RAND Journal of Economics
Vol. 35, No. 3, Autumn 2004
pp. 423–448

Sender or receiver: who should pay to exchange an electronic message?

Benjamin E. Hermalin*

and

Michael L. Katz*

We examine the pricing implications of call externalities, the benefits enjoyed by the recipient of a message sent by another user. We show that, with or without a network-profitability constraint, efficient pricing requires consideration of demands, as well as costs. We present conditions under which equal charges for sending and receiving calls maximize welfare and profits. We also present conditions under which the receiving party's subsidizing the sender maximizes welfare and profits. Finally, we show that menus of pricing options can increase welfare and profits. None of these findings holds in the absence of call externalities.

1. Introduction

■ It takes two to tango. It also takes two to talk on the phone, exchange e-mails, share files, or video conference. In communicating, two (or more) parties take actions, receive benefits, and bear costs. Hence, consumption of a single message by multiple parties gives rise to external effects, also referred to as network effects. Previous authors have distinguished two types of effects: an *access externality,* whereby benefits accrue to existing members of a network when a new user joins and thus can send and receive messages that the original members value exchanging with her; and a *call externality,* which comprises the benefits enjoyed by a user who receives a message initiated by another user (see Hermalin and Katz (2003), Laffont and Tirole (2000), and Taylor (1994) for surveys of telecommunications externalities). In this article we explore the implications of call externalities for socially and privately optimal pricing.

With a few notable exceptions, previous theoretical work on communications pricing has tended to note the possibility of call externalities and then ignore them. This treatment typically is justified by one of two assumptions: either the receiving party enjoys no benefit from a message exchange, or the effects between the parties are internalized. The first assumption clearly is unrealistic. Were it correct, we would never answer the telephone or read our e-mail. The second

* University of California, Berkeley; hermalin@haas.berkeley.edu, katz@haas.berkeley.edu.

The authors wish to thank two referees and the Editor, as well as participants in Berkeley's IDS 270 seminar, May 2001, and the Berkeley-Stanford IO*fest*, October 2001 for valuable comments and suggestions. The authors also wish to acknowledge the support of the Walter A. Haas School of Business, through the Willis H. Booth and Edward & Mollie Arnold Chairs, and the University of California, through a Committee on Research Grant. The views expressed here are solely those of the authors. As such, they do not necessarily reflect those of the chair donors or the University of California.

assumption is applicable only to a limited set of situations in which either the communicating parties behave altruistically or have a repeated relationship.[1]

When message exchange generates benefits and costs for both parties, there are important differences between situations in which either party can initiate a message exchange ("two-way calling") and those in which only one party can do so ("one-way calling"). One-way calling has several interpretations. Message origination could be literally one-sided for technological reasons, as with paging and pay phones. Other technologies are two-way, but in many instances only one of the two parties knows there is value in communicating (e.g., a diner calling a restaurant to make a reservation). In such situations, it is reasonable to view only one of the two parties as the potential message initiator. Other situations, in which both parties know there's a value to communicating and it is technically feasible for either party to initiate a message exchange, are two-way calling situations.

An alternative interpretation of the distinction between one-way and two-way calling models is the following. In our two-way calling model below, a party may strategically delay sending a costly message, hoping the other party will initiate the exchange instead. For low-cost messages, such strategic behavior may be implausible and the situation better approximated by a pair of one-way calling models in which a party sends a message if and only if her expected value of exchange exceeds the price she must pay.

Other theoretical examinations of retail pricing in the presence of call externalities—Hahn (2003), Jeon, Laffont, and Tirole (2004), Kim and Lim (2001), Squire (1973), and Srinagesh and Gong (1996)—implicitly examine the one-way calling case.[2] Although any user in these models can both send and receive messages, the sets of messages sent and received are independent of each other. In effect, each consumer is modelled as participating in a pair of independent one-way situations. Squire examines socially optimal pricing under the assumption that all users have the same expected benefits from receiving a message. In this case, the call externality can be perfectly internalized by setting a receive price equal to the common expected value and then reducing the send price by the same amount. Srinagesh and Gong consider more general distributions and identify implicit welfare tradeoffs for two polar cases. In the first, receivers cannot refuse to accept messages, while in the second, receivers can selectively block low-value messages. As the authors note, these assumptions are often unrealistic. For instance, a consumer receiving a voice telephone message can refuse to answer but cannot base this decision on the value of the call. Hahn characterizes the profit-maximizing nonuniform price schedule for sending messages but does not allow for nonzero receive prices despite the fact that, when both parties benefit from a message exchange, it is feasible—and often optimal—to charge both parties positive prices.[3]

Jeon, Laffont, and Tirole and Kim and Lim, writing contemporaneously with us, allow for positive receive prices and, like us, demonstrate that positive receive prices can be welfare enhancing. These articles formulate the issue quite differently than we do, however. In particular, both articles are primarily concerned with deterministic values of calling (Kim and Lim exclusively so).[4] Even when noise is introduced, as in Jeon, Laffont, and Tirole, it is done in a way so that the

[1] If sending a message triggers a set number of incoming messages in reply, then call externalities will be internalized in the demand for sending messages and, thus, can be captured by the standard analysis of access externalities; see Willig (1979) for details. Hermalin and Katz (2003) develop a simple game-theoretic model in which users (partially) internalize call externalities by engaging in tit-for-tat message initiation.

[2] Acton and Vogelsang (1992) allow for call externalities, but they assume that the incremental value of a message to the receiver depends solely and deterministically on the order in which the message is received. The authors are concerned exclusively with deriving the demand for sending messages, rather than with the welfare effects of pricing.

[3] Examples of positive receive prices in the United States include collect calling, 800 numbers, and most wireless telephone service.

[4] Another difference is that Jeon, Laffont, and Tirole and Kim and Lim take minutes of communication to be the relevant units of analysis, while we take messages exchanged. For telephony, there is value to both approaches. For other electronic messages (e.g., e-mail and data files), messages seem the more appropriate unit of analysis. Unlike the present article, those articles are concerned primarily with distortions that arise from competition among networks possessing market power. In that sense, those articles have more in common with our companion paper, Hermalin and Katz (2001), which also considers network competition.

two parties' expected values of calling are perfectly correlated, which—as we show below—keeps the noise from being a source of inefficiency per se. In contrast, we are concerned primarily with pricing under uncertainty about the parties' values of message exchange and the difficulties this uncertainty creates for setting efficient prices. Finally, none of the articles cited above considers the possible *strategic* game between the parties as to who will be the sender and who the receiver.[5]

In summary, our article extends the literature by characterizing optimal send and receive prices for a variety of situations, including two-way calling and realistic settings in which receivers can refuse to accept messages but must do so on the basis of limited information about incoming messages. We also explore the effects of pricing menus.

After introducing notation and preliminary concepts, we examine one-way calling situations. In the absence of a binding profitability constraint for the communications network, the prices for sending and receiving a message should generally sum to less than the marginal message cost as a means of internalizing the benefits that the sender and receiver each enjoy from the actions of the other. In the presence of a binding profitability constraint, socially optimal pricing entails trading off the effects on the origination and acceptance of messages as the sending and receiving prices are varied along an iso-profit line. When the sender and receiver's expected values of message exchange are independently and identically distributed with an increasing hazard rate, we show that welfare maximization subject to a breakeven constraint entails setting equal send and receive prices. Under these conditions, equal send and receive prices also maximize the network provider's profits. In contrast, when the common distribution of message values has a decreasing hazard rate, we show that it can be socially and privately optimal to require one of the parties to pay more than cost in order to subsidize the other. We also establish that typically it is socially valuable to offer pricing options (e.g., a menu whose different elements apportion the cost burdens differently between sender and receiver) when call externalities are present.

We then consider two-way calling situations. We show that setting unequal send and receive prices creates a pricing option that can stimulate message exchange for much the same reasons as menus did in the one-way calling case (a party can choose to be a sender or receiver). In the two-way calling case, however, this option can lead to socially wasteful strategic delay by the parties. For the case of parties with a common, weakly convex power function distribution of message values, we show that splitting the costs of a message between the sender and receiver equally leads to greater welfare than does billing the sender only. Finally, we show by example that, in other situations, welfare can be greater when the sender is the sole payer than when message costs are equally divided between the sender and receiver.

2. Preliminaries

■ Consider two individuals, A and B, who may wish to communicate. We model the communication between A and B as the exchange of a single message, which can be a telephone call, an SMS message, a data file, a page, or an e-mail, for example. The sender initiates the communication (e.g., places a phone call) and the receiver accepts it (e.g., answers the phone).

We take as given that A and B have access to a network.[6] Doing so greatly simplifies the analysis, and one can often expect call-externality issues to dwarf access-externality issues in practice (e.g., phone service is nearly universal in the United States, but people's calling behavior is price sensitive for certain types of calls, such as long distance). We also abstract from repeated-play considerations: each party is motivated only by his or her private net benefits for a given potential message exchange.

We ignore income effects by assuming that consumers have quasi-linear utility functions. That is, user j's utility is $\chi v_j + y$, where v_j is party j's value or benefit from communication, y

[5] An article written contemporaneously with ours, Kim, Bae, and Won (2002), does analyze a simple version of such a game. Because they do not allow for postponing message exchange, we believe their formulation of the problem is not the natural one for many message-exchange settings (see also footnote 34 below).

[6] For analyses of network connection decisions, see Hahn (2003), Littlechild (1975), Rohlfs (1974), and Squire (1973).

is her consumption of a composite commodity comprising all other goods and services, and χ is equal to one if she exchanges a message and zero otherwise. We also assume that the marginal social utility of a dollar of income to a user is equal across all users, and we take total surplus, the sum of producer and consumer surplus, as our welfare measure.[7]

The cost of exchanging a message between A and B is $m > 0$.[8] One might hold the view that, for local wireline telephony, m is so small as to be meaningless. However, m is of greater significance for wireless and long-distance calling. Moreover, the costs associated with sending other forms of messages, such as video files, can be significant. Indeed, while Internet service providers often offer flat-rate services, many industry observers believe that traffic-sensitive charges will have to be widely introduced. Sonic Telecommunications International Ltd. already offers per-minute billing to television program creators who send video files over its telecommunications network.

