Spam - solutions and their problems

B. Curtis Eaton
Department of Economics
University of Calgary

Ian A. MacDonald*
Commerce Division
Lincoln University

Laura Meriluoto
Department of Economics
University of Canterbury

June 13, 2007

Abstract

We model an email network in which messages are exchanged between consumers and unsolicited commercial email (spam) is sent by a spammer to consumers. We examine the spammer's profit-maximising problem and his responses to pricing and filtering deterrents. We find that sender pays and/or receiver pays pricing can be used to eliminate spam but at the cost of decreased welfare in the consumer-to-consumer network. Filtering alone is unlikely to offer a viable solution, and may exacerbate the problem if the spammer counteracts filtering by sending multiple variants of a message to each consumer. Filtering used in conjunction with sender pays pricing will reduce the size of the spam-eliminating sender pays price.

Keywords: spam, filtering, email, receiver pays pricing, sender pays pricing

JEL classification numbers: L96, L10

*Corresponding author. Address: Commerce Division, Lincoln University, PO Box 84, Lincoln 7647, New Zealand. phone: +64-3-325-3838 ext 8264; fax: +64-3-325-3847; email: macdonai@lincoln.ac.nz

1 Introduction

Unsolicited commercial email or "spam" imposes significant costs on the users of email and their internet service providers (ISPs). According to the CAN-SPAM Act of 2003, spam accounted for 50 percent of all US email traffic in 2003, up from just seven percent two years earlier. This huge growth in spam can be attributed to the fact that it works - spammers can be profitable with response rates of as low as 0.001%[1]. While the direct cost on spammers of sending millions of spam messages is small, the cost imposed by these messages on the system as a whole are significant. ISPs incur costs associated with wasteful consumption of bandwidth, increased demand on mail servers and a coresponding decrease in processor performance. This is turn has forced ISPs and other email network providers to invest in costly filtering technologies and increased infrastructure to overcome congestion of email servers. Consumers in the email network incur direct costs associated with the processing of spam and indirect costs resulting from decreased reliability of email systems and delay of messages[2]. Spam may also impose psychological costs on email users if the messages are offensive or when the sheer volume of messages becomes overwhelming. There is also a risk that spam will adversely affect the volume of beneficial non-spam messages, either because consumers simply stop using the email network altogether or when non-spam messages are caught by spam filters or accidentally deleted by the receiver.

There has been a recent move by many countries including the USA, Canada, New Zealand, India,and the countries of the European Union to enact anti-spam legislation that imposes hefty fines for the companies that send unwanted commercial email (Sorkin, n.d., www.spamlaws.com). The legislation is either

[1] Wall Street Journal, 2004

[2] Tens of thousands of New Zealanders recently experienced 24 hour delays in receiving emails when their ISP was bombarded by spam messages that were not caught by its filters (Chug, 2006).

"opt-in" requiring the senders to get consent from receivers prior to sending commercial email, "opt-out" requiring the sender to cease sending messages to a receiver upon their request, or a hybrid of the two. However, despite a number of spammers being convicted under these laws, legislation of this sort works only when the sender and the receiver fall under the same jurisdiction and so it is unlikely to provide widespread relief from spam.

The primary technological defense against spam is filtering. Filtering is a process whereby ISPs and other institutions attempt to identify and block spam messages prior to forwarding them into consumers' inboxes[3]. Messages can be blocked based on the sender's address, the message's subject and/or the content of the message. Unfortunately filtering is not perfect and the risk that non-spam messages will get caught by filters means that the filters must be set to err on the side of caution. Spammers can also evade filters by hiding their true identity, a practice known as spoofing, hiding the true subject or content of a message by adding characters to disguise certain keywords or sending messages as images rather than text, or by sending a large number of variant messages to each consumer in the hope that at least one of them will evade capture by the filters. Spammers have even started to use viruses that hijack consumers' computers so that they can use them to send spam messages (Griffiths, 2006). Moreover, while filtering might reduce the number of messages finding their way into consumers' inboxes it does little to reduce and may in fact increase the number, and associated costs, of spam messages received by ISPs.

The evident economic solution to the problem is to price spam out of the market (see for example Arrison(2004), Dai and Li (2004), Khong (2004) and Kraut et al (2005)). In theory, a sender pays price in the order of fractions of cents per message could eliminate spam by increasing spammers' per message

[3]Consumers can undertake a similar sort of activity by blacklisting messages from specific addresses.

costs above their expected per message revenue. However, while the agrement in favour of pricing is compelling in principle, to date the economic analysis of pricing solutions has been unsophisticated. In this paper we address this gap in the economic literature by modeling an email network in which email messages are exchanged between consumers and spam is sent by a spammer to consumers. Using this model we examine the incentives faced by the spammer to send spam messages, the impact of spam on consumer welfare, the effects of filters on the amount of spam sent and received, and the impact of uniform sender pays and receiver pays pricing to deter spam[4].

The paper is structured as follows. Section 2 looks at the spammer's problem as a two stage process. In the first stage, the spammer creates a mailing list by randomly drawing consumers from the population onto their mailing list. In the second stage, the spammer chooses the number of message variants to send to each consumer on their list so as to maximise the expected profit per target. Using this framework we are able to determine the impact of filtering and pricing on the number of consumers who are targeted by the spammer and the number of messages sent to each target. Because the spammer can increase the likelihood of eluding the ISP's filtering efforts by sending a number of variants of the message to each consumer, we find that filtering alone is unlikely to offer a viable solution to the spam problem. In fact, filtering leads to an increase in the total number of spam messages sent and sometimes even an increase in the number of spam messages arriving in consumers' inboxes when the spammer's per message cost is small compared to the profit of making a sale. This suggest that the interaction between spammers and ISPs in the use of filters has prisoner's dilemma characteristics. That is to say, collectively everyone (consumers, spammers and ISPs) could be better off if no one used

[4] In two separate papers we construct a model with many spammers and discuss the impact of a wider range of pricing on email networks in more depth.

filters but individually all ISPs have a clear incentive to use them. Filtering is a beneficial tool in the fight against spam when used in conjunction with sender pays pricing, however, because it reduces the magnitude of the required spam-eliminating sender pays price in direct proportion to the effectiveness of the filter which in turn reduces the detrimental impact of pricing emails in the consumer-to-consumer network. We also identify the receiver pays price that, by removing the incentive of consumers to read spam messages, removes the incentive for the spammer to send them.

