Email pricing

B. Curtis Eaton* Ian MacDonald† Laura Meriluoto‡

June 25, 2008

Abstract

We study how the nature of email technology affects the level of efficient and profit-maximizing prices as well as the optimal mix between sender and receiver pricing. We find that consumers cannot be induced to read efficient messages with negative receiver prices and that there is a cap in the size of the subsidy that can induce senders to send efficient messages. This implies that the first-best outcome may not be achievable even with perfectly discriminatory prices. We study efficient uniform sender and receiver prices and find that the sender price decreases to the negative of the sender's processing cost and the receiver price decreases to zero as the maxima in the preference distributions increase. Senders (receivers) tend to pay more than receivers (senders) when the message preference distribution has a lot of mass for messages with higher sender (receiver) value and low receiver (sender) value. Perhaps more surprisingly, we find that even with perfectly symmetric message preference distributions, the optimal uniform prices are asymmetric in that the receiver pays more than the sender. The sum of the efficient uniform receiver and sender prices never cover all the message costs nor do they cover ISP costs when both the maxima in the message preference distributions are sufficiently large. The efficient prices given ISP break-even constraint and the profit-maximizing prices are also asymmetric in that receivers pay more than the senders when the message preference distributions are symmetric.

Keywords: email, receiver pricing, sender pricing

JEL classification numbers: L96, D60

*University of Calgary; eaton@ucalgary.ca.
†Lincoln University; macdonai@lincoln.ac.nz.
‡University of Canterbury; laura.meriluoto@canterbury.ac.nz.

1 Introduction

To many people concept of charging people to send or receive email may be objectionable but the inefficiencies in email networks associated with overuse resulting from the choice to charge zero prices are significant and increasing. When faced with similar inefficiencies in other settings economists usually turn to the power of prices to influence behavior as a primary solution and, we argue, so it should be in email networks as well.

While there is an established literature on pricing in other networks, few papers have looked at the idiosyncratic nature of the technology in email networks. If ISPs are serious about introducing pricing there needs to be a good understanding of what exactly prices can do in email networks, what the efficient and profit-maximizing prices are and how best sender and receiver pricing can be used together to achieve economic goals. We address these issues in this paper.

As a first step towards a better understanding of how prices work in email networks, we discuss the limits the specific technology of email networks places on the set of prices that can influence consumer behavior. One of the most important differences between the technology of email and many other communication networks is that the receiver bears the processing cost of a message whether or not she reads the message. Because the processing cost is sunk it does not influence the receiver's decision to read messages.

We find that consumers cannot be induced to read messages with negative receiver prices and that if senders are given a subsidy to induce them to send messages this subsidy cannot be larger than their processing cost of the message. This implies that the first-best outcome cannot be achieved in email networks even with perfectly discriminatory prices if the minima in the message preference distributions are negative.

We calculate the optimal uniform sender and receiver prices and discuss how and why they differ from those suggested for other networks and from those suggested by others for email networks (for example Hermalin and Katz (2004)). We show that the two prices are asymmetric even when the preferences for sending and receiving messages are distributed identically and independently by a uniform or a beta distribution. Given such preference distributions the receiver pays more than the sender and both efficient prices decrease with the maximum in the preference distribution to the point where, if the maximum is sufficiently large, the efficient receiver price is equal to zero and the sender subsidy is equal to the sender's processing cost. It is also true that if the maximum is sufficiently large the sum of the efficient receiver and sender prices does not cover the ISP's costs. We examine the efficient Ramsey prices given an ISP break-even constraint and show that these too are asymmetric. We discuss the circumstances under which it is efficient to use sender pricing alone, receiver pricing alone or sender and receiver prices together.

We do not attempt to capture the effects of spam in this paper. In fact, we argue that a model of communication demand with a smooth and continuous distribution of preferences, as most papers in this area assume, is ill suited for addressing spam for a number of reasons. Because spam is not targeted to those receivers who value it, most spam messages generate zero receiver utility. This suggests that one possible way to model spam is to represent it as a mass point within the non-spam preference distribution. A further complication arises, however, because with filtering many spam messages that are sent are not received. One really needs to specifically model the behavior of spammers in the face of email pricing as we do in Eaton, MacDonald and Meriluoto (2008).

To help put our paper in context with the existing literature, we now want to briefly discuss the main approaches of modeling demand taken in the communication network literature. Many papers that study pricing in networks assume that consumers are heterogenous (in income or preferences or both) and that their demand for communication depends on the number of consumers in the network, to allow for network effects, but not on their identities. Examples are Squire (1973), Littlechild (1975), Dhebar and Oren (1985), Einhorn (1993) and Hahn (2003). In addition to the assumptions that imply that the utility derived from communication is not dependent on the others' identity, Squire assumes a constant, positive marginal utility for receiving calls. The set-up allows him to examine efficient uniform call and access prices when call and access externalities are present. Littlechild (1975) derives uniform surplus-maximizing and profit-maximizing two-part tariffs when consumers get no utility from receiving a call, i.e. when call externalities are not present, but when access externalities do exist. Dhebar and Oren (1985) study dynamic monopoly pricing in an expanding network. Einhorn (1993) assumes a constant receiver value of a message and analyzes how network effects affect Ramsey prices. Hahn (2003) examines nonlinear pricing (excluding receiver prices) in the presence of network and call externalities.

The more recent studies on network competition make the additional assumption that there are no network effects for usage and access. That is, receivers are assumed not to gain utility from receiving messages and the utility of participating in a network is not dependent on the number of other subscribers. These two further simplifications enable the authors to express the network providers' demands as simple functions of their own and their competitors' prices and therefore to model competition between interconnected networks that exchange regulated or negotiated termination charges. Examples of such papers are Armstrong (1998), Laffont, Ray and Tirole (1998a, 1998b), Kim and Lim (2001) and Carter and Wright (2003). Jeon, Laffont and Tirole (2004) extend analysis of Laffont, Ray and Tirole by adding call externalities and receiver prices to the model and are therefore able to investigate networks' pricing strategies as well as regulation of termination charges and receiver prices.

For some research questions, however, it is necessary to assume a richer preference structure, one where the identity of the sender and a receiver of a message matter to the utility derived from the message. For example, if there are negative external effects a representative consumer model that necessarily assumes positive utility for messages is not suitable. Clearly many communication networks are affected by negative effects including, of course, email networks with the presence of spammers. Rohlfs (1974) studies access decisions with a model where the identity of consumers matters. He derives the conditions for equilibrium user sets for a given a uniform sender price and discusses the possibility of multiple equilibria and start-up problems in network creation. His set-up does not enable him to derive either perfectly discriminatory or uniform welfare- or profit-maximizing prices.