Further, even when the network's cost of transmitting a message is small, users may still incur significant opportunity costs in terms of the time expended exchanging messages (e.g., answering a telemarketing call during dinner). These costs are modelled by allowing the value of v_j to be negative. In the presence of these costs, it can be efficient to charge positive message prices even when the network's cost of transmitting a message is zero, either to discourage message exchange (e.g., a tax on "junk" e-mail) or to raise revenues to subsidize the other party to engage in message exchange.

The parties' values of communicating, v_A and v_B, could be unknown to them at the time they make their send and receive decisions. Specifically, each individual has some prior knowledge (type, signal, etc.), $\omega_j \in \Omega_j$. The pair (ω_A, ω_B) has joint distribution $\Psi(\omega_A, \omega_B)$, which is common knowledge. Each (ω_A, ω_B) vector defines a joint distribution over (v_A, v_B).

Our focus is on whether prices can induce socially optimal outcomes. Under the first-best outcome, total surplus is maximized if all messages for which $v_A + v_B > m$ are exchanged and no messages for which $v_A + v_B < m$ are exchanged. Because v_A and v_B might be realized only after a message is sent and received, the first-best outcome is generally unattainable. A more realistic welfare standard is second-best, or information-constrained, efficiency: a message is exchanged if and only if the social expected value conditional on what the parties know exceeds the cost.

Most of our analysis concerns per-message pricing in which, if the message is successfully exchanged, the sender pays p and the receiver pays r. At several points we are interested in prices that cover the message cost. Prices such that $p + r \geq m$ are said to satisfy the network profitability constraint.[9]

Under the information structure presented above, either party may draw inferences about his or her value of message exchange from the other party's behavior, as the following two examples illustrate. In both, suppose that it is common knowledge that A is perfectly informed and B is uninformed (i.e., $\omega_A = v_A$ and ω_B is uninformative about v_B). Also assume that only A can initiate messages.[10] Because it is common knowledge that A knows v_A, B can rely on A to send a message if and only if $v_A \geq p$.

In the first example, suppose that the two parties always have equal values of message exchange, $v_A = v_B$. Because the two parties have identical message values, B can infer that $v_B \geq p$ whenever A sends a message. Hence, if $p = r = m/2$, it is an equilibrium for A to send a message if and only if $\omega_A \geq p$ and for B always to accept it. Under this equilibrium, a message

[7] Willig (1979) provides a rigorous justification of the use of the consumer surplus approach in the presence of income effects.

[8] Note that we are assuming m is not incurred when a message is sent but not accepted. In the analysis of optimal pricing subject to a network profitability constraint, m is the long-run economic cost incurred by the network. For the analysis of the social optimum absent a network profitability constraint, m can also be interpreted as the social marginal cost of a message, including any congestion effects suffered by network users.

[9] To focus on the effects of call externalities, we assume there are no network fixed costs.

[10] The conclusions of these two examples remain valid under the alternative assumption that only B—the uninformed party—can initiate messages.

is exchanged if and only if $v_A + v_B \geq m$; the first-best outcome is attained through the receiver's drawing inferences from the sender's behavior.[11]

Such inferences need not always be beneficial, however. Consider the opposite extreme in which the two parties' message values are perfectly negatively correlated, with $v_B = \mu - v_A$, where $m < \mu < 2m$. Moreover, suppose it is common knowledge that ω_A is distributed uniformly on $[0, \mu]$. Because $v_A + v_B = \mu > m$, it is efficient for A always to send a message and B always to accept it. Because of the inferences drawn by B from A's behavior, however, it is impossible to support an efficient outcome in the presence of the network profitability constraint: if A sends a message, then B infers that $v_A \geq p$ and B's expected payoff from accepting the message is

$$\begin{aligned}\mu - \mathbb{E}\{v_A \mid v_A \geq p\} - r &= \mu - \frac{\mu + \max\{0, p\}}{2} - r \\ &= \frac{\mu - 2m - \max\{0, p\}}{2} + p,\end{aligned}$$

where we have used the fact that $r = m - p$. B's expected payoff is nonnegative if and only if $p \geq 2m - \mu$, which implies that messages will be exchanged with probability $2(1 - m/\mu) < 1$. This probability shrinks to zero as $\mu \to m$. In other words, the ability to make inferences can lead to a "lemons" problem (Akerlof, 1970) that essentially eliminates message exchange even though all exchanges are socially desirable.

As these two examples illustrate, the possibility of complex inferences makes the analysis of a general model difficult.[12] Progress can be made if one assumes each party's information is relevant only for predicting his or her own value of communicating, because then neither party draws on the behavior of the other to form inferences about his or her own value of message exchange. To that end, we make the following assumption throughout our analysis.

Assumption 1. For $i = A, B$, $\Omega_i \subseteq \mathbb{R}$ and $v_i = \omega_i + \eta_i$, where η_i is a random variable satisfying $\mathbb{E}\{\eta_i \mid \omega_i\} \equiv \mathbb{E}\{\eta_i \mid \omega_i, \omega_j\} \equiv 0$.

Under this assumption, ω_i contains no additional information useful to j in predicting the expected value of v_j *conditional* on the information he or she already possesses, ω_j. Observe, however, that, *ex ante*, the parties' information, ω_i and ω_j, could be correlated. Observe too that ω_i is now party i's expected value of communicating.

Assumption 1 strikes us as reasonable when applied to the sender because the receiver's decision whether to accept often will not convey much information to the sender about the value to her of having sent the message. This lack of learning is particularly likely in markets where there are multiple potential senders and the receiver does not know the identity of the party sending the message at the time he decides to accept. In this latter situation, Assumption 1 plausibly applies to the receiver as well: it may be unreasonably complex for the receiver to draw inferences about the message's value to him from the fact that a message was sent. When, as here, the receiver knows who is sending him a message (recall there are only two parties in our model, and in other situations the receiver may have caller ID or see the sender's e-mail address), the assumption that the receiver infers nothing from being sent a message could be less innocuous. On the other hand, if the sender is known to send messages that vary greatly in value to the receiver, or if the receiver's message value strongly depends on his current opportunity cost of time (e.g., whether he's eating dinner), then the receiver's estimation of a message's value to him may be little influenced by the sender's having initiated the message.

[11] A variant of this example illustrates the potential value of a junk-e-mail tax. Suppose $m = 0$ and fix $r = 0$. Maintain the informational assumptions, except assume $v_B = v_A - \alpha$, $\alpha > 0$ (e.g., α is the receiver's fixed cost of opening and reading an e-mail). Assume, too, that $\mathbb{E}\{v_A \mid v_A \geq \alpha/2\} \geq \alpha$. Then the first best requires that e-mails be sent and accepted only when $v_A + (v_A - \alpha) \geq 0$ or $v_A \geq \alpha/2$. This result is achieved by charging a tax on sending e-mails equal to $\alpha/2$. A tax can be optimal when (i) the expected social value of message exchange is an increasing function of the sender's expected value, and (ii) the sender's value sometimes exceeds the social value.

[12] We examine the effects of such inferences more fully in Hermalin and Katz (2003).

To avoid trivial cases, we also make the following assumption throughout our analysis.

Assumption 2. $\Pr\{\omega_A + \omega_B > m\} > 0$ and $\underline{\omega}_A + \underline{\omega}_B < m$, where $\underline{\omega}_i$ is the infimum of the support of ω_i, $i = A, B$.

If the first part of the assumption were violated, then no message exchange would be efficient. If the second part were violated, then all message exchange would be efficient and, as we shall see, this pattern of exchange could trivially be supported by simple pricing.

3. One-way calling decisions

■ In this section we suppose that only party A can send a message. B's action is either to accept or to reject any message sent by A. An equilibrium comprises a pair of mutual best-response functions: a sending rule for A and an acceptance rule for B.

□ **Socially optimal pricing.** We first examine the information-constrained social optimum without concern as to whether the network provider earns nonnegative profits.[13] As the left panel of Figure 1 shows, it is information-constrained efficient to exchange all messages above the line $\omega_A + \omega_B = m$.[14] The right panel of Figure 1 shows the set of messages exchanged when users face prices p and r. Under Assumption 1, a weakly dominant strategy for A is to send a message if and only if $\omega_A \geq p$. Likewise a weakly dominant strategy for B is to accept the message if and only if $\omega_B \geq r$. In the case illustrated, p and r sum to m and thus satisfy the network-profitability constraint. But even if that constraint did not bind, it is clear that the space above the "L" defined by prices p and r cannot match the space of messages that are efficient to exchange. More formally, simple pricing can support information-constrained efficient message exchange if and only if there exist p and r such that $\Pr[\{\omega_A + \omega_B \geq m\} - \{\omega_A \geq p, \omega_B \geq r\}] = 0$ and $\Pr[\{\omega_A + \omega_B < m\} \cap \{\omega_A \geq p, \omega_B \geq r\}] = 0$.