In section 3 we model a simple consumer-to-consumer (c2c) email network in order to identify how the various email pricing regimes designed to combat spam affect the welfare of c2c messages. We find that the spam-eliminating sender pays price reduces the welfare of c2c messages if it is more than twice the difference between cost of each message to the ISP and the receiver and the expected receiver value of messages. Therefore, if the cost per message is less than the expected receiver value of c2c messages, the sender pays price will always lead to losses in the c2c network, and there is a clear trade-off between the direct benefits of eliminating spam and the losses of eliminating c2c messages. Furthermore, we find that the receiver pays price that eliminates spam leads to losses in the c2c network if it is more than twice the difference between the total cost per message and the expected sender value of all sent messages. Again, if the expected sender value of sent messages exceeds the total cost of each message, we have a clear trade-off between welfare gains due to eliminating spam and welfare losses due to eliminating c2c messages.

Section 4 concludes.

2 The spammer's problem

The objective of this section is to determine how filtering, receiver pays pricing and sender pays pricing affect the number of consumers the spammer targets, the number of messages the spammer sends to each target and the number of messages each target receives in their inbox.

The spammer is interested in selling his product to consumers and, in order to do so, must make contact with a consumer who is interested in purchasing the product. In order for such a contact to be made two things must occur. First, the spammer must place an email message in the consumer's inbox by both utilising their address and eluding any spam filters that are in place. Second, the interested consumer must read the message. Importantly, we assume that consumers cannot be identified by their tastes for spam and so the spammer cannot target his messages. Instead the spammer must contact consumers at random and this indiscriminate sending of messages means that for every message that finds its way into an interested buyer's inbox, many more are likely to be filtered or received by uninterested consumers.

Each spam message that is sent costs the spammer c^{spam} to process and send[5]. This per message cost for the spammer is certainly small and likely to be very close to zero. Each spam message costs the ISP c^U to transmit and each message that arrives in a consumer's inbox costs that consumer c^R to process. The spammer pays a per message sender pays price $p^S \geq 0$ to the ISP and the receiver of a message pays a per message receiver pays price of $p^R \geq 0$ to the ISP upon opening the message. We assume that a proportion α of all consumers receive benefit of $\rho^{spam} > 0$ from reading spam messages associated with gaining product information and reduced search costs. For simplicity we assume that

[5]In reality many of the spammers costs (such as access/bandwidth, labour, hardware, development of ways to avoid filtering, etc.) will be lumpy. For simplicity we model them as a constant per message marginal cost.

ρ^{spam} is the same for all α interested consumers. Spam messages will be read ($\theta = 1$) if $\rho^{spam} \geq p^R$ and will not be read ($\theta = 0$) otherwise. Again, although the spammer does not know the preferences of any particular consumer, he does have complete information about the benefit interested consumers receive from his message and about the magnitude of the receiver pays price. This means that the spammer can determine *ex ante* whether or not his messages will be read by those consumers who are interested. The spammer makes a profit of π associated with making a sale from each interested customer that reads his spam message.

For simplicity, because we are looking at constrained-optimal pricing and are not modeling ISP decisions per se, we treat the ISP as if it were a single entity that services all participants in the email network.

Filtering technologies employed by the ISP block messages that are from particular origins or that contain certain words or phrases in the subject line or body of the message. The spammer does not know the exact filtering technologies employed by the ISP but can try to evade them by avoiding words or phrases that are likely to be caught and/or by sending a number of variants of the message, perhaps from different origins. We capture the essence of filtering by assuming that any message sent by a spammer has a probability q of being filtered and $(1 - q)$ of getting through to a consumer's inbox. By sending multiple variants of a message to a consumer, the spammer increases the likelihood of getting at least one message in the receiver's inbox. With n messages sent to a consumer, the probability of at least one message getting through to her inbox is $(1 - q^n)$. Each message the ISP is successful in filtering saves a consumer c^R but still costs the spammer and the ISP c_s^{spam} and c^U respectively to process.

The spammer has a two-stage problem. In stage 1, the spammer generates a mailing list, by drawing with replacement from a population of N addresses. Denote the size of the mailing list by M. In stage 2, the spammer chooses how many messages to send to each consumer on his list. Denote the number of messages sent by the spammer to each of the consumers on his list by n. We use backward induction to solve the problem.

2.1 Stage 2 - Optimal number of messages per target

We introduce the stage 2 problem in continuous form even though it is not defined in the absence of filtering and does not perform well when the optimal number of messages is less than one. The reason we do this is because the continous model allows for the derivation of interesting closed-form comparative static results. We believe that the discrete form of the model better represents the reality of the problem, particularly when the number of messages sent to each consumer is likely to be small, and so we also provide a descretised representation of the spammer's per target message choice.