MacDonald and Meriluoto (2005) also have a preference structure where the identity of the parties to a message exchange matters. They derive efficient caller and receiver prices and show that all access effects are internalized with such prices so that access price can efficiently be set equal to marginal cost. This structure, however, does not allow for the analysis of uniform welfare- or profit-maximizing monopoly prices or competition amongst network providers. So far there have been no studies that have derived total welfare meaningfully, that is without artificially pinning down taste parameters, in a model where the identity of network participants matter for the utility of others. The result would have to be a function where the nature of demand could be easily altered, such as with ordinary demand systems where the slope, intercept or elasticity of the demand function can be changed.

This paper follows the approach of Hermalin and Katz (2004) in that the unit of interest is a message, not a consumer. Message preferences are distributed in a known way and no effort is made to link messages to the consumers who are sending or receiving them. This framework allows for the determination of uniform prices that are first-best efficient or second-best efficient (subject to informational or other constraints) and for the study of monopoly prices perfectly discriminatory usage prices. However, it is not possible to use this framework for analyzing utility of access and access pricing because for such studies the unit of interest must be the consumer. Consequently, it is hard to see how such framework can be used to address issues around competition where the utility each consumer derives from belonging to the competing networks must be known. Hermalin and Katz (2004) examine uniform Ramsey pricing in a set-up where each message generates utility for both the sender and the receiver and the network provider knows the distribution of these message preferences.

The rest of the paper is structured as follows. Section 2 presents the model assumptions on preferences and costs. Section 3 describes the social and private surpluses of messages given some arbitrary sender and receiver prices, and describes what are the constraints set by the email technology on prices. Section 4 describes the general conditions for social welfare maximizing prices, Ramsey prices and profit-maximizing prices. Section 5 presents social welfare maximizing prices, Ramsey prices and profit-maximizing prices for a uniform distribution. Section 6 concludes.

2 Model set-up

2.1 Preferences

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, they choose whether or not to open and read the message.

The benefits associated with a message from sender s to receiver r are captured by the pair (σ_{sr}, ρ_{sr}), where σ_{sr} is the benefit the sender gets if the message is read, and ρ_{sr} 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_{sr} and ρ_{sr} can be positive, negative, or zero.

We assume that the email message heading has enough of information to enable the receiver to correctly estimate its value to her.

For our purposes the identities of the receiver and/or the sender of a message are of no concern – all that matters are the magnitudes of the benefits associated with the message. Hence, we can describe the potential benefits of the email network by a density function $M(\sigma, \rho)$. The potential benefits are of course realized only if the message is sent and read. For convenience we assume that there is a continuum of messages, and that the density function is continuous and independent in its two arguments. So, we can express the density of message preferences as $M(\sigma, \rho) = f(\sigma)g(\rho)$.

2.2 Costs

There is a constant per message cost, $C = c^U + c^S + c^R$, that can be broken down to three components. When an email message is sent, the ISP incurs a marginal cost, c^U. This cost is incurred regardless of whether or not the receiver actually reads the message. This assumption differs from that made by the existing literature of telephone pricing and email pricing. As the cost of a call is realized only if the call is answered and therefore the physical connection is made, it is sensible to assume that the network provider's cost requires both the caller and receiver to act. In email networks, however, a message is transmitted without the receiver's consent, and chews up bandwidth whether or not it is read. However, the literature on email pricing has not previously adopted our assumption. Hermalin and Katz (2004) assume a per message cost m which is incurred only if the sent message is accepted. Loder et. al. (2006) do not include an ISP cost. In perfect information equilibrium, however, both approaches are equivalent because in our model messages are sent only if the sender anticipates them to be read. However, if we introduce asymmetry of information, such as would be reasonable at least if some network participants were spammers who do not know which consumers will respond to their message, the assumption of the ISP cost being incurred regardless or whether or not the receiver reads the message becomes important. In fact, the ISP cost of unwanted spam is one of the major costs of spam.

For every outgoing message, the sender incurs a processing cost c^S. This assumption is equivalent to that of Loder et. al. and effectively also to that of Hermalin and Katz who assume that the preference parameters are net of any cost associated with sending or opening and reading a message.

For every incoming message, the receiver incurs a processing cost c^R, regardless of whether or not the message is actually read. In contrast, Loder et. al. (and effectively Hermalin and Katz due to the cost being lumped up with the benefit of reading a message) assume that this cost is incurred only if the receiver reads the message. Thus, their assumption is very much in line with the current telephone technology but not with the email technology. This assumption will affect the main results of the model.

3 Message surplus

3.1 Social surplus of a message

The social surplus of a message (σ, ρ) that is sent and read is

$$ss(\sigma, \rho) = \sigma + \rho - (c^S + c^R + c^U). \tag{1}$$

Cost-benefit optimality requires that this message be sent and read if $ss(\sigma, \rho) \geq 0$, i.e. if $\rho \ \geq C - \sigma$ where $C = c^S + c^R + c^U$, and that it not be sent if $ss(\sigma, \rho) < 0$.[1] In Figure 1, those messages that are on and above the line $\rho \ \geq C - \sigma$ are efficient and those that are below the line are inefficient. Notice that it is possible that some efficient messages have a negative value to the sender or to the receiver. This analysis is the same as in Loder et. al. (2006). However, in Hermalin and Katz (2004), the message space is restricted to the positive quadrant.

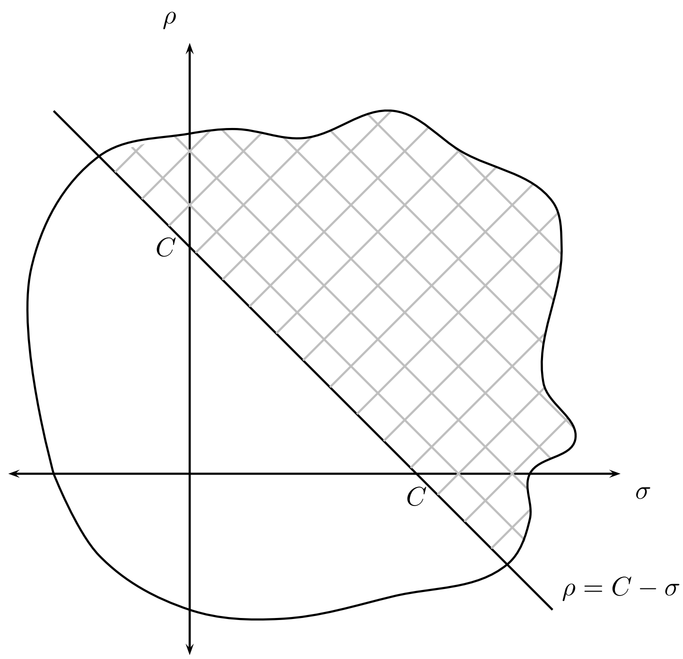

Figure 1: Possible message space

3.2 Private surplus of a message

We assume now that the ISP can charge sender prices (p^S) payable upon sending the message and/or receiver prices (p^R) payable upon opening a message. Notice that the ISP can only observe if the message has been opened but not if it has been read. This, as we will see, implies that negative prices do not serve the useful purpose of inducing consumers to read messages that it may have in other networks, such as telephone network.