Unless the distribution $\Psi(\cdot, \cdot)$ meets rather stringent conditions, it is impossible to achieve efficient message exchange using only simple pricing. But should there exist a pair of send and receive prices summing to more than m that supports the social optimum, clearly there also exists a pair of send and receive prices summing to m that supports the social optimum. Moreover, in the *absence* of a network-breakeven constraint, the socially optimal send and receive prices often will sum to strictly less than m. To see why, consider Figure 2, which illustrates the welfare effects of lowering r by Δ starting from a point at which $p + r = m$. The shaded triangle below the m–m line represents inefficient message exchange triggered by the price decrease (i.e., those messages for which the sum of the expected values of exchange is less than marginal cost). The cross-hatched area above the m–m line represents welfare-enhancing messages triggered by the price decrease (i.e., those for which the sum of the expected values of exchange exceeds marginal cost). Note that the efficiency loss from an inefficient message is no more than Δ. The efficiency gain from a welfare-enhancing message can, however, clearly be much larger than Δ. Hence, unless drawing values for message exchange in the shaded triangle is much more likely than drawing values in the cross-hatched area, a small reduction in price is welfare improving; that is, there exist conditions under which prices that violate the network-profitability constraint are optimal.

Making use of Assumption 2 and the conditions for information-constrained efficient exchange above, one can say more when ω_A and ω_B are independently distributed.

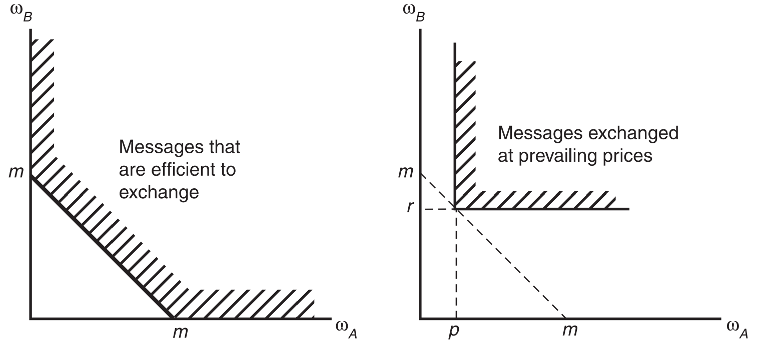
Proposition 1. Suppose the parties' expected values of message exchange, ω_A and ω_B, are independently distributed with positive densities defined everywhere on their supports. Use of a single price pair cannot achieve information-constrained efficient message exchange. If the network

[13] Doing so implicitly assumes that any funds needed to cover the network provider's costs can be raised in a nondistortionary manner. To the extent that subsidy funds are costly, they increase the *social* cost and, thus, should be accounted for in m.

[14] For convenience, we have drawn the figures for cases in which all messages have nonnegative values to both parties. Nothing in our analysis precludes messages with negative values.

FIGURE 1

EFFICIENT VERSUS EQUILIBRIUM MESSAGE EXCHANGE



provider is not subject to a profitability constraint, then socially optimal prices satisfy $p + r < m$. If the network provider is subject to the profitability constraint $p + r \geq m$, then socially optimal prices satisfy $p + r = m$.

Proof. See the Appendix.

Our main interest is in characterizing second-best pricing in the presence of a network-profitability constraint, so-called Ramsey pricing (Ramsey, 1927). We begin by examining settings in which ω_A and ω_B are independently distributed with differentiable distribution functions, $\Psi_A(\cdot)$ and $\Psi_B(\cdot)$, respectively. Let $S_i(\omega) \equiv 1 - \Psi_i(\omega)$ be the survival function. If $\psi_i(\cdot)$ is the density associated with $\Psi_i(\cdot)$, then $S_i'(\omega) = -\psi_i(\omega)$. In what follows, we assume for each $i = A$ or B that $\lim_{\omega\to\infty} \omega S_i(\omega) = 0$.[15]

Consider prices p and $r = m - p$, which satisfy the profitability constraint. Expected welfare, $\mathbb{E}W$, is

$$\begin{aligned}
\mathbb{E}W &= \int_{m-p}^{\infty} \left(\int_{p}^{\infty} \big(\omega_A + \omega_B - m\big)\psi_A(\omega_A)d\omega_A \right) \psi_B(\omega_B)d\omega_B \\
&= [1 - \Psi_B(m - p)] \int_{p}^{\infty} \omega_A \psi_A(\omega_A)d\omega_A \\
&\quad + [1 - \Psi_A(p)] \int_{m-p}^{\infty} \omega_B \psi_B(\omega_B)d\omega_B - m[1 - \Psi_A(p)][1 - \Psi_B(m - p)].
\end{aligned}$$

Using the definition of $S_i(\cdot)$ and integrating by parts,

$$\begin{aligned}
\mathbb{E}W &= S_B(m - p) \left[pS_A(p) + \int_{p}^{\infty} S_A(\omega)d\omega \right] \\
&\quad + S_A(p) \left[(m - p)S_B(m - p) + \int_{m-p}^{\infty} S_B(\omega)d\omega \right] - mS_A(p)S_B(m - p),
\end{aligned}$$

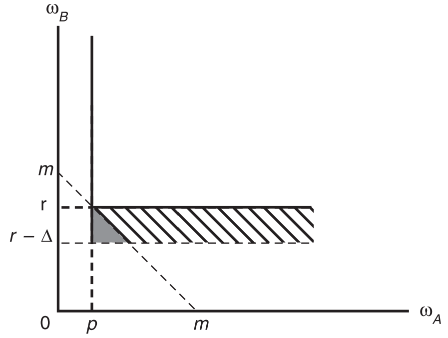
which simplifies to

$$\mathbb{E}W = S_B(m - p) \int_{p}^{\infty} S_A(\omega)d\omega + S_A(p) \int_{m-p}^{\infty} S_B(\omega)d\omega. \tag{1}$$

[15] The limit condition is clearly satisfied by any distribution with a bounded support. It is also satisfied by, for example, the exponential, the normal, the log-normal, the logistic, and Pareto distributions (the last only if the elasticity parameter exceeds one). A distribution that does not satisfy this limit condition is the Cauchy (if $\Psi(\cdot)$ were Cauchy, then $\lim_{\omega\to\infty} \omega S_i(\omega) = 1/\pi$).

FIGURE 2

WELFARE EFFECTS OF LOWERING THE RECEIVE PRICE



Suppose that $\Psi_A(\cdot)$ and $\Psi_B(\cdot)$ are identical. Differentiating expression (1) with respect to p yields the first-order condition

$$S'(p)\int_{m-p}^{\infty} S(\omega)d\omega - S'(m-p)\int_{p}^{\infty} S(\omega)d\omega = 0, \tag{2}$$

where $S(\cdot)$ is the common survival function.

This first-order condition has a simple and intuitive interpretation. Consider the two components of the first term. $S'(p)$ is the reduction in the sender's probability of initiating a message exchange as the send price, p, is increased, and $\int_{m-p}^{\infty} S(\omega)d\omega$ is the receiver's expected consumer surplus from message exchange. The second term represents the increase in the receiver's probability of accepting a message as the receive price, $m - p$, is reduced times the sender's expected consumer surplus from message exchange.

Notice that it can never be socially optimal to set p or r outside of the supports of Ψ_A and Ψ_B, respectively. In particular, if either p or r were set above the supports of Ψ_A and Ψ_B, respectively, expected welfare would be zero, which cannot be optimal given Assumption 2. If the send price were set below the support of Ψ_A, then $S'(p)$ in (2) would be zero and the derivative of welfare would be positive. A similar argument shows that the derivative of welfare with respect to p is negative when $m - p$ is below the support of Ψ_B.

Clearly, $p = m/2$ solves the first-order condition, equation (2). By construction, $r = m - p = m/2$. If the left-hand side of (2) is nonnegative for all $p < m/2$ and nonpositive for all $p > m/2$, then equal send and receive prices are a global optimum. Thus, equal prices are optimal if

$$\frac{S'(p)}{\int_p^{\infty} S(\omega)d\omega} \tag{3}$$

is nonincreasing in p.

By definition, the hazard rate of $\Psi(\cdot)$ is $\psi(\omega)/S(\omega) = -S'(\omega)/S(\omega)$. A nondecreasing hazard rate is therefore equivalent to $S'(\omega)/S(\omega)$ being nonincreasing in ω, which is to say $S(\cdot)$ being log concave. By Lemma A1 in the Appendix, the log concavity of $S(\cdot)$ implies the log concavity of $\int_p^{\infty} S(\omega)d\omega$. By the definitions of the survival function and log concavity, both terms in the following expression are then nonincreasing, negative functions of p:

$$\left(\frac{S'(p)}{S(p)}\right)\left(\frac{-S(p)}{\int_p^{\infty} S(\omega)d\omega}\right). \tag{4}$$

Hence, the product of these two terms must be positive and nondecreasing, from which it immediately follows that expression (3), which is negative one times this product, is nonincreasing.[16]

[16] We are very grateful to an anonymous referee for suggesting this decomposition.

Therefore, if the common distribution has a nondecreasing hazard rate, then prices that divide the cost of a message evenly between the sender and receiver are socially optimal given a breakeven constraint. Moreover, if the common distribution has an everywhere nondecreasing hazard rate that is strictly increasing around $m/2$, then this is the unique socially optimal price pair. Examples of distributions with everywhere increasing hazard rates include the normal distribution, the logistic distribution, and the uniform distribution, the last of which corresponds to a linear "demand" for messages.

Now consider a distribution with a constant hazard rate. First note that the support of any distribution with a nonincreasing hazard rate must have a finite infimum, $\underline{\omega}$, and an infinite supremum.[17] In the case of a constant hazard rate, the survival function must be $S(\omega) = e^{(\underline{\omega}-\omega)/\mu}$ for some positive constant μ. Hence, $\mathbb{E}W = 2\mu e^{(2\underline{\omega}-m)/\mu}$ for all $p \in [\underline{\omega}, m - \underline{\omega}]$. Therefore, when message exchange values are independently and identically distributed with a constant hazard rate, *any* send price between $\underline{\omega}$ and $m - \underline{\omega}$ is socially optimal in the presence of a network profitability constraint.[18]

Finally, consider a distribution with an everywhere decreasing hazard rate (e.g., the Pareto distribution, which corresponds to a constant elasticity demand for messages). If the hazard rate is everywhere decreasing, then $S(\cdot)$ is log convex, and it is straightforward to modify Lemma A1 to show that the log convexity of $S(\cdot)$ implies the log convexity of $\int_p^\infty S(\omega)d\omega$. In this case, expression (3) is increasing in p and it is socially optimal to set prices at a "corner," that is, to set either the send or receive price equal to $\underline{\omega}$ and set the complementary price equal to $m - \underline{\omega} < \infty = \bar{\omega}$. Because of the assumed symmetry, it does not matter whether the send or receive price is set at the infimum. Observe that this last point means that the sender and receiver optimally face different prices even when they have the *same* constant elasticity of demand for message exchange (i.e., when Ψ is the Pareto distribution).

