separating bonds can reveal recipient type, thereby facilitating targeting. As a first contact mechanism, the ABM uses private knowledge of first versus subsequent contact to penalize only messages from strangers and not all messages globally.

6 Why not Always Seize the Bond?

Among heterogeneous recipients, if a bond is attached to a mistargeted communication, it will be seized. Further communications of this type provide no content benefit. For correctly targeted communications, however, can a recipient rationally forgo a bond from a welcome sender who has sufficient surplus to risk first contact $\sum_{i=0}^{\infty} E[s] - b \geq c_s$? If not, the ABM resembles a uniformly costly tax. It can outperform the tax on subsequent contact and in signaling recipient type, but not in setting a less costly screen for matching senders.

At issue is "first contact moral hazard." A recipient can seize the bond then whitelist matching senders. First contact bonds then become sunk costs. Further, a sender threat not to send is not credible if the sender surplus condition above holds. 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. Four conditions mitigate recipient interest in seizing bonds from welcome or matching senders.

The first exception yields a recipient signal of a relationship specific investment. As in the tax section, let sender values be weakly correlated with those of receivers, and let there exist a pooling equilibrium such that different recipients choose the same bond. Without loss of generality, let type \mathcal{B} recipients declare a bond b_G despite having a preference for type \mathcal{B} messages. Because values are correlated, a type \mathcal{G} sender prefers to interact with a type \mathcal{G} recipient (or a type \mathcal{U} recipient); likewise, a type \mathcal{G} recipient prefers to be so identified. Let correlated sender \mathcal{G} enjoy surplus increment $\Delta_s \equiv \sum_{i=1}^{\infty} \delta^i (s_G^G - s_B^G)$ from subsequent interaction with recipients of type \mathcal{G} over type \mathcal{B} . We sup-

pose further that the recipient merely needs to uniquely identify herself with an unambiguous reply in order for subsequent gains to be realized. If a reply is costless to forge, any recipient \mathcal{B} can continue to mimic a recipient \mathcal{G} , serving as a parasite in an effort to secure $\frac{\Delta_s}{2}$. Therefore, as a credible signal, recipient \mathcal{G} refunds a portion of b_G . Any refund amount is possible – including a subsidy larger than the bond – in which the subsequent recipient \mathcal{G} share of total incremental benefit $\Delta_r + \Delta_s$ exceeds the refund. Note that the minimum refund cannot be less than the subsequent value of the mismatched relationship $\sum_{i=1}^{\infty} \delta^i r_B^G$ to prevent a parasite recipient \mathcal{B} from forging a reply. The correlated increment to recipient surplus Δ_r of a correct match makes this possible. In a transactions cost sense Williamson (1985), b_G represents a hostage offered by the sender to identify a superior opportunity, while the refund offered by the recipient represents a relationship-specific investment to realize these opportunity gains.

The second exception represents an investment in reputation, again in the context of pooled recipients, but does not require repeated interaction with a given sender. Rather, it requires *ex post* verification of bond forbearance in the case of type \mathbb{G} senders, and recipient expectation of interaction with *other* type \mathbb{G} senders. Since message content resolves the information asymmetry regarding sender type, the classification of sender type and recipient behavior in response to sender type are both verifiable. To show that recipients can rationally forgo the bond, we borrow the framework of Rogerson (1983).

In a model of advertising, he shows that a firm can signal a high quality product by “investing” in high fixed costs and advertising. While irrational in the context of a fully informed consumer, the investment signals to uninformed consumers that the firm expects increased sales. This increase would not be possible if unfavorable word-of-mouth from a low quality product hurt future sales. Investment then becomes a credible signal of a valuable transaction.

In the context of the ABM, recipients can rationally forgo the bond as an investment in reputation conditional on the expectation of increased communication from other \mathbb{G} senders. Consider more refined sender heterogeneity. Suppose that among the population of good senders $E[r^G] > c_r$ are those who can and cannot afford bond seizure. Thus we simply let hi types have surplus $E[s^{GH}] \geq c_r + b$ and lo types have surplus $E[s^{GL}] \geq c_r$ but $E[s^{GL}] < c_r + b$ with relative hi and lo proportions of λ and $(1 - \lambda)$. Note that the proportion of \mathbb{B} is uninteresting as their bonds will always be seized. Then a recipient can forgo the bond on a hi type whenever $\lambda b < (1 - \lambda)E[r^{GL}]$. This equilibrium is sustainable so long as the first contact population distribution does not change sufficiently over time to invalidate the test. In effect, each act of recipient forbearance is a signal of investment in reputation analogous to a firm’s signal

of type through an investment in advertising. Rasmusen (2001) observes a similar effect: a contract writer's good reputation can lead a contract reader to accept risky terms without reading them. The analog, in our context, is to risk posting a bond given a recipient's reputation for refunding it.