The number of messages per target sent by the spammer is found by maximizing expected profit per consumer w.r.t n:

$$\max_{\{n\}} \Pi = (1 - q^n)\theta\alpha\pi - (c^{spam} + p^S)n. \tag{1}$$

The profit-maximising number of messages per target is:

$$\begin{aligned} n^*(q) &= \frac{ln\left(-\frac{A}{ln(q)}\right)}{ln(q)} \text{ if } \theta = 1 \text{ and } A \leq \ln(q) \\ n^*(q) &= 0 \text{ otherwise} \end{aligned} \tag{2}$$

where $A = \frac{c^{spam}+p^S}{\alpha\pi}$ and the expected number of messages received by each targeted consumer is:

$$\begin{aligned} n^{inbox}(q) &= \frac{(1-q)ln\left(-\frac{A}{ln(q)}\right)}{ln(q)} \text{ if } \theta = 1 \text{ and } A \leq \ln(q) \\ n^{inbox}(q) &= 0 \text{ otherwise.} \end{aligned} \tag{3}$$

The maximized profit equals:

$$\Pi(n^*) = (1 - q^{\frac{ln\left(-\frac{A}{ln(q)}\right)}{ln(q)}})\alpha\pi - \frac{(c^{spam} + c_s^P)ln\left(-\frac{A}{ln(q)}\right)}{ln(q)} \text{ if } \theta = 1 \text{ and } A \leq \ln(q) \quad (4)$$
$$\Pi(n^*) = 0 \text{ otherwise}$$

and $\frac{\partial \Pi(n^*)}{\partial q} < 0$.

2.1.1 The impact of receiver pays pricing

Recall that for a receiver pays price $p^R \geq \rho^{spam}$, spam messages will not be read ($\theta = 0$) even by those consumers who are interested in purchasing the spammer's product and so no spam will be sent. Any price $p^R < \rho^{spam}$ will not affect the decisions of interested consumers and so will have no impact on the number of messages sent by the spammer. We define the spam-eliminating receiver pays price as:

$$p^R_{spam} = \rho^{spam}. \quad (5)$$

2.1.2 The impact of filtering

Assuming $\theta = 1$, the profit-maximising number of messages sent varies with q in the following way:

$$\frac{\partial n^*}{\partial q} = -\frac{1 + \ln\left(-\frac{A}{\ln(q)}\right)}{(\ln(q))^2 q} \quad (6)$$

and

$$\frac{\partial^2 n^*}{\partial q \partial A} = -\frac{1}{A(\ln(q))^2 q} < 0. \quad (7)$$

We can see from (6) that the number of spam messages increases as filtering is made more effective if $q < e^{-Ae}$ and decreases as filtering is made more effective if $q \geq e^{-Ae}$. Equation (7) shows that the smaller is A the faster is the change in n^* resulting from an improvement in the filtering technology.

The relationship between n^*, n^{inbox} and q is illustrated in Figure 1 for $A = 0.04$. Notice that n^* is increasing with q up to $q = 0.9$ and reaches a maximum of 9.2 messages per targeted consumer and n^{inbox} is increasing for $q \leq 0.46$ and remains above one for $q \leq 0.89$.

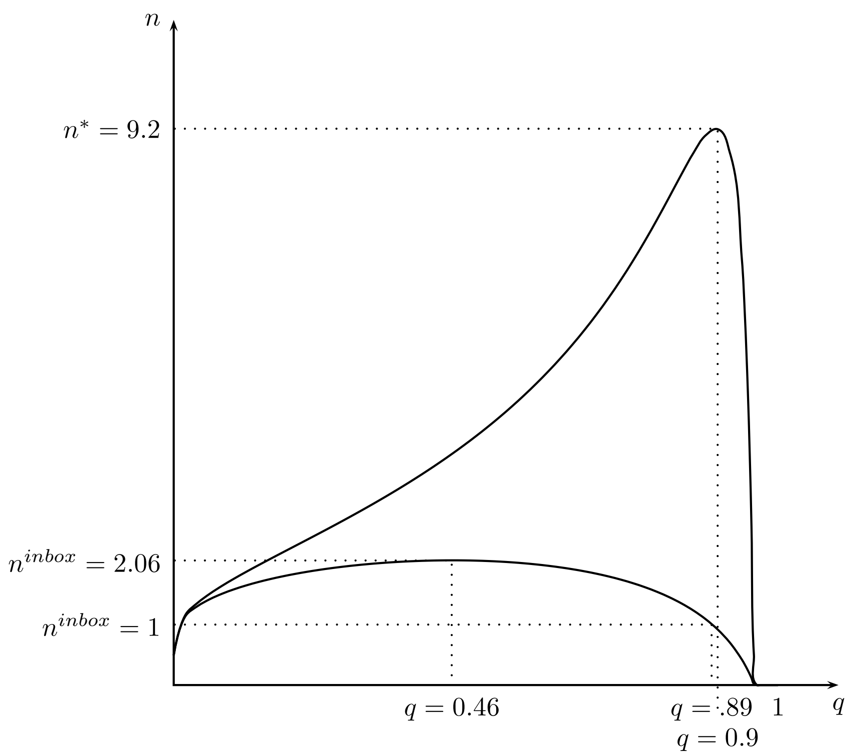

Figure 1: n^* and n^{inbox} for $A = 0.04$

2.1.3 The impact of sender pays pricing

Assuming $\theta = 1$, the profit-maximising number of messages sent varies with p^S in the following way:

$$\frac{\partial n^*}{\partial p^S} = \frac{\partial n^*}{\partial A} = -\frac{1}{(\ln(q))^2 A} < 0 \tag{8}$$

and

$$\frac{\partial n^*}{\partial \left(p^S\right)^2} = \frac{\partial^2 n^*}{\partial A^2} = \frac{1}{\left(\ln(q)\right)^2 A^2} > 0. \tag{9}$$

It is intuitively obvious that sender pays pricing deters spam and this is supported by (8). Equation (9) shows that the reduction in spam resulting from increases in p^S is most significant when p^S is small. Equation (7) shows that, the larger is q, the more effective is sender pays pricing in reducing the number of messages per target.