Summarizing the analysis above, we have the following proposition.

Proposition 2. Suppose the network is subject to a profitability constraint and the parties' expected values of message exchange, ω_A and ω_B, are independently and identically distributed.

(i) If the hazard rate is everywhere increasing, then prices that divide the cost of a message equally between the sender and receiver are the unique socially optimal prices.
(ii) If the hazard rate is constant, then any prices such that $\underline{\omega} \le p \le m - \underline{\omega}$ and $r = m - p$ are socially optimal, where $\underline{\omega}$ is the infimum of the common support.
(iii) If the hazard rate is everywhere decreasing, then there are two socially optimal price pairs: one in which the send price equals $\underline{\omega}$ and one in which the receive price equals $\underline{\omega}$. In each case, the complementary price is set at $m - \underline{\omega}$.

One can also derive sufficient conditions for a monotone hazard rate expressed in terms of the underlying density function, $\psi(\cdot)$. Consider a nondecreasing hazard rate. An obvious condition is that $\psi(\cdot)$ be nondecreasing. A more interesting one is that $\psi(\cdot)$ be log concave. By a theorem of Prékopa (1971) (Pecarić, Proschan, and Tong, 1992), if the density function is log concave, then the survival function is log concave as well.[19]

---

[17] It is readily shown that $S(\omega) = e^{-\int_{\underline{\omega}}^{\omega} h(z)dz}$, where $h(\cdot)$ is the hazard rate. Take $h(\cdot)$ to be nonincreasing. If $\underline{\omega} = -\infty$, then, for any $\omega > -\infty$, $\int_{\underline{\omega}}^{\omega} h(z)dz > \lim_{z\downarrow-\infty}(\omega - z)h(\omega) = \infty$; that is, the distribution is not well defined. Now suppose the supremum, $\bar{\omega}$, were finite. Then $\int_{\underline{\omega}}^{\bar{\omega}} h(z)dz = \int_{\underline{\omega}}^{\hat{\omega}} h(z)dz + \int_{\hat{\omega}}^{\bar{\omega}} h(z)dz$ for some $\hat{\omega} \in (\underline{\omega}, \bar{\omega})$. The first integral on the right-hand side of the equation is finite. Because $h(\cdot)$ is nonincreasing, the second integral on the right-hand side is less than or equal to $h(\hat{\omega})(\bar{\omega} - \hat{\omega}) < \infty$. Thus, $S(\bar{\omega}) > 0$, which contradicts the definition of the survival function.

[18] Recall that $\underline{\omega} < m/2$ by Assumption 2.

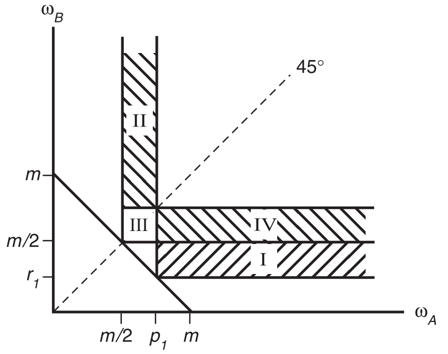
[19] Prékopa's theorem states that if B_1 and B_2 are measurable subsets of $\mathbb{R}$ under the measure induced by $\psi(\cdot)$ and $\psi(\cdot)$ is log concave, then

$$\log \Pr\{\alpha B_1 + (1-\alpha)B_2\} \ge \alpha \log \Pr\{B_1\} + (1-\alpha)\log \Pr\{B_2\},$$

$\alpha \in [0, 1]$. Setting $B_k = [p_k, \infty)$, one obtains $\alpha B_1 + (1-\alpha)B_2 = [\alpha p_1 + (1-\alpha)p_2, \infty)$, and it follows that $S(\cdot)$ is log concave.

FIGURE 3

EFFECTS OF A MOVE TOWARD EQUAL SEND AND RECEIVE PRICES



Corollary 1. Suppose that ω_A and ω_B are independently and identically distributed according to the density function $\psi(\cdot)$. If $\psi(\cdot)$ is log concave, then prices that divide the cost of a message equally between the sender and receiver are socially optimal given a breakeven constraint.

The intuition underlying this result can be seen graphically using Figure 3. Consider a move from prices (p_1, r_1), with $p_1 > r_1$, to equal prices. Messages in area I are lost, but messages in areas II and III are gained. By symmetry, messages in area II are equivalent to messages in area IV. Consider any point, (ω_A, ω_B), in area I and match that point with a new point, $\big(\omega_A - (p_1 - m/2),\ \omega_B + (p_1 - m/2)\big)$. The new point lies in areas III or IV. By construction, (ω_A, ω_B) yields the same welfare as the new point, but the new point lies closer to the 45° line than (ω_A, ω_B). When the common density is log concave, each point closer to the 45° line gets at least as much weight as the corresponding point farther from that line (see Lemma A2 in the Appendix). It follows that the expected welfare gains, areas III and IV, are at least as large as the expected welfare loss, area I, when $\psi(\cdot)$ is weakly log concave. A similar argument applies to cases in which the initial send price is less than the receive price.[20]

Proposition 2 assumes that ω_A and ω_B are independently and identically distributed. Suppose that ω_A and ω_B are distributed identically but not independently. If (ω_A, ω_B) pairs closer to the 45° line are relatively more likely than are pairs away from the 45° line, then equal prices are optimal. If closer pairs are not more likely, then unequal prices are in some cases optimal. For instance, if $\omega_B \equiv \omega_A$, then equal send and receive prices would always be optimal regardless of the other properties of $\Psi(\cdot, \cdot)$ because all weight is on the 45° line. Consider, in contrast, perfect negative correlation. In particular, suppose that ω_A has a marginal distribution that is symmetric about mean μ and that $\omega_B \equiv 2\mu - \omega_A$. In this case, the distribution function is critical. Because the social return from a completed call is constant at $2\mu - m$, the social problem reduces to maximizing the probability of a completed call,

$$\Pr\{p \leq \omega_A \leq 2\mu - (m - p)\} = \Psi^A(2\mu - m + p) - \Psi^A(p)$$

(where $\Psi^A(\cdot)$ is the marginal distribution of ω_A). The problem becomes one of finding an interval of length $2\mu - m$ on the support of ω_A with maximum probability. If the marginal density were a triangular, "$\wedge$-shaped" density on $[0, 2\mu]$, then equal prices would be optimal. If, however, this density were "$\vee$ shaped" on $[0, 2\mu]$ with a minimum at μ, then it would be optimal to make either A or B pay the full cost.

[20] This figure can also be used to illustrate the result that equal prices are socially optimal when $\psi(\cdot)$ is nondecreasing. Fixing ω_A, every message in area IV is more valuable than any message in area I (i.e., has a greater ω_B value). When the density is nondecreasing, the mass of lost messages in area I is no greater than the mass of gained messages in area IV, and therefore expected welfare rises when moving from unequal prices to equal prices.

For a class of distribution functions, we can also characterize asymmetric cases.

Proposition 3. Suppose that the parties' expected values of message exchange, ω_A and ω_B, are independently distributed according to differentiable distribution functions with associated densities $\psi_A(\cdot)$ and $\psi_B(\cdot)$, respectively. If

(i) $\psi_j(\omega)$ crosses $\psi_i(\omega)$ once from above at $\hat{\omega} \geq m/2$ (i.e., $\psi_j(\omega) > \psi_i(\omega)$ if $\omega < \hat{\omega}$ and $\psi_j(\omega) < \psi_i(\omega)$ if $\omega > \hat{\omega}$) and

(ii) the hazard rates associated with $\psi_i(\cdot)$ and $\psi_j(\cdot)$ are nondecreasing,

then party i pays more than party j under any socially optimal pricing scheme that satisfies the network profitability constraint.

Proof. See the Appendix.

The following example satisfies the assumptions of Proposition 3: $m = 1$ and both ω_A and ω_B are distributed on the unit interval, the former as ω_A^κ, with the constant $\kappa \geq 2$, and the latter uniformly. It is readily shown that (a) ψ_B crosses ψ_A once from above at or above $\omega = 1/2$, and (b) the hazard rates are nondecreasing. It thus follows that it is optimal for the sender to pay more than the receiver. For instance, if $\kappa = 2$, then $p \approx .554$ and $r \approx .446$.

Condition (i) in Proposition 3 implies that Ψ_i dominates Ψ_j in the sense of first-order stochastic dominance (see the proof of Lemma A3 in the Appendix). It is, however, a stronger assumption than first-order stochastic dominance. A stronger assumption is required because what matters is not only who values communicating more in expectation, but also the relative dispersions of their expected values of communication. If, say, j's expected value were very tightly clustered about some $\tilde{\omega}_j > m/2$, then one would do best to set j's price at approximately $\tilde{\omega}_j$ and, thus, i's price *below* $m/2$ even if i on average had the higher expected value of communicating.