Messages that are sent and read give the sender and receiver the following private surpluses:

$$s^S = \sigma - c^S - p^S$$

[1] Since the sender's benefit is 0 when her message is not read, the social surplus of a message that is sent and not read is $-(c^I + c^R + c^U)$. So it is never optimal for such a message to be sent.

and

$$s^R = \rho - c^R - p^R$$

Given (p^S, p^R), messages in the following set will be exchanged

$$SR(p^S, p^R) \equiv \{(\sigma, \rho) | \sigma \geq \max(c^S + p^S, 0), \rho \geq \max(p^R, 0)\}. \quad (2)$$

Notice that messages are read when $\rho \geq p^R$ and not when $\rho \geq p^R + c^R$ because c^R is sunk when the decision to read a message is made.

Figure 2 illustrates a possible distribution of preferences (without any consideration for density) in the (σ, ρ) space. With some arbitrary uniform prices p^S and p^R, efficient messages in $abdeca$ and in $ijghfi$ are not exchanged and inefficient messages in eik are exchanged.

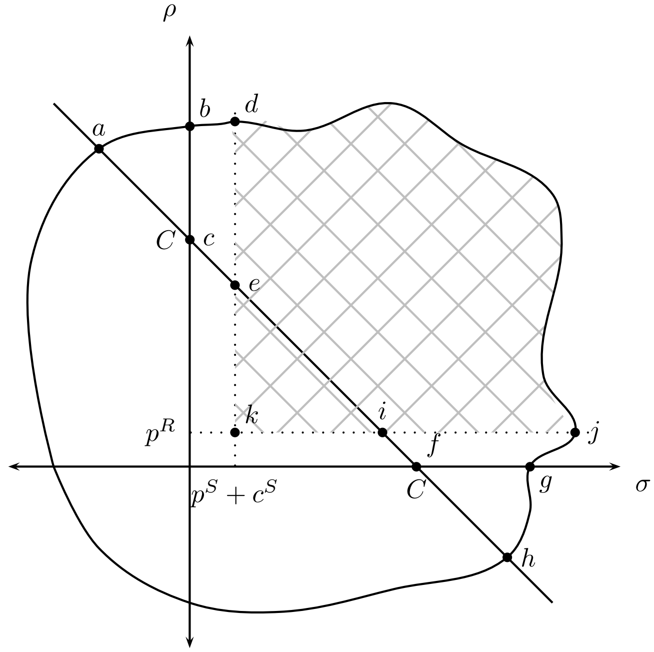

Figure 2: Messages that are sent and read given some combination of p^S and p^R

It is commonly considered, however, that we should be able to achieve efficient message exchange with optimal pricing. In the situation depicted in Figure 3, we could consider using perfectly discriminatory pricing to internalize all external effects. Such prices would involve, for instance, setting negative sender prices to induce the messages in abc to be sent, as well as negative receiver pricing to induce messages in fgh to be opened and read. We would need positive sender prices (or receiver prices) to discourage the bad messages in cfO to be sent. For example, MacDonald and Meriluoto (2005) suggest using prices $p^S = \sigma - c^R - c^U$ and $p^R = \rho - c^S - c^U$ (or $p^R = \rho$) to ensure that only and all good messages are exchanged. While this type of pricing may be valid for telephone networks, it is not valid in email networks for two key reasons.

First, 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 price cannot induce the receiver to read a message for which $\rho < 0$. Such a price would be willingly accepted, but no consumer would find it in their best interest to actually read such a message and therefore this type of pricing can merely distribute income but not induce efficiency.

Second, it is conceivable that the optimal sender price might satisfy $p^S < -c^S$, 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.

Thus, while prices, discriminatory or uniform, can be used to control the exchange of message in the positive quadrant in Figures 1 and 3, they cannot be used to induce efficient messages in areas *abc* and *fgh*.

It is useful to look at the current situation where no prices are being used, so $p^S = p^R = 0$. Given zero prices, the messages for which $\sigma \geq c^S$ will be sent if the sender can infer that they will be read. As the receiver's processing cost is sunk at the time of her decision to open and read the message, the message is read if $\rho \geq 0$. So, with zero pricing, the messages that are send and read are given in the crosshatched area in Figure 3.

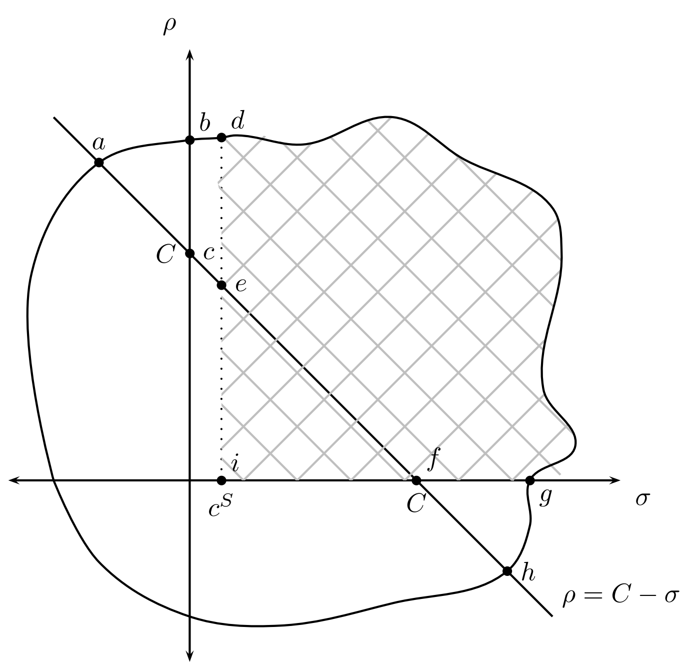

Figure 3: Messages that are sent and read given zero prices

With zero pricing, therefore, the messages in areas *abdec* and *fgh* are not exchanged despite being welfare improving. The messages in *abdec* are not sent because their value to the sender, net of the processing cost, is negative. The messages in fgh are not sent because their value to the receiver is negative and will therefore not be opened and read. Furthermore, messages in efi are exchanged despite being welfare deteriorating. These messages give both the sender and the receiver a positive utility, but the utilities sum up to less than the total cost of the message.