2.1.4 A discrete representation of spammer behaviour

If the spammer is restricted to choose the number of messages to send, n in $\{0, 1, ..., \infty\}$, the spammer sends no messages ($n^* = 0$) if $\Pi(1) < 0$, sends one message ($n^* = 1$) if $\Pi(1) \geq 0$ and if $\Pi(1) > \Pi(2)$, etc. Generally, $n^* = n$ if $\theta = 1$ and if

$$q^n(1-q) < A \leq q^{n-1}(1-q). \tag{10}$$

The expected number of messages that actually make it into the inbox of a consumer on the spammer's mailing list equals:

$$n^{inbox} = (1-q)n^*(q). \tag{11}$$

We illustrate the combinations of A and q for which the spammer chooses to send exactly n messages in an iso-message region mapping in Figure 2. The two inequalities in (10) define the iso-message region n bounded from below by iso-message boundary $n+1$ and from above by iso-message boundary n.

If $q = 0$ the spammer sends at most one message to each consumer on his list and all of these messages arrive in the consumers' inboxes. When $q > 0$ and $A > 0.25$, the spammer sends at most one spam message to each consumer on his mailing list and the expected number of messages received by each targeted consumer is $(1-q)$ if $q < 1-A$ and zero if $q \geq 1-A$. For $q > 0$ and $A < 0.25$, the descrete model and the continous model behave similarly and so the flavour

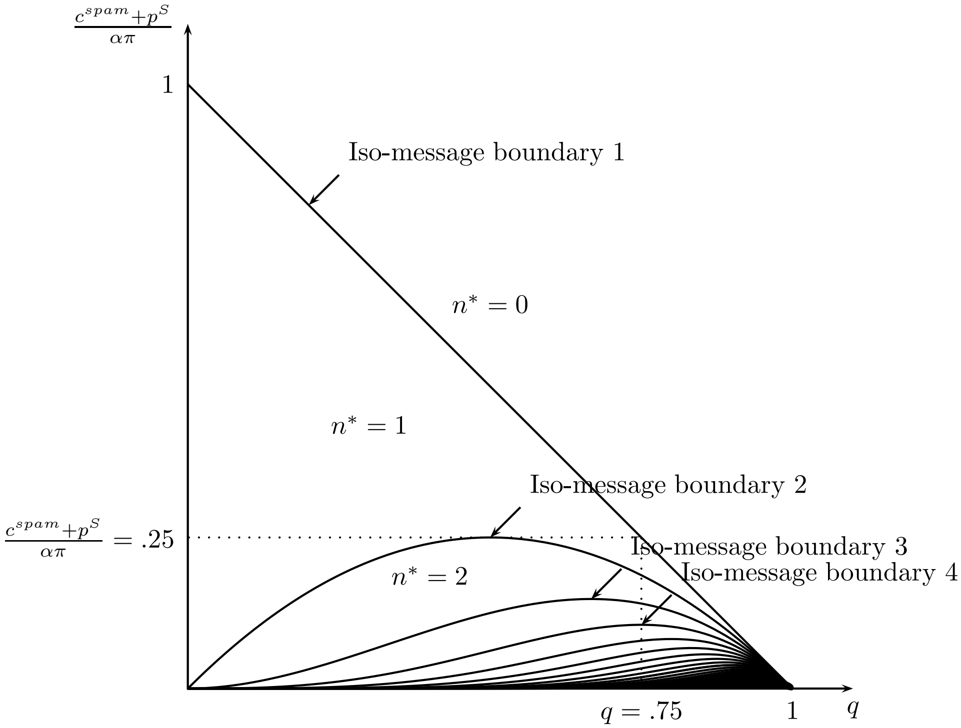

Figure 2: Iso-message regions and boundaries in a discrete representation of the model

of the comparative static results derived in equations (6)-(9) is evident in Figure 2 for this region of the parameter space. If $A < 0.25$, there is a range of values for q such that the spammer sends two or more messages to each consumer on his mailing list. For very small A the volume of spam increases rapidly as q increases from zero and starts to decrease only when q takes on values close to one. When $A < 0.25$, each targeted consumer can expect to receive more than one message in their inbox if $q < \frac{(n-1)}{n}$ and less than one message in their inbox if $q > \frac{(n-1)}{n}$. Clearly, if $q < \frac{(n-1)}{n}$ consumers are worse off in terms of the number of spam messages they receive in the presence of filtering than they would be in the absence of filtering because of the perverse incentives that filtering provides to the spammer. For example, from equations (10) and (11) we see that if $A = 0.1$ a filter that blocks 50% of all spam will result in 3 messages being sent by the spammer to each consumer on his list and each of these consumers can expect to receive 1.5 messages in their inbox.

2.2 Stage 1 - Optimal size of mailing list

In stage 1, the spammer chooses the size of its mailing list M at a total cost $C(M)$. The stage 1 objective function for the spammer is

$$V \equiv M\Pi(n^*) - C(M) \tag{12}$$

where $\Pi(n^*)$ is the expected profit per consumer from stage 2. The equilibrium condition is simply

$$\Pi(n^*) = C'(M). \tag{13}$$

We assume that the spammer builds his mailing lists by drawing addresses with replacement from the entire population N of email users and this assumption generates a specific cost function.