To understand better the role of distributions, it is again helpful to consider the extreme cases of perfect positive and negative correlation. Let $\omega_B \equiv \alpha + \beta\omega_A$. Suppose first $\beta > 0$; that is, there is positive correlation as might apply if there were essentially common expected value to exchanging the message (e.g., where to meet for dinner). It is readily shown that the information-constrained socially optimal breakeven prices are

$$p = \frac{m - \alpha}{1 + \beta} \quad \text{and} \quad r = \frac{\beta m + \alpha}{1 + \beta}.$$

Observe that these prices do not depend on the distributions of ω_A and ω_B. In particular, it is possible with positive correlation that the party who faces the higher price is likely to have the lower expected value of message exchange (e.g., $\alpha + \beta\omega$ crosses ω from above, $\alpha/(1-\beta) > m/2$, and most of the weight is on $\omega_A \geq \alpha/(1-\beta)$). Observe, too, that when $\alpha \neq 0$, the socially optimal prices are nonzero even if $m = 0$; that is, it is beneficial to have one party subsidize the other.

Now consider $\beta < 0$ (i.e., perfect *negative* correlation). One might think of this example as a situation in which A needs a favor from B or is trying to get B to honor a costly obligation—the more valuable the request is to A, the less B's expected value from receiving the request. Given $\beta < 0$ and the network profitability constraint, messages are exchanged under the information-constrained social optimum if and only if

$$\Pr\left\{\frac{p - (m - \alpha)}{|\beta|} \geq \omega_A \geq p\right\} > 0$$

for some p. If that probability is positive for some p, then the optimal p solves

$$\max_p \int_p^{\frac{p-(m-\alpha)}{|\beta|}} \Big[(1 + \beta)\omega_A - (m - \alpha)\Big]\psi(\omega_A)d\omega_A. \tag{5}$$

Clearly, the solution to (5) depends, generically, on the distribution function when the expected values of message exchange are negatively correlated.

□ **Profit-maximizing prices.** Having characterized socially optimal prices, we now examine a profit-maximizing network. To do so, it is useful to divide the monopolist's problem into two steps. First, suppose the monopolist's profit margin is fixed at π per message exchanged: $p+r-m=\pi$. When ω_A and ω_B are independently distributed, the profit maximizer chooses a price pair from this line that maximizes the probability of message exchange, $S_A(p)S_B(r) = S_A(p)S_B(\pi+m-p)$. The corresponding first-order condition for an interior maximum is

$$S_A'(p)S_B(\pi+m-p) - S_A(p)S_B'(\pi+m-p) = 0. \tag{6}$$

When $S_A(\cdot) \equiv S(\cdot) \equiv S_B(\cdot)$, $p = (\pi+m)/2$ is a solution, which implies $r = (\pi+m)/2$. Equation (6) is sufficient, as well as necessary, if the left-hand side is a decreasing function of p. It is apparent that this condition is satisfied when the hazard rate is increasing (i.e., $S(p)$ is log concave).[21]

One can also apply the logic of our earlier welfare analysis to a common distribution function with a constant or everywhere decreasing hazard rate. With an everywhere decreasing hazard rate, for instance, there are two profit-maximizing price pairs for a given profit-margin: (i) $p = \underline{\omega}$ and $r = \pi + m - \underline{\omega}$, and (ii) $r = \underline{\omega}$ and $p = \pi + m - \underline{\omega}$.

Once optimal prices p and r for a given π are found, the second step is to maximize expected profits with respect to π. Define $p^*(\pi)$ as a solution to $\max_p S_A(p)S_B(\pi+m-p)$. For any $\pi > 0$, $p^*(\pi)$ is a send price that maximizes profits subject to the constraint that $p+r-m=\pi$. The profit maximizer's problem is to

$$\max_{\pi} S_A\big(p^*(\pi)\big)S_B\big(\pi+m-p^*(\pi)\big)\pi.$$

By Assumption 2, there exist price pairs that yield positive profits. Hence, the profit-maximizing value of π is positive, and the usual monopoly distortion of inefficiently few sales holds.

To summarize:

Proposition 4. Suppose the parties' expected values of message exchange, ω_A and ω_B, are independently and identically distributed. The sum of the profit-maximizing send and receive prices exceeds the sum of welfare-maximizing send and receive prices. Moreover,

(i) If the hazard rate is everywhere increasing, then prices that divide the cost of a message equally between the sender and receiver are the unique profit-maximizing prices.
(ii) If the hazard rate is constant, then $S(\omega) = e^{(\underline{\omega}-\omega)/\mu}$, μ a positive constant and $\underline{\omega}$ the infimum of the common support, and the profit-maximizing margin is μ. Any prices such that $\underline{\omega} \le p \le \mu + m - \underline{\omega}$ and $r = \mu + m - p$ maximize profits.
(iii) If the hazard rate is everywhere decreasing, then there are two profit-maximizing price pairs: in one the send price equals $\underline{\omega}$ and in the other the receive price equals $\underline{\omega}$.

A profit-maximizing network inefficiently sets a positive margin. In the cases covered by this proposition, however, the profit-maximizing monopolist imposes no further distortion in terms of the division of the cost and markup between sender and receiver. In other cases, there can be two types of distortion. In addition to setting the markup too high, the profit-maximizing network may distort the relative levels of the send and receive prices. This second distortion arises because—for a given margin—the profit maximizer chooses the relative prices to maximize the probability that a message will be exchanged, while the welfare maximizer takes into account the value of the messages exchanged.

To understand this second distortion more fully, it is informative to consider the behavior of a profit-maximizing network restricted to setting a margin of $\bar{\pi}$ as $\bar{\pi}$ goes to zero. Suppose that a limit of $p^*(\bar{\pi})$ exists as $\bar{\pi}$ goes to zero. If $S_A(\cdot)$ and $S_B(\cdot)$ are continuous, then the limit must equal

[21] The graphical analysis of Figure 3 again illustrates the forces at work when the common underlying *density* function is increasing or log concave. The only differences from the earlier welfare analysis are that the monopolist attaches equal dollar value to any two points that involve message exchange and m is replaced by $m+\pi$.

$p^*(0)$ and maximize $S_A(p)S_B(m - p)$.[22] Maximization of the probability of message exchange can lead to different prices than does maximization of the social objective function, expression (1). Consider, for instance, the example examined after Proposition 3, with $S_A(\omega) = 1 - \omega^2$, $S_B(\omega) = 1 - \omega$, and $m = 1$. The prices $p = .554$ and $r = .446$ maximize welfare, but $p = .577$ and $r = .423$ maximize $S_A(p)S_B(1 - p)$. As this example demonstrates, even highly competitive markets might perform poorly in the presence of call externalities.

□ **Menus.** So far we have assumed that the network sets a single pair of send and receive prices. In some instances, the network may be able to offer a menu of price pairs from which the sender then chooses. In this subsection we examine the welfare effects of adding options to an arbitrary menu. The potential value of menus is that they provide scope for a sender with a high-value message to assume a greater share of the cost in order to increase the odds that the receiver accepts.

Fix a single price pair, (p_1, r_1), and consider the consequence of adding an option, (p_2, r_2). Under this menu, if A chooses to send a message and picks option i, she pays p_i and B pays r_i if he accepts (the receiver knows which receive price he faces when he makes his decision). If $p_2 > p_1$ and $p_2 + r_2 \geq m$, then welfare must be weakly enhanced by adding the option: all messages that would have been exchanged under the single price-pair regime will be exchanged under the menu, and some surplus-generating messages that would not have been exchanged may now be exchanged.[23] For example, if ω_A and ω_B are independently distributed, additional messages are exchanged when there exists an $\hat{\omega}_A$ such that $\Psi_A(\hat{\omega}_A) < 1$ and $(\hat{\omega}_A - p_1)S_B(r_1) < (\hat{\omega}_A - p_2)S_B(r_2)$.

Intuitively, the sender can be thought of as a monopoly producer of messages that she "sells" to the receiver at a price of r. Allowing her to charge a lower price cannot reduce welfare, and it can raise welfare to the extent that the reduced price is closer to the monopoly price for some types of sender. This monopoly metaphor also points out that adding a receive price to a menu that is greater than the lowest existing price can reduce welfare by giving the monopoly sender added scope to exercise market power.

While the discussion has been couched in terms of adding a new option to a single price pair, the argument applies with equal force when adding a new option to a menu comprising any number of price pairs. Hence, we have established the following.

Proposition 5. Consider an arbitrary menu of price pairs. Adding an option for which the send price is greater than any existing send price and the sum of the send and receive prices is greater than or equal to the marginal cost of a message weakly raises welfare. Expanding the menu by adding an option with a send price that is less than an existing send price, or for which the sum of the send and receive prices is less than the marginal cost of a message, can reduce welfare.

Offering a menu can also be more profitable than offering a single price pair. An analysis similar to that which led to Proposition 5 demonstrates the following: profits weakly rise when an option is added for which (i) the send price is greater than any existing send price and (ii) the sum of the send and receive prices is greater than or equal to the sum of any existing option.

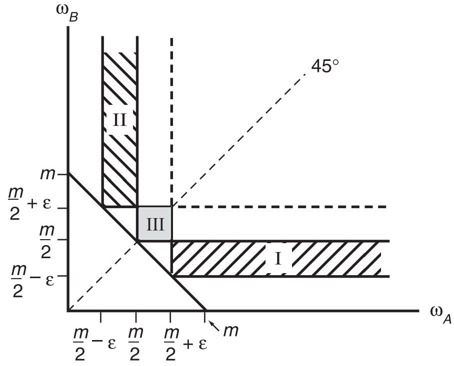
□ **A pair of one-way calling models.** Thus far, we have assumed that only A can initiate a message. We now consider a situation similar to the one examined by earlier authors: either party can initiate a message but does so without regard for the possibility that the other party might also initiate a message. The analysis goes beyond that of earlier authors by measuring welfare in a way that recognizes the substitutability of a message sent in one direction for a message sent in the other. Assuming naive behavior, party i initiates a message whenever $\omega_i \geq p$. Because the

[22] Let p_0 denote the limit. Suppose counterfactually that $p_0 \neq p^*(0)$. By the definition of $p^*(0)$, $S_A(p^*(0)) \times S_B(m - p^*(0)) > S_A(p_0)S_B(m - p_0)$. By the continuity of $S_A(\cdot)$ and $S_B(\cdot)$ and the fact that $p^*(\bar{\pi})$ tends to p_0, it must be the case that $S_A(p^*(0) + \bar{\pi}/2)S_B(m - p^*(0) + \bar{\pi}/2)\bar{\pi} > S_A(p^*(\bar{\pi}))S_B(m - p^*(\bar{\pi}))\bar{\pi}$ for all sufficiently small $\bar{\pi}$. This inequality contradicts the definition of $p^*(\bar{\pi})$.