4 Optimal prices

4.1 Cost-Benefit Optimal Uniform Sender and Receiver Prices

We will first investigate the combination of uniform sender and receiver prices in the range $p^S \geq -c^S$ and $p^R \geq 0$ that maximize total social surplus of the email network.

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 exchanged. The total number of such messages is given by the double integral $\int_{p^S+c^S}^{\sigma^{\max}} \int_{p^R}^{\rho^{\max}} g(\rho)f(\sigma)d\rho d\sigma$. The total surplus is the definite double integral of the surplus per message over $\rho \in [p^R, \rho^{\max}]$ and $\sigma \in [p^S + c^S, \sigma^{\max}]$.

$$TSS(\rho, \sigma, p^S, p^R) = \int_{\sigma=p^S+c^S}^{\sigma^{\max}} \int_{\rho=p^R}^{\rho^{\max}} (\rho + \sigma - C)\, g(\rho)f(\sigma)d\rho d\sigma. \tag{3}$$

Integrating by parts gives

$$\begin{aligned} TSS(\rho, \sigma, p^S, p^R) = &\left[\rho^{\max} - p^R G(p^R) - \int_{p^R}^{\rho^{\max}} G(\rho)d\rho\right] \left(1 - F\left(p^S + c^S\right)\right) \\ &+ \left[\sigma^{\max} - \left(p^S + c^S\right) F\left(p^S + c^S\right) - \int_{p^S+c^S}^{\sigma^{\max}} F(\sigma)d\sigma\right] \left(1 - G\left(p^R\right)\right) \\ &- C\left(1 - F\left(p^S + c^S\right)\right)\left(1 - G\left(p^R\right)\right). \end{aligned} \tag{4}$$

TSS can be interpreted as follows. The first line is the expected surplus of send and read messages to the receiver. The first term on the first line is the expected receiver value of the messages that the receivers would willingly read. However, not all of these messages are sent, and therefore this term is multiplied by the number of messages actually sent. The second line is the expected surplus of messages to the sender. The first term on the second line is the expected sender value of messages that the sender would willingly send if they were read. This is multiplied by the number of messages actually read. The last line is the total cost of all messages sent and received.

Maximising (4) w.r.t. $p^S + c^S$ [2] and p^R subject to the constraints $0 \leq p^S + c^S \leq \sigma^{\max}$ and $0 \leq p^R \leq \rho^{\max}$ yields the following "best-response functions":

[2] This is equivalent to maximizing w.r.t. p^S.

$$p^S + c^S = C - \frac{\left(\rho^{\max} - p^R G\left(p^R\right) - \int_{p^R}^{\rho^{\max}} G(\rho) d\rho\right)}{1 - G\left(p^R\right)} \tag{5}$$

and

$$p^R = C - \frac{\left(\sigma^{\max} - \left(p^S + c^S\right) F\left(p^S + c^S\right) - \int_{p^S + c^S}^{\sigma^{\max}} F(\sigma) d\sigma\right)}{1 - F\left(p^S + c^S\right)}. \tag{6}$$

These best response functions show that the sender price is set to equal the total cost of the message minus the expected receiver value of the messages minus the sender's processing cost and that the receiver price is set to equal the total cost of messages minus the expected sender value of the messages. Notice the asymmetry in the two prices caused by the fact that the receiver's processing cost is sunk but the sender's processing cost is not. After deriving the general conditions for Ramsey prices and profit-maximizing prices, we will turn out attention to specific preference distributions to gain further insights about the three types of prices.

4.2 Ramsey prices

Consider now the welfare-maximizing prices subject to the ISP's break-even constraint, that is, the Ramsey prices. Most research in communication network pricing show that efficient prices do not cover the network provider's cost. We also show for the case of uniform distribution that the efficient prices sum up to less than the total cost of the message. However, as only a part of the total cost of a message is ISP cost, this inequality does not necessarily mean that the ISP would not break even. We will discuss this aspect further in subsection 5.2 where we examine Ramsey prices for a uniform distribution.

Assume for now that the efficient prices do not cover all ISP costs. Given that the break-even constraint will be binding, the Ramsey price can be obtained by substituting the ISP's break-even constraint $p^R + p^S = c^U$ into the total surplus in (4):

$$\begin{aligned} TSS(\rho, \sigma, p^S, p^R = c^U - p^S) = \\ \left[\rho^{\max} - (c^U - p^S) G(c^U - p^S) - \int_{(c^U - p^S)}^{\rho^{\max}} G(\rho) d\rho\right] \left(1 - F\left(p^S + c^S\right)\right) \\ + \left[\sigma^{\max} - \left(p^S + c^S\right) F\left(p^S + c^S\right) - \int_{p^S + c^S}^{\sigma^{\max}} F(\sigma) d\sigma\right] \left(1 - G\left(c^U - p^S\right)\right) \\ - C\left(1 - F\left(p^S + c^S\right)\right)\left(1 - G\left(c^U - p^S\right)\right). \end{aligned} \tag{7}$$

Maximizing (7) with respect to p^S gives the FOC that implicitly defines the Ramsey prices:

$$
\begin{aligned}
\frac{\partial TSS}{\partial p^S} &= (1 - F(p^S + c^S))g(c^U - p^S)(c^U - p^S) \\
-f(p^S + c^S) &\left[\rho^{\max} - (c^U - p^S)G(c^U - p^S) - \int_{(c^U - p^S)}^{\rho^{\max}} G(\rho)d\rho\right] \\
&-(1 - G(c^U - p^S))\left[f(p^S + c^S)(p^S + c^S)\right] \\
+g(c^U - p^S) &\left[\sigma^{\max} - \left(p^S + c^S\right) F\left(p^S + c^S\right) - \int_{p^S + c^S}^{\sigma^{\max}} F(\sigma)d\sigma\right] \\
+C &\left[(1 - G(c^U - p^S))f(c^S + p^S) - (1 - F(p^S + c^S))g(c^U - p^S)\right]. \qquad (8)
\end{aligned}
$$

The interpretation of (8) is as follows. Line 1 represents the change in expected surplus of received messages. The first term is the increase in expected receiver surplus caused by the effective reduction in the receiver price keeping the number of messages sent constant, and the second term is the reduction in receiver surplus caused by the reduction of the number of messages sent keeping the expected receiver surplus per message constant. Line 2 represents the change in expected consumer surplus of sent messages. The first term in line 2 is the reduction in expected surplus caused by the rising sender price keeping the number of messages read constant, and the second term is the increase in consumer surplus of sent messages caused by the increase in the number of messages that are read due to the resulting reduction in receiver pays price. Line 3 represents the change in the total cost of the messages when the composition of messages that are both sent and read changes. The first term is the reduction in cost caused by the reduction in the messages that are sent keeping the messages that are willingly read constant, and the second term is the increase in cost caused by the increase in the messages that are read keeping the messages that are sent constant.