Let the cost of a draw be v. If the spammer already has a sample of size M, the probability of getting a unique name in the next draw is $\frac{N-M}{N}$ and, since the spammer is drawing with replacement, this probability remains constant until he has found an additional unique name. The expected number of draws, x, required to find one more unique name is therefore found from $x\frac{N-M}{N} = 1$ or $x = \frac{N}{N-M}$. The marginal cost of a generating a unique name is

$$C'(M) = \frac{vN}{N-M} \tag{14}$$

and $C(M)$ is the indefinite integral of (14):

$$C(M) = vN(\ln(N) - \ln(N-M)). \tag{15}$$

Notice that $C'(M) > 0$ and $C''(M) > 0$.

Using this formulation, the stage 1 equilibrium mailing list M^* is

$$M^* = \frac{(\Pi(n^*) - v)\,N}{\Pi(n^*)}. \tag{16}$$

Equation (16) shows that when the cost of drawing an additional address is zero, the spammer targets the entire population ($M^* = N$), but when the cost of a draw is positive the spammer targets only a portion of the population ($M^* < N$). The impact of changing the filter's effectiveness on the size of the spammer's mailing list is

$$\frac{\partial M^*}{\partial q} = \frac{vN\frac{\partial \Pi(n^*)}{\partial q}}{\Pi(n^*)^2} < 0. \tag{17}$$

The impact of changing the sender pays price on size of the spammer's mailing list is

$$\frac{\partial M^*}{\partial p^S} = \frac{\partial M^*}{\partial A} = \frac{vN\frac{\partial \Pi(n^*)}{\partial A}}{\Pi(n^*)^2} < 0. \tag{18}$$

2.2.1 Impact of filtering on total spam volume

Let the total number of messages sent by the spammer be

$$T = n^* M^*.$$

Assuming $\theta = 1$, filtering influences both the size of the spammer's mailing list and the number of messages sent to each consumer on the mailing list in the following way:

$$\frac{\partial T}{\partial q} = \frac{\partial n^*}{\partial q} M^* + \frac{\partial M^*}{\partial q} n^*. \tag{19}$$

The first term on the left hand side of equation (19) is the effect of changing the effectiveness of filtering on the number of messages sent to each consumer on the spammer's mailing list. From equation (6) we know that the number of spam messages targeted at each consumer on the spammer's list is an increasing function of q for $q < e^{-Ae}$ and a decreasing function of q for $q > e^{-Ae}$. The second term on the left hand side of equation (19) is the effect of changing the effectiveness of filtering on the number of consumers on the spammer's mailing list. From (17) we know that the size of the spammer's mailing list is unambiguosly a decreasing function of q. For $q > e^{-Ae}$, therefore, (19) is certainly negative but for $q < e^{-Ae}$ we cannot comment on whether the list effect outweighs the per target effect without pinning down parameter values. We know for certain, however, that as q increases the total volume of spam starts to decline before n^* starts to decline.

2.2.2 Impact of sender pays pricing on total spam volume

Assuming $\theta = 1$, increasing the sender pays price unambiguosly reduces the total volume of spam because both the size of the spammer's mailing list and the number of messages sent to each consumer on the mailing list, in equations (8) and (18) respectively, are decreasing functions of sender pays price:

$$\frac{\partial T}{\partial p^S} = \frac{\partial T}{\partial A} = \frac{\partial n^*}{\partial A} M^* + \frac{\partial M^*}{\partial A} n^* < 0. \tag{20}$$

3 The consumer network

The purpose of this section is to set out a framework in which we can analyze the impact in the c2c network of various pricing solutions to the spam problem. We analyse an email network in which consumers send and receive messages to and from other consumers and receive messages from spammers. The composition of the network is fixed and every consumer in the network has the ability to both receive and send messages. Consumers choose whether or not to send messages to other consumers, and if they receive a message from another consumer or the spammer, choose whether or not to open and read the message. For our purposes the identities of the receiver and/or the sender of a c2c message are of no concern – all that matters are the magnitudes of the benefits associated with those messages that are potentially sent and read.

The benefits associated with a message are captured by the pair (σ, ρ), where σ is the benefit the sender gets if the message is read, and ρ is the benefit the receiver gets from reading the message. If the message is not read, the sender's benefit is 0. In principle, both σ and ρ can be positive, negative, or zero. The ISP incurs a marginal cost c^U when any c2c message is sent regardless of whether or not the receiver actually reads the message. For any outgoing message, the sender incurs a processing cost c^S and for any incoming message, regardless of whether or not the message is actually read, the receiver incurs a processing cost c^R. We assume that consumers can identify the benefit and cost parameters associated with every message for which they are either the receiver or the sender. The receiver of a message can identify this information without opening the message.

Given a uniform sender pays price p^S, the private surplus the sender gets from an outgoing message that is read is

$$ps^S(\sigma, p^S) = \sigma - c^S - p^S. \tag{21}$$

Given a uniform receiver pays price p^R that is payable only if the receiver opens the message, the private surplus that the receiver gets from receiving, opening and reading an incoming message is

$$ps^R(\rho, p^R) = \rho - c^R - p^R. \tag{22}$$

Because c^R is incurred when the message enters the receiver's mailbox but before she makes a decision to read the message, it has no impact on the decision to read the message. Therefore, those messages for which both $ps^S \geq 0$ $(\sigma \geq c^S + p^S)$ and $ps^R \geq -c^R$ $(\rho \geq p^R)$ will be sent and read. The social surplus of a message (σ, ρ) that is sent and read is

$$ss(\sigma, \rho) = \sigma + \rho - (c^S + c^R + c^U). \tag{23}$$

Although it is possible to observe whether or not a message has been opened, it is not possible to observe whether or not the receiver actually reads the message. Hence a negative receiver pays price cannot induce the receiver to read a message for which $\rho < 0$ and so achieves nothing useful that cannot be achieved by a 0 receiver pays price. For this reason we consider only non-negative receiver pays prices: $p^R \geq 0$. It is conceivable that the optimal sender pays price might be one such that $p^S + c^S < 0$ or $p^S < -c^S$, because such prices induce senders to send efficient messages for which the benefit to the sender is negative but the benefit to the receiver is positive and large. But such prices create an incentive to manufacture and send phoney messages as, in effect, a commercial activity. For this reason we restrict attention to sender pays prices that satisfy $p^S > -c^S$. Therefore, given (p^S, p^R), messages in the following set will be sent and read

$$SR(p^S, p^R) \equiv \{(\sigma, \rho) | \sigma \geq \max(c^S + p^S, 0), \rho \geq \max(p^R, 0)\}. \tag{24}$$