[23] Note that the sender would never choose p_2 if $r_2 \geq r_1$.

FIGURE 4

MESSAGE EXCHANGE WHEN EITHER PARTY CAN INITIATE A CALL



parties are behaving nonstrategically, there is no great interest in who is the sender and who is the receiver when both parties desire to initiate a message—it is simply assumed that the message is exchanged in that case.

When either party can originate a message, unequal send and receive prices create options that are similar to those created by offering a pricing menu under one-way calling. With unequal send and receive prices, a user can adopt unequal thresholds for initiating and accepting messages. Intuitively, such options are valuable because they increase the likelihood that a sender with a highly valuable message will be able to communicate it successfully.

Figure 4 illustrates the effects of moving from equal send and receive prices to unequal ones, $m/2+\varepsilon$ and $m/2-\varepsilon$. When the prices are equal, calls are completed when (ω_A, ω_B) lie above the L whose corner is $(m/2, m/2)$. When the prices are unequal, calls are completed when (ω_A, ω_B) lies in the *union* of the regions above two Ls, one with corner $(m/2-\varepsilon, m/2+\varepsilon)$, the other with corner $(m/2+\varepsilon, m/2-\varepsilon)$. Hence, comparing unequal to equal prices, one gains areas I and II but loses area III. As drawn, areas I and II dwarf area III; so, provided the joint distribution over (ω_A, ω_B) does not put too much weight near the 45° line, expected welfare is increased by moving to unequal send and receive prices. Formally, we have the following.

Proposition 6. Suppose the joint-density function for the parties' expected values of message exchange, ω_A and ω_B, is continuous and strictly positive at $(m/2, m/2)$. Then under a pair of one-way calling models (i.e., two-way calling with nonstrategic users), the socially optimal breakeven send and receive prices are not equal.

Proof. See the Appendix.

Observe that we have not said which of $m/2-\varepsilon$ and $m/2+\varepsilon$ is the send price and which is the receive price. For the analysis in this section, which one is which is irrelevant even when the joint density for expected message values is asymmetric. If $m/2-\varepsilon$ is the send price, then the upper L region in Figure 4 corresponds to messages initiated by A and accepted by B. If $m/2+\varepsilon$ is the send price, then that upper L region corresponds to messages initiated by B and accepted by A. The sets of messages exchanged are identical. This logic establishes the following.

Proposition 7. Under a pair of one-way calling models, if prices $p = k$ and $r = m - k$ are socially optimal for some constant k, then prices $p = m - k$ and $r = k$ are also socially optimal.

A remaining question is the value of k in Proposition 7. At the level of abstraction in this subsection, a definite value cannot be established. However, if the joint density over (ω_A, ω_B) pairs puts sufficiently little weight on pairs for which $\max\{\omega_A, \omega_B\} > m$, then $k < m$.[24] Proposition

[24] Differentiating the expected gain from moving from equal prices to prices $m/2+\varepsilon$ and $m/2-\varepsilon$, expression (A9) in the Appendix, and evaluating it at $\varepsilon = m/2$ reveals that if sufficiently little weight is put on pairs for which $\max\{\omega_A, \omega_B\} > m$, welfare is improved by reducing ε from $m/2$.

6 tells us $k > m/2$. Hence, a k in $(m/2, m)$ is optimal if the joint distribution does not put too much weight on large expected values of message exchange. Graphically, recall Figure 4: as ε gets large, area III, the loss, becomes large relative to areas I and II, the gains, from increasing ε.

4. Two-way calling with strategic users

■ As in the preceding subsection, suppose that either A or B can initiate communication, but now suppose they behave strategically. That is, each player faces the choice of either sending a message or waiting for the other party to send one. If the send price exceeds the receive price, then waiting is preferable conditional on the other party's initiating soon. The danger, however, is that both wait and valuable communication is significantly delayed or fails to occur.

To model this situation, we consider the following highly stylized waiting game. The values of ω_A and ω_B are drawn from the same bounded support, normalized to the unit interval, [0, 1]. The message-exchange cost, m, is scaled accordingly, so that $m \in (0, 1)$. Further assume it is common knowledge that the expected message values are drawn independently with distribution functions $\Psi_A(\cdot)$ and $\Psi_B(\cdot)$, each of which is differentiable with positive density, $\psi_A(\cdot)$ and $\psi_B(\cdot)$, respectively, everywhere on (0, 1). The value of communicating decays over time. Specifically, if communication is initiated and accepted at t, then it is worth $\omega_A e^{-\delta t}$ and $\omega_B e^{-\delta t}$ to parties A and B, respectively, where δ is the common decay rate. As before, at most one message is exchanged.

Our analysis proceeds by making welfare comparisons of alternative breakeven pricing regimes. That is, we take prices as exogenously given and characterize the consumers' equilibrium strategies.[25] The solution concept is Bayesian-Nash equilibrium.

□ **Equilibria under alternative pricing regimes.** One regime of interest is equal cost sharing (i.e., $p = m/2 = r$). In this regime, there is no advantage to delay, thus sending is immediate if at least one party's value of communicating exceeds $m/2$. That message is accepted if the receiver's value exceeds $m/2$.[26] Hence, message exchange is immediate if and only if $\min\{\omega_A, \omega_B\} \geq m/2$.

Now, and for the rest of this subsection, we consider a sender-pays-more regime, $p > m/2 > r$.[27] We assume that, if both parties send a message at the same moment, each party has an equal chance of being billed as the sender, with the other party being billed as the receiver (i.e., the expected payment is $(p + r)/2$). Because delay is costly and $p > r$, it clearly is a dominant strategy for party i to accept a message sent at t if and only if $\omega_i e^{-\delta t} \geq r$, that is, whenever $t \leq (\ln \omega_i - \ln r)/\delta$. Hence, if party i sends a message at time t and it is rejected, party i believes it will also be rejected if sent later.

Observe that, if it is rational for party i to send a message at t, then $t \leq (\ln \omega_i - \ln p)/\delta$. Because $r < p$, if party i is willing to send at t, she is also willing to accept at t. Let $T \equiv -\ln p/\delta$ be the last time at which initiating communication could be rational.

For each possible value of ω_i, party i's sending strategy is a distribution over times at which i plans to send a message if message exchange has not already occurred. Let $\Sigma_i(\cdot, \omega)$ denote the distribution of send times conditional on having expected message value ω. Because, in any equilibrium, j must play a best response and, hence, be able to construct one by forming expectations about the times at which he might receive a message, we limit attention to strategies for i such that $\Sigma_i(t, \cdot)$ is a measurable function.

Lemma 1 (messages exchanged with positive probability). If $p < 1$, then, in any equilibrium, each party expects to be sent, with positive probability, a message that he or she would accept.

[25] In the next subsection, we derive properties that an equilibrium must possess *if* it exists. In the subsequent subsection, we provide an example of an equilibrium.

[26] We do not worry about who is the sender and who is the receiver when both parties' values exceed $m/2$ and, thus, both are willing to send immediately. It is assumed that the message is exchanged (i.e., there is no "busy signal") and the cost, m, is borne evenly.

[27] A receiver-pays-more regime (i.e., $r > m/2 > p$) is symmetric in a sense to be discussed at the end of this subsection.

Proof. Define $\mathcal{M}_i(t) \equiv \{\omega \mid \Sigma_i(t,\omega) > 0\}$, the set of types for party i who may initiate communication at time t or earlier. We first show $\Psi_A(\mathcal{M}_A(T)) > 0$ and $\Psi_B(\mathcal{M}_B(T)) > 0$. Suppose, to the contrary, that $\Psi_j(\mathcal{M}_j(T)) = 0$ for at least one of $j = A$ or B. Because i expects never to receive messages, her best response is to send a message at time 0 if $\omega_i > p$ and never send a message if $\omega_i < p$. Consider a $t_\varepsilon > 0$ such that $re^{\delta t_\varepsilon} < p$ and $pe^{\delta t_\varepsilon} < 1$. By hypothesis, there is a positive measure of values of ω_j who never send a message in the candidate equilibrium but for whom $\omega_j e^{-\delta t_\varepsilon} > p$. Suppose these types deviate to a strategy of sending at t_ε. The expected gain from deviating is $(\omega_j e^{-\delta t_\varepsilon} - p)(\Psi_i(p) - \Psi_i(re^{\delta t_\varepsilon})) > 0$, a contradiction. A similar analysis would apply if i expected never to be sent a message that she would accept: i's best response would have to be to send only at $t = 0$, but then, for a positive measure of ω_j, j would be in a situation in which he wished to deviate to sending messages at some t_ε that some types of i would accept. This last contradiction establishes the lemma. *Q.E.D.*

Define $F_i(t) = \int_0^1 \Sigma_i(t,\omega)\psi_i(\omega)d\omega$. $F_i(\cdot)$ is the *ex ante* distribution over the times party i intends to initiate communication. As such, it is monotone and right continuous (i.e., $\lim_{t \downarrow t_0} F(t) = F(t_0)$). In fact, as we show, it is continuous, but it may have an "atom" at zero.