Setting the FOC in (8) equal to zero and rearranging gives

$$
\begin{aligned}
&g(c^U - p^S)(1 - F(p^S + c^S))\left(C - c^U + p^S - \frac{\sigma^{\max} - (p^S + c^S)F(p^S + c^S) - \int_{p^S + c^S}^{\sigma^{\max}} F(\sigma)d\sigma}{1 - F(p^S + c^S)}\right) \\
&= f(p^S + c^S)(1 - G(c^U - p^S))\left(C - p^S - c^S - \frac{\rho^{\max} - (c^U - p^S)G(c^U - p^S) - \int_{c^U - p^S}^{\rho^{\max}} G(\rho)d\rho}{1 - G(c^U - p^S)}\right) \qquad (9)
\end{aligned}
$$

Thus, the Ramsey sender price (and thus effectively the receiver price) is set such that the change in total welfare caused by the increase in messages that are willingly read equals the total change in welfare caused by the reduction in messages that are willingly sent.

4.3 Profit-maximizing prices

A monopolist ISP profit is given by

$$\begin{aligned}\pi(p^S, p^R) &= \int_{p^S+c^S}^{\sigma^{\max}} \int_{p^R}^{\rho^{\max}} (p^R + p^S - c^U) g(\rho) f(\sigma) d\rho d\sigma \\ &= [1 - G(p^R)][1 - F(p^S + c^S)](p^R + p^S - c^U). \end{aligned} \tag{10}$$

Maximizing (10) with respect to p^R and p^S yield the following FOCs that implicitly define the profit-maximizing prices for the interior solution:

$$g(p^R)(p^R + p^S - c^U) = [1 - G(p^R)] \tag{11}$$

and

$$f(p^S + c^S)(p^R + p^S - c^U) = [1 - F(p^S + c^S)]. \tag{12}$$

As usual, the monopolist must balance the decreased quantity of messages with the increase in revenue on infra-marginal messages.

5 Uniform density functions for message preferences

Let us now consider a specific distribution function for the message preferences. Assume that σ is distributed uniformly in $\sigma \in uni\,[0, \sigma^{\max}]$ with density equal to $\frac{1}{\sigma^{\max}}$ and ρ is distributed uniformly in $\rho \in uni\,[0, \rho^{\max}]$ with density equal to $\frac{1}{\rho^{\max}}$. Notice that the density functions are independent but not necessarily identical.

5.1 Optimal uniform prices with uniform distribution for message preferences

Assuming that $p^R \geq 0$ and $p^S \geq -c^S$, all the messages for which $\rho \in [p^R,\, \rho^{\max}]$ and $\sigma \in [p^S + c^S,\, \sigma^{\max}]$ will be send and read and the surplus of each such message is $\rho + \sigma - C$.

The total surplus is

$$
\begin{aligned}
TSS(\rho,\sigma,p^S,p^R) &= \int_{\rho=p^R}^{\rho^{\max}} \int_{\sigma=p^S+c^S}^{\sigma^{\max}} \frac{(\rho+\sigma-C)}{\rho^{\max}\sigma^{\max}} d\rho d\sigma \\
&= \frac{1}{\rho^{\max}\sigma^{\max}} \int_{\sigma=p^S+c^S}^{\sigma^{\max}} \left(\int_{\rho=p^R}^{\rho^{\max}} (\rho+\sigma-C)\, d\rho \right) d\sigma \\
&= \frac{1}{\rho^{\max}\sigma^{\max}} \int_{\sigma=p^S+c^S}^{\sigma^{\max}} \left(\frac{(\rho^{\max})^2}{2} - \frac{(p^R)^2}{2} + (\rho^{\max}-p^R)(\sigma-C) \right) d\sigma \\
&= \frac{1}{\rho^{\max}\sigma^{\max}} [\left(\frac{(\rho^{\max})^2-(p^R)^2}{2} \right) (\sigma^{\max}-(p^S+c^S)) \\
&\quad + \left(\frac{(\sigma^{\max})^2-(p^S+c^S)^2}{2} \right) (\rho^{\max}-p^R) \\
&\quad - (\rho^{\max}-p^R)(\sigma^{\max}-(p^S+c^S))\, C] \\
&= \frac{(\rho^{\max}-p^R)(\sigma^{\max}-(p^S+c^S))}{\rho^{\max}\sigma^{\max}} \left(\frac{\rho^{\max}+p^R}{2} + \frac{\sigma^{\max}+(p^S+c^S)}{2} - C \right) \qquad (13)
\end{aligned}
$$

Maximising (13) w.r.t. p^S+c^S and p^R subject to the constraints $0 \le p^S+c^S \le \sigma^{\max}$ and $0 \le p^R \le \rho^{\max}$ yields the following best-response functions:

$$p^S + c^S = C - \frac{\rho^{\max}}{2} - \frac{p^R}{2} \qquad (14)$$

and

$$p^R = C - \frac{\sigma^{\max}}{2} - \frac{p^S+c^S}{2}. \qquad (15)$$

Because the expected receiver value of all read messages given uniform distribution is $\frac{\rho^{\max}}{2} + \frac{p^R}{2}$, we can confirm the general result that the sender price is set equal to the total cost of the message minus the expected receiver value minus the seller's processing cost. Similarly, because the expected sender value of all sent messages is $\frac{\sigma^{\max}}{2} + \frac{p^S+c^S}{2}$, we can confirm the general result that the receiver price is set equal to the total cost of the message minus the expected sender value of all sent messages.

The interior solution is given by

$$(p^S+c^S)^* = \frac{2C}{3} - \frac{2\rho^{\max}}{3} + \frac{\sigma^{\max}}{3} \qquad (16)$$

and

$$p^{R*} = \frac{2C}{3} - \frac{2\sigma^{\max}}{3} + \frac{\rho^{\max}}{3} \qquad (17)$$

and is valid if $0 \le 2C - 2\rho^{\max} + \sigma^{\max} \le 3\sigma^{\max}$ and if $0 \le 2C - 2\sigma^{\max} + \rho^{\max} \le 3\rho^{\max}$, that is if $\max[-2C+2\sigma^{\max}, C-\sigma^{\max}] \le \rho^{\max} \le C + \frac{\sigma^{\max}}{2}$. Two other solutions are found where one of the two non-negativity constraints holds with equality and the other one is not binding. These are

$$(p^S + c^S)^* = 0 \tag{18}$$

and

$$p^{R*} = C - \frac{\sigma^{\max}}{2}, \tag{19}$$

which is valid if $\sigma^{\max} < \min[-2C + 2\rho^{\max}, 2C]$, and

$$(p^S + c^S)^* = C - \frac{\rho^{\max}}{2} \tag{20}$$

and

$$p^{R*} = 0, \tag{21}$$

which is valid if $\rho^{\max} < \min[-2C + 2\sigma^{\max}, 2C]$. The fourth solution is found where both non-negativity constraints hold with equality, that is

$$(p^S + c^S)^* = 0 \tag{22}$$

and

$$p^{R*} = 0, \tag{23}$$

and is valid if $\sigma^{\max} \geq 2C$ and if $\rho^{\max} \geq 2C$. The conditions for the four types of solutions are illustrated in Figure 5.1.