In the analysis that follows, we assume a distribution of c2c messages that is consistent with (24). Specifically, let the preferences for sending messages $\sigma \in [0, \sigma^{\max}]$ be given by a distribution function $F(\sigma)$ with $F(0) = 0$, $F(\sigma^{\max}) = 1$ and $f(\sigma) = \frac{\partial F(\sigma)}{\partial \sigma}$. Also, let the preferences for receiving messages $\rho \in [0, \rho^{\max}]$ be given by a distribution function $G(\rho)$ with $G(0) = 0$, $G(\rho^{\max}) = 1$ and $g(\rho) = \frac{\partial G(\rho)}{\partial \rho}$. Assume that both $F(\sigma)$ and $G(\rho)$ are independent, continuous, increasing and twice differentiable. In Figure 3 we illustrate a possible distribution of c2c messages, without consideration of the density of messages, and identify the messages that will be sent and read. For zero prices the sent and read region is defined by $\sigma \geq c^S$ and $\rho \geq 0$ and is illustrated with the entire lined and cross-hatch area. Note that those messages that are above the line $\rho = C - \sigma$ improve welfare and those that are below the line reduce welfare. With positive prices p^S and p^R the sent and read region shrinks to the cross-hatch area resulting in both welfare benefits and losses. Naturally the net change in welfare depends on the distribution of messages and the magnitude of the processing costs in a way that is described in the subsections that follow.

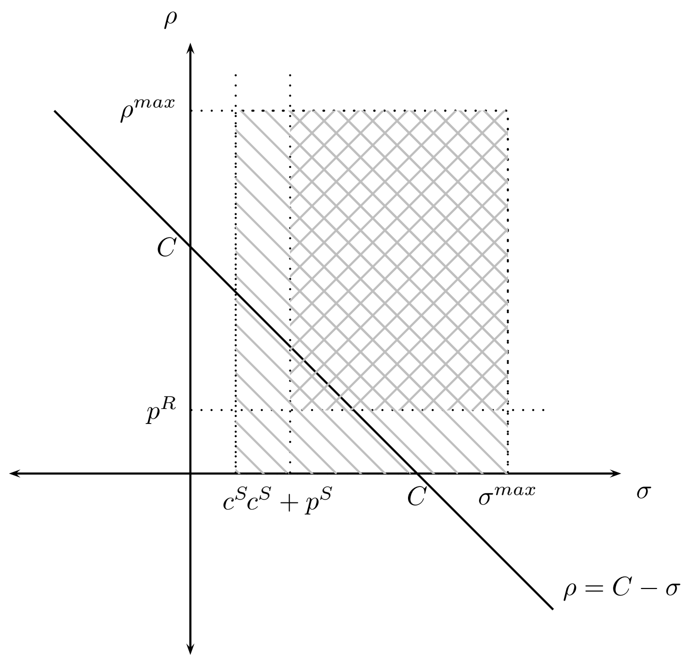

Figure 3: c2c message space

3.1 Total surplus of c2c messages

The social welfare of a c2c messages that are sent and read, given in (23), is $\rho + \sigma - C$ where $C = c^S + c^R + c^U$. For a distribution of preferences and arbitrary uniform sender pays and receiver pays prices $p^R \geq 0$ and $p^S + c^S \geq 0$, we can construct the social welfare function for the c2c network as a whole[6]. To begin, we know that all the messages for which $\rho \in [p^R, \rho^{\max}]$ and $\sigma \in [p^S + c^S, \sigma^{\max}]$ will be sent and read. The total number of these messages is given by the double integral $\int_{p^S+c^S}^{\sigma^{\max}} \int_{p^R}^{\rho^{\max}} g(\rho) f(\sigma) d\rho d\sigma$ and the total surplus of these

[6]The problem is symmetric around p^R and $p^S + c^S$, not p^R and p^S, and we present the analysis for the first set of 'prices'.

messages is the definite double integral of the surplus per message over $\rho \in [p^R, \rho^{\max}]$ and $\sigma \in [p^S + c^S, \sigma^{\max}]$.

$$\begin{aligned} TSS(\rho, \sigma, p^S, p^R) &= \int_{\sigma=p^S+c^S}^{\sigma^{\max}} \int_{\rho=p^R}^{\rho^{\max}} (\rho + \sigma - C)\, g(\rho) f(\sigma) d\rho d\sigma \qquad (25) \\ &= \left[\rho^{\max} - p^R G(p^R) - \int_{p^R}^{\rho^{\max}} G(\rho) d\rho\right] \left(1 - F\left(p^S + c^S\right)\right) \\ &\quad + \left[\sigma^{\max} - \left(p^S + c^S\right) F\left(p^S + c^S\right) - \int_{p^S+c^S}^{\sigma^{\max}} F(\sigma) d\sigma\right] \left(1 - G\left(p^R\right)\right) \\ &\quad - C\left(1 - F\left(p^S + c^S\right)\right)\left(1 - G\left(p^R\right)\right) \end{aligned}$$