A third reason to forgo the bond introduces scarcity and with it competition. Let sender attention be limited to N communications streams, say for personalized messages of type \mathbb{G} , in a context of $N + M$, $M \geq 1$ recipients. Then first contact moral hazard implies that, *ceteris paribus*, senders prefer recipients with lower posted bonds as $E[s^G] - b_1 > E[s^G] - b_2$ with $b_2 > b_1$. The presence of M uncompensated recipients implies Bertrand price competition to secure the benefit of valuable communication. This reduces the monopoly price established in Section 4. Lower bonds, even past the point of subsidy, are possible as recipients bid away their surplus. This property scales arbitrarily in the number of senders, stopping at the point where total sender attention exceeds total recipient demand. In short, while a recipient may have a monopoly on her own attention, she may be a price taker in the market for attention as a whole.

A final reason to forgo the bond arises in the context of sender and receiver role reversal. If the value from a communications stream is jointly created, being realized only after a recipient responds, the monopsonist in one direction becomes the monopolist in the other. Having solved the first contact matching problem, each can reciprocally post take-it-or-leave-it offers. For a finite game, as in Proposition 3, bargaining power rests with the last party in position to make this offer. In the infinite horizon game, however, Nash bargaining implies that each party would divide the expected surplus. But, since communications itself cannot be the subject of bargaining except through communication, a decision by both parties to whitelist the other and forgo the bond represents a decision to economize on transaction costs that could overwhelm joint surplus. That one party might need to send k times as often as the other does not matter if both parties are in fact indispensable to an infinite or randomly terminal communications stream.

7 Conclusions

This article presents mechanisms designed to improve the value of communication. The context for our analysis is multi-party, point-to-point media with call externalities and a two-sided information asymmetry. Among strangers generally, recipients do not know sender qualities and senders do not know recipient tastes. 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. Our main finding is that an attention bond designed to promote valuable communication can outperform mechanisms designed merely to block low value content or curb externalities.

Applying the Coase theorem, the mechanism operates by giving recipients a technologically conferred right in their attention. This not only affords recipients the ability to price their opportunity costs of message processing, it also allows them to substantially internalize the call externalities of senders. Internalization is achieved by posting a take-it-or-leave-it bond price on messages from unrecognized senders. Recipients then exercise their option to refuse non-conforming messages. Since unwanted communication is itself the problem, this sequence minimizes negotiation yet places substantial information in the market based on signaling and screening.

Non-pooling bond choices signal recipient types. These signals promote efficient targeting by senders based on information that filters and taxes generally do not recover. The bond also functions as a screen. Willingness to post a specific recipient bond distinguishes matched senders from unmatched senders based on a combination of recipient signals and private sender knowledge. This sequence also shifts the recipient burden from *ex ante* classification, a potentially unwinnable arms race, to one of *ex post* verification, a much easier task. Verification of the implied bond promise to send matching messages also discourages senders from misrepresenting their true purpose.

If recipient bond choices are pooling, a specific bond can serve as a transactions hostage, refunded by the recipient in order to distinguish her true type. Her confirmation signal implies that the future expected value of communication exceeds the present value of the bond. Otherwise, the bond serves as side payment, which consummates the price of the exchange. Among the extensions, we show that bonds taking negative values can affect purchases in the opposite direction.

Decentralized mechanisms offer at least three advantages relative to centralized mechanisms. First, as noted, individuals can set prices based on their private costs and private preferences, improving efficiency. Second, decentralization allows individual sender and receiver pairs to distinguish first contact from second contact. The mechanism lets users adapt to any level of prior knowledge, from no information to full information. Better information reduces distortion relative to mechanisms that cannot improve based on past experience. Third, matched communications markets exhibit subtle multiple externalities. Although a tax might correct for message pollution caused by a single sender type, finding a single efficient tax is probably infeasible for

multiple types.

Acknowledgments

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 Lones Smith.

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.
- 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.