Intuitively, when both $\sigma^{\max}$ and $\rho^{\max}$ are larger than $2C$ the exchange of messages should be encouraged by setting prices at their minima because, on average, the value of a message is greater than the cost of the message for both the sender and receiver. Setting prices above their minima would lead to a larger welfare loss due to eliminating "good" messages than a welfare gain due to eliminating "bad" messages. When both $\sigma^{\max}$ and $\rho^{\max}$ are small, there are more "bad" messages than "good" messages at minimum prices, and higher prices therefore increase welfare. When $\rho^{\max} \leq \min[-2C + 2\sigma^{\max}, 2C]$, receivers do not value messages very much compared to the cost of messages but senders do. Since a positive sender price eliminates messages with a small value to the sender and because the expected receiver value $\frac{\rho^{\max}}{2}$ is small in this case, a sender price increases welfare. Receiver price is not used because even if it would eliminate messages with small receiver value, these eliminated messages have a high expected value $\frac{\sigma^{\max}}{2}$to the sender. Similarly, when $\sigma^{\max} \leq \min[-2C + 2\rho^{\max}, 2C]\rho^{\max}$, senders do not value messages very much compared to the cost of those messages but receivers do. Since a positive receiver price eliminates messages with a small value to the receiver and because the expected sender value $\frac{\sigma^{\max}}{2}$ is small in this case,

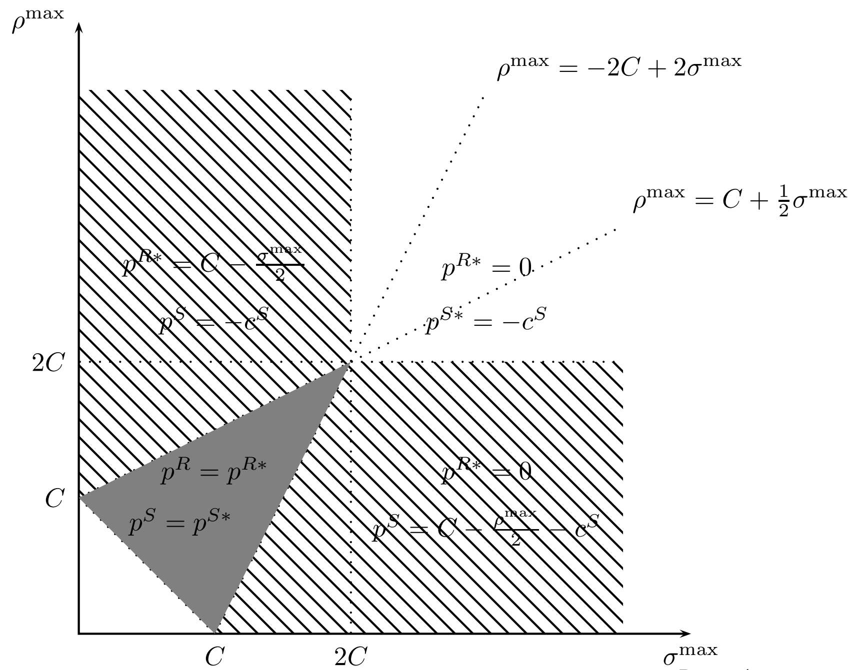

Figure 4: Optimal prices as functions of $\sigma^{\max}$ and ρ^{max}. $p^{R*} = \frac{1}{3}(2C - 2\sigma^{\max} + \rho^{\max}), p^{S*} = \frac{1}{3}(2C - 2\rho^{\max} + \sigma^{\max} - c^S.)$

a receiver price increases welfare. The sender price is set at its minimum because even if a higher price would eliminate messages with small sender value, these eliminated messages have a high expected value $\frac{\rho^{\max}}{2}$ to the receiver.

The two efficient prices sum up to

$$p^{S*} + p^{R*} = \frac{4c^U + c^S + 4c^R - \sigma^{\max} - \rho^{\max}}{3} \tag{24}$$

for the interior solution, and they cover the total cost of a message if

$$\rho^{\max} \leq C - 3c^S - \sigma^{\max}. \tag{25}$$

As the participation constraint of consumers is violated everywhere where (25) holds, the interior solution prices never cover the total cost of messages. It is trivial to see that the prices in the other solutions never cover the cost of the message, either.

Consider now identical distributions for message preferences, that is $\sigma^{\max} = \rho^{\max} = \gamma$. The efficient prices are now

$$p^{S*} = \frac{2C}{3} - \frac{1\gamma}{3} - c^S \tag{26}$$

and

$$p^{R*} = \frac{2C}{3} - \frac{1\gamma}{3} \tag{27}$$

if $\frac{1}{2} \geq \gamma \leq 2C$, and

$$p^{S*} = -c^S \tag{28}$$

and

$$p^{R*} = 0 \tag{29}$$

if $\gamma \geq 2C$. When $\gamma \leq 2C$, the receiver price is above zero and falling in γ and when $\gamma \leq 2C$ the receiver price is equal to zero. The sender price is set below the receiver price by the amount of the sender's processing cost. We will examine Ramsey prices to show that this asymmetry in efficient prices prevails once we impose a break-even constraint for the ISP. Hermalin and Katz (2004) examine the same message preferences but impose break-even constraint for the ISP. They find the sender price and the receiver price optimally divide the cost of the message between the sender and the receiver. The reason for the asymmetry in our prices stems from our assumption that c^R is sunk at the time the receiver makes her decision, something that Hermalin and Katz do not assume.