Note that (25) is defined for $p^R \leq \rho^{\max}$ and for $p^S + c^S \leq \sigma^{\max}$. The total surplus can be interpreted as follows: the first term of (25) is the expected surplus of sent and read messages to the receiver; the second term is the expected surplus of messages to the sender; and the last term is the total cost of all messages sent and received. The expected surplus of sent and read messages to the receiver is found by multiplying the receivers' expected value of all the messages that would be read by the proportion of those messages that are actually sent. Likewise, the expected surplus of messages to the sender is found by multiplying the senders' expected value of all messages that would willingly be sent if they were read by the proportion of those messages that are actually read. When neither receivers nor the senders pay for the messages, which is the norm currently, the surplus is equal to

$$\begin{aligned} TSS|_{p^S=0.p^R=0} &= \left[\rho^{\max} - \int_0^{\rho^{\max}} G(\rho) d\rho\right] \left(1 - F\left(c^S\right)\right) \qquad (26) \\ &\quad + \left[\sigma^{\max} - \left(c^S\right) F\left(c^S\right) - \int_{c^S}^{\sigma^{\max}} F(\sigma) d\sigma\right] \\ &\quad - C\left(1 - F\left(c^S\right)\right) \end{aligned}$$

3.2 Welfare effect of introducing a receiver pays price

We now want to look at how the spam-eliminating receiver pays price affects welfare in the c2c market. The surplus for c2c messages, given an arbitrary receiver pays price p^R and a zero sender pays price is

$$\begin{aligned} TSS|_{p^S=0} &= \left[\rho^{\max} - p^R G(p^R) - \int_{p^R}^{\rho^{\max}} G(\rho) d\rho\right]\left(1 - F\left(c^S\right)\right) \\ &+ \left[\sigma^{\max} - c^S F\left(c^S\right) - \int_{c^S}^{\sigma^{\max}} F(\sigma) d\sigma\right]\left(1 - G\left(p^R\right)\right) \\ &- C\left(1 - F\left(c^S\right)\right)\left(1 - G\left(p^R\right)\right) \end{aligned} \tag{27}$$

The change in surplus going from a zero to a positive receiver pays price is therefore

$$\begin{aligned} & TSS|_{p^S=0} - TSS|_{p^S=0.p^R=0} \\ =\ & -\left[p^R G(p^R) - \int_0^{p^R} G(\rho) d\rho\right]\left(1 - F\left(c^S\right)\right) \\ & -\left[\sigma^{\max} - c^S F\left(c^S\right) - \int_{c^S}^{\sigma^{\max}} F(\sigma) d\sigma\right] G\left(p^R\right) \\ & + CG\left(p^R\right)\left(1 - F\left(c^S\right)\right). \end{aligned} \tag{28}$$

The overall change in welfare associated with the introduction of a receiver pays price is ambiguous. The first term in (28) is negative, represents the decrease in expected receiver value of messages as a result of not reading messages for which $\rho < p^R$, and is large if $G(\rho)$ is heavily weighted between 0 and p^R and/or p^R is large. The second term is also negative, represents the decrease in expected sender value associated with the reduction in messages that are sent because they are no longer read, and is likely to be large if $\sigma^{\max}$ is large and/or $F(\sigma)$ is heavily weighed towards large realisations of σ and/or p^R is large. The third term is positive, represents the cost saving due to the reduction in the number of messages sent and read, and is likely to be large if $G(\rho)$ is heavily weighted towards small values of ρ and/or p^R is large.

For many distributions, $\left[p^R G(p^R) - \int_0^{p^R} G(\rho) d\rho\right] \approx G(p^R)\frac{p^R}{2}$ because, for small receiver-pays prices, the cumulative distribution of receiver value is approximately linear[7], the total effect on welfare can be approximated by

$$\begin{aligned} & TSS|_{p^S=0} - TSS|_{p^S=0.p^R=0} && (29) \\ \approx & -\left[\left(\sigma^{\max} - c^S F(c^S) - \int_{c^S}^{\sigma^{\max}} F(\sigma)d\sigma\right) + \left(\frac{p^R}{2} - C\right)\left(1 - F(c^S)\right)\right] G(p^R). \end{aligned}$$

The total surplus of c2c messages decreases with the introduction of a receiver pays price if

$$p^R > 2\left[C - \frac{\left(\sigma^{\max} - c^S F(c^S) - \int_{c^S}^{\sigma^{\max}} F(\sigma)d\sigma\right)}{(1 - F(c^S))}\right], \quad (30)$$

which holds if the receiver pays price required to eliminate spam is more than twice the difference between the total cost and the expected sender value of all sent messages[8]. Notice that the total surplus of c2c messages goes down with the introduction of spam-eliminating receiver pays price always if the expected sender value of sent messages exceeds the total cost of each message. Moreover, when the spam-eliminating receiver pays price reduces the welfare of c2c messages, the larger is the receiver pays price and/ or more heavily weighted is $F(\sigma)$ between 0 and p^R the larger is the decrease in c2c welfare c*eteris paribus*.

3.3 Welfare effect of introducing a sender pays price

We now want to look at how the spam-eliminating sender pays price affects welfare in the c2c market. The surplus for c2c messages, given an arbitrary price p^S and a zero receiver pays price is

$$\begin{aligned} TSS|_{p^R=0} = & \left[\rho^{\max} - \int_0^{\rho^{\max}} G(\rho)d\rho\right]\left(1 - F\left(p^S + c^S\right)\right) \\ & + \left[\sigma^{\max} - \left(p^S + c^S\right) F\left(p^S + c^S\right) - \int_{p^S+c^S}^{\sigma^{\max}} F(\sigma)d\sigma\right] \\ & - C\left(1 - F\left(p^S + c^S\right)\right) \end{aligned}$$

[7] For the uniform distribution it is exactly linear.