Lemma 2 (continuity). In any equilibrium, for $i = A, B$, $F_i(\cdot)$ is continuous on $(0, T)$. Moreover, if $F_i(0) > 0$ for $i = A$ or B, then $F_j(0) = 0$ for $j \neq i$.

Proof. See the Appendix.

Intuitively, if j expected to be sent a message precisely at t_0 with strictly positive probability, j would do better not to send at the same time but to delay slightly, because that would strictly increase the odds of being a receiver without incurring substantial message decay. This same logic also holds for times shortly before t_0 (if $t_0 > 0$), so that j never sends in an interval $(t_0 - \eta, t_0]$ for some $\eta > 0$. But then it cannot be a best response for i to send at t_0—if she accelerates her sending time toward $t_0 - \eta$, then there is no cost in terms of missing an opportunity to be a receiver (j isn't sending in the interval $(t_0 - \eta, t_0]$), but there is a benefit because (i) at least as many j will accept at a time earlier than t_0 as at t_0, and (ii) the value of the message has decayed by less. Hence, the only possible atom is at $t = 0$.

The same intuition that a party wants to accelerate her sending times when she doesn't expect to receive a message just prior to her planned sending time also explains why there are no "gaps" between times at which, *ex ante*, a party might possibly send. If i had a gap, then j wouldn't send during it—j would accelerate. But if j won't send during that gap, then i shouldn't plan to send at the "end" of the gap; she should also accelerate. Formally, we have the following.

Lemma 3 (no gaps). Consider $i = A, B$ and $t_2 > t_1 > t_0$. In any equilibrium, if $F_i(\cdot)$ is strictly increasing on (t_1, t_2) and the probability that j accepts at t_2 is positive, then $F_i(t_1) > F_i(t_0)$.

Proof. See the Appendix.

Because $F_i(\cdot)$ is bounded and monotone on $(0, T]$, it is differentiable, with density $f_i(\cdot)$, almost everywhere. Observe that $f_i(\cdot)$ is Riemann integrable on the interval $(0, T]$ or any subinterval. As before, it is useful to work with the survival function $Q_i(t) \equiv 1 - F_i(t)$.

Party i's equilibrium best-response send time conditional on having type (expected message value) ω_i must be a solution to

$$\max_{t \in \mathcal{T}} Q_j(t) \times \Pr\{j \text{ would accept at } t \mid j \text{ hasn't sent by } t\} \times (\omega_i e^{-\delta t} - p) + \int_0^t (\omega_i e^{-\delta z} - r) f_j(z) dz, \quad (7)$$

where $\mathcal{T} = [0, T]$ if $Q_j(0) = 1$ and, by Lemma 2, $\mathcal{T} = (0, T]$ if $Q_j(0) < 1$. Observe that program (7) treats the possibility of the two parties' calling simultaneously as a zero-probability event, which it is in light of Lemma 2: over $\mathcal{T}$, $F_j(\cdot)$ and, hence, $Q_j(\cdot)$ are continuous. We let $t_i^*(\omega_i)$ denote a solution of (7).[28] We let $t = \infty$ denote a decision never to send a message.

We can now establish that the more valuable exchanging a message is to him or her, the (weakly) sooner a party sends a message.

Lemma 4 (weak monotonicity). Consider party i and two expected values of message exchange, $\omega_1 < \omega_2$. In any equilibrium, either (i) $t_i^*(\omega_1) \geq t_i^*(\omega_2)$, or (ii) the probability that j accepts the message at time $t_i^*(\omega_1)$ equals zero.

Proof. See the Appendix.

Combining results, we see that no meaningful mixing can occur.

Lemma 5 (no mixing). In any equilibrium, the parties play pure strategies with respect to sending times that have positive probabilities of being accepted.

Proof. Suppose i mixes over t_1 and t_2 (and possibly other times as well) when her type is ω_0, where $t_1 < t_2$ and t_2 is accepted with positive probability. By Lemma 4, if $\omega_1 > \omega_0$, then $t_i^*(\omega_1) \leq t_1 < t_2$; that is, ω_1 prefers not to send a message at t_2. Similarly, if $\omega_2 < \omega_0$, then ω_2 prefers not to send a message at t_1. But then party i almost surely does not send a message in the interval (t_1, t_2), a contradiction of Lemma 3. *Q.E.D*

Last, we observe that with positive probability both parties delay sending messages.

Lemma 6 (delay happens). In any equilibrium, there exist $\hat{\omega}_i$, for $i = A, B$, such that $t_i^*(\hat{\omega}_i) > 0$ and a message sent by i at $t_i^*(\hat{\omega}_i)$ will be accepted by $j \neq i$ with positive probability.

Proof. Given that with positive probability both parties send messages that have positive probabilities of acceptance (Lemma 1) and they cannot both send these messages only at $t = 0$ (Lemma 2), the only case to consider is an equilibrium in which one party, say i, sends messages with a positive probability of acceptance only at $t = 0$ and the other party, j, does not. By Lemma 2, j sends *no* messages at time 0. Fix an $\omega_j > p$ and let $t' > 0$ be the equilibrium time at which such a j sends a message (by Lemma 1, such a time exists for some ω_j). Consider an alternative time $t'' \in (0, t')$. By assumption (t'', t') is a gap in which i never sends. By our now-familiar argument that a player should accelerate his sending time if he would otherwise face a gap of never receiving, ω_j should deviate to t''. Therefore, neither party sends at $t = 0$ all his or her messages that have positive probabilities of acceptance. *Q.E.D*

To summarize, the results above imply the following.

Proposition 8. Any equilibrium of the two-way calling game has the following properties:

(i) There is message exchange at time 0 for some expected message values, but there is delay for others.
(ii) The parties play pure strategies such that, except possibly at time 0, the time at which a party plans to send a message if not sent a message earlier is a strictly decreasing function of her expected value of exchanging a message.
(iii) For each party, i, there is a cutoff value, $\omega_i^\ell > p$, such that party i (a) sends a message with a positive probability of acceptance if $\omega_i > \omega_i^\ell$, and (b) never sends a message with a positive probability of acceptance if $\omega_i < \omega_i^\ell$.

In words, finding (iii) in the proposition states that there exist expected message values such that a party values exchanging a message by more than the cost of sending one, but never sends a message because she waits to receive one at a lower price and the depreciated message value eventually falls below the send price.[29]

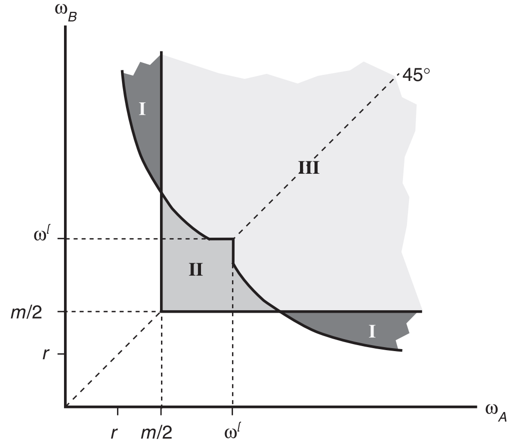
---

[28] A solution to the program may fail to exist, in which case $Q_j(\cdot)$ cannot have been generated by an equilibrium strategy.

[29] Part (iii) follows from Lemmas 1 and 4.

FIGURE 5

MESSAGES EXCHANGED WHEN $p > m/2 > r$ AS OPPOSED TO WHEN $p = m/2 = r$



The findings of the analysis above are illustrated in Figure 5.[30] Messages exchanged if $p > r$ are those in areas I and III. Messages exchanged if $p = r$ are those in areas II and III. The lower envelope of areas I and III takes a largely curved shape—approaching the straight lines $\omega = r$—because messages initiated by low-ω senders occur later and, thus, are accepted only by receivers for whom $\omega > re^{\delta t}$.[31] Because anyone who would send at a given time will also accept, this envelope line hits the $\omega = \omega^\ell$ line before the 45° line, creating a "corner" near the 45° line. Area II must exist because $\omega^\ell > p > m/2$. A welfare comparison of the pricing regimes with $p > m/2 > r$ and $p = m/2 = r$ must balance the mass in area I with that in area II. It must also account for the welfare loss due to delay under the $p > r$ regime.

The receiver-pays-more case, $p < r$, is symmetric in the following sense. Instead of delaying sending, the parties delay receiving. Likewise, instead of being willing to receive until the value of message exchange falls below r, they are willing to send continually until the value of message exchange falls below p. That is, suppose in a sender-pays-more regime with prices $p_1 > r_1$, the equilibrium sending times are $t_i^*(\cdot)$ and the equilibrium intervals for accepting messages are $[0, T^*(\omega_i)]$. Then in a receiver-pays-more regime with prices $r_2 = p_1$ and $p_2 = r_1$, there is an equilibrium in which each party i sends a message continually over the interval $[0, T^*(\omega_i)]$ and first accepts a message at $t_i^*(\omega_i)$. Observe that our assumption that message costs are incurred only if a message is exchanged is critical to this result.

□ **Welfare analysis for power-function distributions.** At the level of generality of the previous subsection, we are unable to compare welfare levels between pricing regimes. In the present subsection, we therefore impose more structure on the problem by assuming $\Psi_A(\omega) = \Psi_B(\omega) = \omega^\gamma$, where $\gamma > 0$. Even with this additional structure, the fact that the receive decision is time-dependent when $r > 0$ makes it difficult to obtain closed-form solutions for expected equilibrium payoffs. For tractability, we compare an equal-sharing regime with a sender-pays regime (i.e., $p = m$ and $r = 0$).