The two efficient prices sum up to

$$p^{S*} + p^{R*} = \frac{4C}{3} - \frac{2\gamma}{3} - c^S \tag{30}$$

if $\frac{C}{2} < \gamma \leq 2C$ and to $p^{S*} + p^{R*} = -c^S$ if $\gamma > 2C$. That is, the efficient prices sum up to less than the full cost of the message. However, it is not clear that they do not cover the ISP cost c^U. In fact, the ISP makes non-negative profit whenever

$$\frac{4C}{3} - \frac{2\gamma}{3} - c^S > c^U \tag{31}$$

or when

$$C + 3c^R > 2\gamma \tag{32}$$

Let us now determine how the efficient prices are affected by changes in the distribution of message preferences. First, assume that $\max[-2C + 2\sigma^{\max}, C - \sigma^{\max}] \leq \rho^{\max} \leq C + \frac{\sigma^{\max}}{2}$, i.e. that we have an interior solution. Now let $\sigma^{\max}$ increase keeping $\rho^{\max}$ constant, which implies that the average sender value of a message increases but the average receiver value stays constant. Thus we are stretching the uniform distribution to the right simultaneously reducing the density at any point. We can see from () that the optimal sender price goes up until $\sigma^{\max} = C + \frac{\rho^{\max}}{2}$ when the optimal sender price equals $p^{S*} + c^S = C - \frac{\rho^{\max}}{2}$ and any increases on $\sigma^{\max}$ thereafter have no impact on the optimal price. Similarly, let $\rho^{\max}$ increase keeping $\sigma^{\max}$ constant. The optimal receiver price increases until $\rho^{\max} = C + \frac{\sigma^{\max}}{2}$ when the optimal receiver price becomes $p^{R*} = C - \frac{\sigma^{\max}}{2}$ and any increases thereafter have no impact on the optimal price. Also, if both $\sigma^{\max}$ and $\rho^{\max}$ are increased simultaneously by the same amount, the optimal sender and receiver prices fall linearly from $p^{R*} = p^{S*} + c^S = \frac{C}{3}$ at $\sigma^{\max} = \rho^{\max} = \frac{C}{2}$ to $p^{R*} = p^{S*} + c^S = 0$ at $\sigma^{\max} = \rho^{\max} = 2C$.

As an example, assume that $\sigma^{\max} = 1.2C$ and $\rho^{\max} = 1.5C$. The optimal prices are now $((p^S + c^S)^*, p^{R*}) = (\frac{0.2C}{3}, \frac{1.1C}{3})$ and they are illustrated in Figure 5. The shaded area includes all sent messages of which the messages in the crosshatched area reduce social surplus. Notice that the prices are set such that the shaded area is a square. This graphical explanation holds for other distributions as well. Also, for the messages at the left and bottom boundaries of the sent message space (where the senders are indifferent between sending and not sending a message and where the receivers are indifferent between reading and not reading a message, respectively), there is an equal number of good and bad messages. This graphical explanation is specific to the assumption of uniform distribution where all the observations receive the same weight.

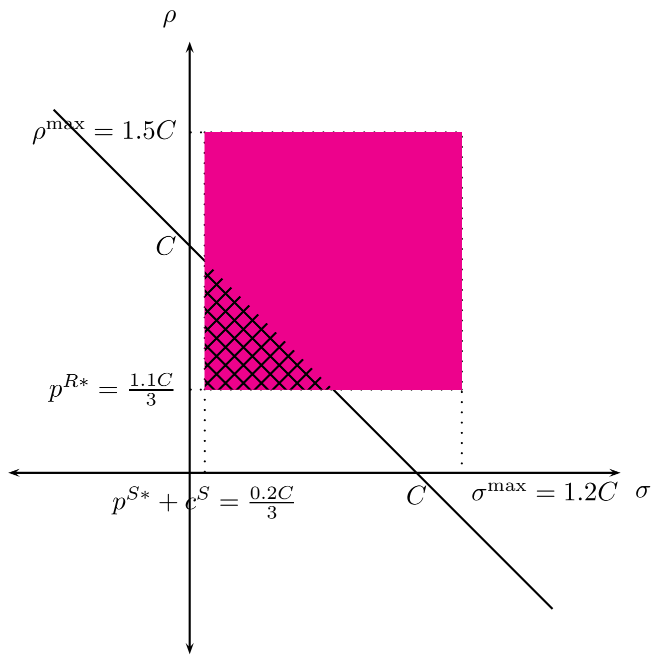

Figure 5: Efficient prices with uniform distribution when $\rho^{\max} = 1.5C$ and $\sigma^{\max} = 1.2C$

The graphical interpretation used in the above example can also be used to explain why optimal prices are set to their minima in other cases. If the price that would set the number of good messages equal to the number of bad messages at either left or the lower boundary of the sent message space is non-positive, the optimal price is set equal to its minimum.

5.2 Ramsey prices with uniform distribution for message preferences

Ramsey prices are found by maximizing total surplus in (13) with respect to p^S subject to break-even constraint: $p^R + p^S \geq c^U$. While we have shown that the efficient prices never cover the total cost of the message, we cannot show unambiguously that they do not cover the ISP cost.

If $\rho^{\max}$ and $\sigma^{\max}$ are such that we have an interior solution for the social welfare maximizing prices, then the ISP cost is covered whenever

$$\rho^{\max} \leq C + 3c^R - \sigma^{\max}. \tag{33}$$

For the corner solutions with one positive price, the ISP makes non-negative profit if $\sigma^{\max} \leq 2c^R$ or $\rho^{\max} \leq 2c^R$. Intuitively, because the efficient prices take into account the external effect to the other party and therefore the prices are the smaller the larger is the other party's utility, the ISP break-even constraint fails to hold when the maxima in the message preference distributions are large.

Figure 5.2 illustrates the areas in the $\rho^{\max}, \sigma^{\max}$ space where the first best efficient prices violate the ISP break-even constraint.

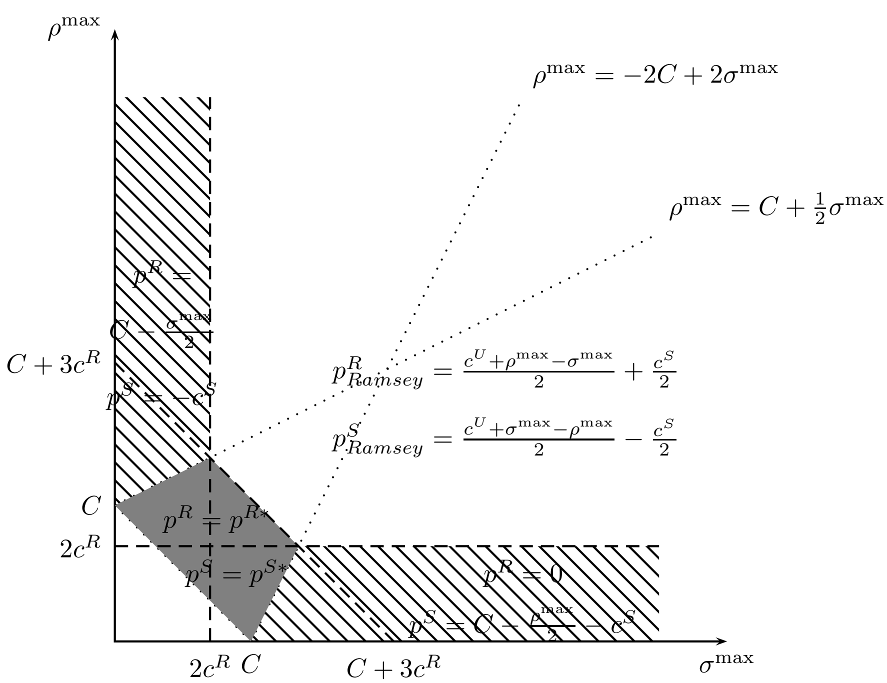

Figure 6: Parameter conditions for optimal prices given ISP break-even constraint