[8] Notice that this form of expected sender value excludes the messages that give zero value due to not being sent.

and the change in surplus going from a zero to a positive sender pays price is

$$
\begin{aligned}
& TSS|_{p^R=0} - TSS|_{p^S=0.p^R=0} \qquad (31)\\
= \quad & -\left[\rho^{\max} - \int_0^{\rho^{\max}} G(\rho)d\rho\right]\left(F\left(p^S+c^S\right) - F(c^S)\right) \\
& -\left[(p^S+c^S)F\left(p^S+c^S\right) - c^S F(c^S) - \int_{c^S}^{p^S+c^S} F(\sigma)d\sigma\right] \\
& +C\left(F\left(p^S+c^S\right) - F(c^S)\right)
\end{aligned}
$$

The overall change in welfare associated with the introduction of a sender pays price is ambiguous. The first term in (31) is negative, represents the decrease in expected receiver value of messages as a result of having fewer messages sent, and is likely to be large if $\rho^{\max}$ is large and/or $G(\rho)$ is heavily weighed towards large realisations of ρ and/or $F(\sigma)$ is heavily weighted towards realisations between p^S and c^S. The second term is also negative, represents the decrease in expected sender value, and is likely to be large if p^S is large and/or $F(\sigma)$ is heavily weighted towards realisations between p^S and c^S. The third term is positive, represents the cost saving due to the reduction in the number of messages sent, and is likely to be large if p^S is large and/or $F(\sigma)$ is heavily weighted towards realisations between p^S and c^S.

Again, because $\left[(p^S+c^S)F\left(p^S+c^S\right) - c^S F(c^S) - \int_{c^S}^{p^S+c^S} F(\sigma)d\sigma\right] \approx \left(F(c^S+p^S) - F(c^S)\right)\left(c^S + \frac{p^S}{2}\right)$, the total effect on welfare can be approximated by

$$
\begin{aligned}
& TSS|_{p^R=0} - TSS|_{p^S=0.p^R=0} \qquad (32)\\
\approx \quad & -\left(F\left(p^S+c^S\right) - F(c^S)\right)\left[\left(\rho^{\max} - \int_0^{\rho^{\max}} G(\rho)d\rho\right) + c^S + \frac{p^S}{2} - C\right].
\end{aligned}
$$

The total surplus of c2c messages goes down with the introduction of spam-eliminating price if

$$
p^S > 2\left[\left(C - c^S\right) - \left(\rho^{\max} - \int_0^{\rho^{\max}} G(\rho)d\rho\right)\right] \qquad (33)
$$

or if the sender pays price required to eliminate spam (or to reduce spam) is more than twice the difference between the cost of each message to the ISP and the receiver and the expected receiver value of messages. This is always the case when the expected receiver value of messages exceeds the cost of them to the ISP and the receiver. When the change in c2c welfare is negative, the larger is the spam-eliminating sender pays price and/or the more density there is between c^S and $p^S + c^S$ the more the welfare or c2c messages deteriorates with the introduction of the sender-pays price.

4 Conclusion

We have examined receiver pays pricing, sender pays pricing and filtering solutions to the spam problem and have highlighted the trade-offs that are inherent in the use of each. Receiver pays and/or sender pays pricing can be used to eliminate spam but at the potential cost of decreased welfare in the c2c network. Receiver pays pricing works by removing the incentive of the receivers of spam messages to open them and sender pays pricing works by rendering spam unprofitable. In addition to eliminating spam, both pricing solutions reduce the number of messages exchanged between consumers and so, depending on parameter values, can lead to a significant loss of welfare in the c2c network. Filtering alone is unable to offer a viable solution to the spam problem if spammers counteract filtering by sending multiple variants of a message to each consumer. In fact, our model suggests that filtering can be counterproductive by leading to an increase in the total volume of spam and in sometimes even the number of spam messages arriving in consumers' inboxes. When used in conjunction sender-pays pricing, however, filtering will reduce the size of the spam-eliminating sender-pays price and thus reduce the degree to which the c2c network is affected.

While we believe our analysis provides a good first step towards understanding the economic implications of solutions to the spam problem, there is still much work to be done. We have made no attempt in this paper to model the interaction between ISPs and so we are unable to comment on the likelihood that ISPs could agree to implement pricing solutions rather than, for example, using differential pricing to attract business. No doubt there are also technical issues associated with ensuring that senders or receivers of emails are unable to evade payment of prices. Finally, because we do not provide an empirical study of the spam problem we are unable to comment on whether the total net effect of any attempt to control spam will be positive or negative or whether, if positive, the net benefit warrants the investment that would be required to make these pricing solutions work.

References

[1] Arrison, Sonia (2004) "Canning Spam: An economic solution to unwanted email". Pacific Research Institute.

[2] Dai, R and Li, K. (2004) "Shall we stop all unsolicited email messages?" Proceedings of First Email and Anti-spam Conference.

[3] Griffiths, Peter (2006) "Email gangs bombard web in 'spam wars'" *The Press*, Christchurch, 29 November 2006.

[4] Khong, Dennis (2004) "An economic analysis of spam laws", *Erasmus Law and Economics Review 1*, February, 23-45.

[5] Kraut, Robert et. al. (2005) "Pricing electronic mail to solve the problem of spam", *Human-computer Interaction 20*, 195-223.

[6] Sorkin, D. E., n.d., retrieved from the www site www.spamlaws.com on 15 March 2007.

[7] The Can-Spam Act of 2003 (US Public Law 108-187), retrieved from the www site http://www.spamlaws.com/federal/can-spam.shtml on 15 March 2007.

[8] The Wall Street Journal, 2004