To find out the Ramsey prices that hold when the ISP break-even constraint binds, substitute $p^R = c^U - p^S$ into TSS and maximize w.r.t. p^S.

$$\begin{aligned} TSS_{Ramsey} = \\ \frac{(\rho^{\max} - c^U + p^S)(\sigma^{\max} - p^S - c^S)}{\rho^{\max}\sigma^{\max}} \left(\frac{\rho^{\max} + c^U - p^S}{2} + \frac{\sigma^{\max} + p^S + c^S}{2} - C \right) \\ = \frac{\sigma^{\max} + \rho^{\max} + c^S - C}{2\rho^{\max}\sigma^{\max}} \left((\rho^{\max} - c^U + p^S)(\sigma^{\max} - p^S - c^S) \right) \end{aligned} \tag{34}$$

so Ramsey prices are

$$p^S_{Ramsey} = \frac{c^U + \sigma^{\max} - \rho^{\max}}{2} - \frac{c^S}{2} \tag{35}$$

and

$$p^R_{Ramsey} = \frac{c^U + \rho^{\max} - \sigma^{\max}}{2} + \frac{c^S}{2} \tag{36}$$

so if the uniform distributions are identical, such that $\rho^{\max} = \sigma^{\max} = \gamma$, then

$$p^S_{Ramsey} = \frac{c^U}{2} - \frac{c^S}{2} < \frac{c^U}{2} + \frac{c^S}{2} = p^R_{Ramsey} \tag{37}$$

5.3 Profit-maximizing prices with uniform distribution for message preferences

Given uniform distributions given above, the profit-maximizing prices are given as

$$p^S = \frac{2\sigma^{\max} - \rho^{\max}}{3} + \frac{c^U - 2c^S}{3} \text{and} \tag{38}$$

$$p^R = \frac{2\rho^{\max} - \sigma^{\max}}{3} + \frac{c^U + c^S}{3}. \tag{39}$$

As with Ramsey prices, the profit-maximizing uniform prices are asymmetric such that the sender less than the receiver. When the preference distributions are identical, the difference in the two prices is the sender's processing cost as with Ramsey prices. However, while the sum of the Ramsey prices equate by construction to the total cost of the message, the profit-maximizing prices are larger.

6 Conclusions

In this paper we have shown how the nature of email technology affects the level of efficient and profit-maximizing prices as well as the optimal mix between sender and receiver pricing. The first-best outcome may not be achievable even with perfectly discriminatory prices because consumers cannot be induced to read efficient messages with negative receiver prices and there is a limit in the size of the subsidy that can induce senders to send efficient messages. Efficient uniform sender prices are decreasing in the magnitude of maximum sender value of the message distribution and at a minimum are equal to the negative of the sender's processing cost. Efficient prices are similarly decreasing in the magnitude of the maximum receiver value and at a minimum are equal to zero. Even with perfectly symmetric message preference distributions the optimal uniform prices are asymmetric in that the receiver pays more than the sender. However, the sender price is relatively large compared to the receiver price when the message preference distribution has a lot of mass for messages with high sender value and low receiver value.

We also show that the sum of the efficient uniform receiver and sender prices fails to cover all the message costs but does cover the ISP's costs when both the maxima in the message preference distributions are sufficiently small. The efficient prices given an ISP break-even constraint and the profit-maximizing prices are also asymmetric in that receivers pay more than the senders when the message preference distributions are symmetric.

References

[1] Armstrong, M., 1998, "Network Interconncection in Telecommunications." *The Economic Journal* 108(448), 545-564.

[2] Carter, M. and J. Wright, 1999, "Interconnection in Network Industries." *Review of Industrial Organiziation* 14 (1), 1-25.

[3] Dhebar, A. and S. S. Oren, 1985, "Optimal Dynamic Pricing for Expanding Networks." *Marketing Science* 4(4), pp. 336-351.

[4] Eaton, B. C., MacDonald, I. A. and L. Meriluoto, 2008, "Spam - solutions and their problems." *University of Calgary Working Paper* 2008-09.

[5] Einhorn, Michael A., 1993, "Biases in Optimal Pricing with Network Externalities." *Review of Industrial Organization* 8, pp. 741-746.

[6] Doyle, C. and J. Smith, 1998, Market structure in mobile telecoms: qualified indirect access and the receiver principle. *Information Economics and Policy* 10, 471-488.

[7] Hahn, J.-H., 2003, Nonlinear pricing of telecommunications with call and network externalities. Intenational Journal of Industrial Organization 21, 949-967.

[8] Hermalin, B. E. and M. L. Katz, 2004, "Sender or receiver: who should pay to exchange an electronic message?" *RAND Journal of Economics* 35(3), pp. 423-448.

[9] Jeon, D.-S., Laffont, J.-J. and J. Tirole, 2004, "On the "receiver-pays" principle". *RAND Journal of Economics* 35, 85-110.

[10] Kim, J.-Y. and Y. Lim, 2001, "An Economic Analysis of the Receiver Pays Principle." *Information Economics and Policy* 13, 231-260.

[11] Laffont, J.-J., P. Rey and J. Tirole, 1998a, "Network competition: I. Overview and nondiscriminatory pricing" *RAND Journal of Economics* 29(1), 1-37.

[12] Laffont, J.-J., P. Rey and J. Tirole, 1998b, "Network competition: II. Price discrimination." *RAND Journal of Economics* 29(1), 38-56.

[13] Littlechild, S., 1975, "Two-part Tariffs and Consumption Externalities." *Bell Journal of Economics and Management Science* 6, 661-670.

[14] Loder, T., Van Alstyne, M. and R. Wash "An Economic Response to Unsolicited Communication," *Advances in Economic Analysis & Policy* 6(1), Article 2, 2006.

[15] MacDonald, I. A. and L. Meriluoto, 2005, "Efficient Usage and Access Pricing in Telephone Networks." *International Journal of Industrial Organization* 23(7-8), pp. 615-623.

[16] Rohlfs, J., 1974, "A theory of interdependent demand for a communications service". *Bell Journal of Economics and Management Science* 5(1), Spring 1974, pp. 16-37.

[17] Squire, L., 1973, "Some aspects of optimal pricing for telecommunications." *Bell Journal of Economics and Management Science* 4, 515-525.