Consider the sender-pays regime. Recall that $t_i^*(\omega)$ is player i's equilibrium planned time to send a message if player j has not sent one earlier expressed as a function of player i's expected value of message exchange. Assuming that the equilibrium value of $Q_j(\cdot)$ is differentiable, the

[30] Figure 5 is drawn assuming that $\Psi_A(\cdot) \equiv \Psi_B(\cdot)$ and the equilibrium is symmetric.

[31] If $r = 0$, then the lower envelope coincides with the appropriate axis for $\omega \geq \omega^\ell$. That is, if $r = 0$, then each part of area I is a rectangle between zero and $m/2$ and beyond ω^ℓ. Although setting $r = 0$ increases area I and shrinks area II relative to the graph shown, it eliminates neither the second area nor the further welfare loss due to delay.

first-order condition for the best-response program, expression (7), is[32]

$$-\dot{Q}_j(t_i^*(\omega_i))m - Q_j(t_i^*(\omega_i))\delta\omega_i e^{-\delta t_i^*(\omega_i)} = 0. \tag{8}$$

One can rewrite (8) as

$$-\frac{\dot{Q}_j}{\omega_i Q_j} = \frac{\delta}{m}e^{-\delta t}. \tag{9}$$

Suppose the equilibrium is symmetric. Dropping the subscripts for convenience, and using the fact that $t^*(\cdot)$ is nonincreasing, we have

$$\begin{aligned} Q(t^*(\omega_0)) &= \Pr\{t^*(\omega) \geq t^*(\omega_0)\} \\ &= \Pr\{\omega \leq \omega_0\} \\ &\equiv \Psi(\omega_0). \end{aligned}$$

Using this result in (9), the differential equation becomes

$$-\frac{\dot{Q}}{\Psi^{-1}(Q)Q} = \frac{\delta}{m}e^{-\delta t}. \tag{10}$$

For a power-function distribution, (10) becomes

$$-\frac{\dot{Q}}{Q^{(\gamma+1)/\gamma}} = \frac{\delta}{m}e^{-\delta t},$$

the solution to which is

$$\frac{\gamma}{Q^{1/\gamma}} = K - \frac{1}{m}e^{-\delta t^*(\omega)}, \tag{11}$$

K a constant.

Using the fact that $t^*(1) = 0$ (by Proposition 8), equation (11) implies that $K = (1 + m\gamma)/m$. Solving (11) for $t^*(\cdot)$,

$$t^*(\omega) = \frac{1}{\delta}\ln\left(\frac{\omega}{\omega(1 + m\gamma) - m\gamma}\right). \tag{12}$$

A consumer with value ω_0 will never send a message after time t_0 if $\omega_0 e^{-\delta t_0} < m$. Consequently, the last type to send a message is ω^ℓ, where ω^ℓ solves

$$\omega e^{-\delta t^*(\omega)} = m. \tag{13}$$

Some algebra on (12) and (13) yields

$$\omega^\ell = \frac{m\gamma + m}{m\gamma + 1}. \tag{14}$$

Recall that $m < 1$. Hence, (14) confirms $\omega^\ell > p = m$.

For reference, observe that the last possible message time under sender pays is

$$t^\ell \equiv t^*(\omega^\ell) = \frac{1}{\delta}\ln\left(\frac{1 + \gamma}{m\gamma + 1}\right).$$

[32] When $r = 0$, the conditional probability in (7) is one: all types are always willing to accept a message.

We now compare expected welfare under the two pricing regimes. When $p = m/2 = r$, expected welfare is

$$W_E = \int_{m/2}^{1} \int_{m/2}^{1} (\omega_A + \omega_B - m)\psi(\omega_A)\psi(\omega_B)d\omega_A d\omega_B. \tag{15}$$

Under the sender-pays regime, expected welfare is[33]

$$\begin{aligned} W_S &= 2 \times \int_{\omega^\ell}^{1} \left(\int_0^{\omega} [e^{-\delta t^*(\omega)}(\omega + \mu) - m]\psi(\mu)d\mu \right) \psi(\omega)d\omega \\ &= 2 \times \int_{\frac{m(1+\gamma)}{m\gamma+1}}^{1} \left(\int_0^{\omega} \left[\frac{\omega(1+m\gamma) - m\gamma}{\omega}(\omega + \mu) - m \right] \psi(\mu)d\mu \right) \psi(\omega)d\omega. \end{aligned} \tag{16}$$

Recalling that $\psi(\omega) = \gamma\omega^{\gamma-1}$, computations reveal

$$W_E - W_S = \frac{m}{4^\gamma(1+\gamma)} \left[2^\gamma m^\gamma (2+\gamma) - m^{2\gamma} - 2^{1+\gamma} m^{\gamma-1}\gamma + 4^\gamma \left(\gamma - \left[\frac{m(1+\gamma)}{1+m\gamma} \right]^{2\gamma} \right) \right]. \tag{17}$$

Further calculations (sketched in the Appendix) demonstrate that this expression is positive for all $\gamma \geq 1$. That is, expected welfare is greater when the message cost is evenly divided than when the sender alone pays.

Proposition 9. Suppose that the expected message values are independently and identically distributed with distribution function $\Psi(\omega) = \omega^\gamma$, where $\gamma \geq 1$. Then equilibrium under the equal-division regime yields greater expected surplus than does the symmetric equilibrium of the sender-pays regime.

Proof. See the Appendix.

The intuition underlying Proposition 9 can be seen by considering the case of a uniform distribution ($\gamma = 1$). For this case, the proposition is the consequence of two factors. First, for relatively large costs (i.e., m near one), the likelihood that either player ever initiates a call becomes too small under a sender-pays regime. Second, even when costs are not that large, there is still the problem of strategic delay. To see the importance of delay, fix $t^*(\omega) \equiv 0$. Then the notional value of expected welfare is

$$\begin{aligned} 2 \times \int_{\omega^\ell}^{1} \left(\int_0^{\omega} [e^{-\delta \times 0}(\omega + u) - m]du \right) d\omega &= 2 \times \int_{\frac{2m}{m+1}}^{1} \left(\int_0^{\omega} (\omega + u - m)du \right) d\omega \\ &= \frac{1 + 2m - 6m^3 + 3m^4}{(1+m)^3}. \end{aligned} \tag{18}$$

It can be shown that (18) exceeds (15) if and only if $m < .513$. In words, except for delay, sender pays would yield greater welfare than would equal division when the cost of a message is low (roughly less than the expected value of a message to either party). But there is strategic delay, which is why sender pays is dominated by equal division for all m when $\gamma = 1$.[34]

[33] The double integral corresponds to the area beneath the 45° line in Figure 5. We calculate the expected value when A sends a message first and then double it to include when B sends first.

[34] Kim, Bae, and Won (2002) consider a game that can be interpreted as one of two-way message exhange. It is similar to the one we analyze here, except they consider only uniformly distributed values and, critically, assume that no messages can be exchanged after time 0 (i.e., the calling opportunity disappears). Without scope for delay, they find that sender pays dominates equal division when $m \leq .58$. The intuition is that the strategic alternative of not calling is worse

In contrast to Proposition 9, numerical calculations demonstrate that expected welfare can be greater under sender pays than under equal division for concave power-function distributions (i.e., $\gamma < 1$) and low values of the cost parameter, m. For example, if $\gamma = .9$, then sender pays dominates equal division for $m < .21343$. Equal division, however, continues to dominate sender pays for $m > .21343$. Intuitively, when the parties' values tend to be low, the likelihood that *both* parties will have high values for sending and receiving is also small. In this case, there is a benefit to creating options, as we saw earlier in our analysis of nonstrategic calling (the option here is whether to be an active sender or a passive receiver). When the send price gets too large, however, the probability that either party has a value above that price gets too low—at which point equal send and receive prices are superior.

5. Conclusion

■ In discussions of telecommunications pricing, it often is asserted that the principle of cost-causative pricing implies that it is efficient for the sender to pay the marginal cost of exchanging a message. In the absence of call externalities: the sender can correctly be viewed as the "cost causer"; with constant marginal cost, the efficient price is cost based in the sense that it can be set without reference to demand;[35] and pricing menus have no social value.

The presence of call externalities changes all of these findings. First, one could just as well assert that the receiver causes the costs by accepting the message.[36] Second, efficient prices must internalize the external effects across the two parties to a message exchange. Hence, with or without a network profitability constraint, efficient pricing requires consideration of demand conditions, as well as cost conditions, even when marginal cost is constant. Third, our analysis demonstrates that frequently it is not socially optimal to have one party bear the full marginal costs of exchanging a message. Fourth, pricing menus can be socially valuable.

Our analysis is based on a number of restrictive assumptions. While the lessons drawn from our simple model will remain, further insights into optimal pricing could be gained by exploring at least three important extensions.

One is to allow for multiple senders and receivers. In practice, even when any given message involves just two parties, there are typically large sets of potential senders and receivers. For some technologies, such as e-mail and caller ID, both the sender and receiver know with whom they are communicating. In this case, our analysis goes through unchanged if pair-specific pricing is feasible. But although telecommunications firms often offer special rates for specific pairs (e.g., MCI's old "Friends & Family" plan or the differential rate that MCI charges on calls home when using an MCI calling card), in many instances they do not. Moreover, the network is unlikely to have detailed information about valuation distributions at the pair level. While it is straightforward to extend our model of individual user behavior to the case of multiple senders and receivers, it is difficult to characterize optimal prices fully because of the need to aggregate across user pairs.

The second extension is to consider access pricing and consumption decisions. A natural model would be as follows: (1) the network sets access, send, and receive prices; (2) consumers decide whether to subscribe to the network; (3) each consumer i learns the value of ω_i; and (4) consumers make message-exchange decisions. Interesting questions include whether access pricing should subsidize message pricing or vice versa.

The third extension is to consider the case of multiple networks. The recognition of call externalities can fundamentally affect the analysis of network interconnection. For instance, the possibility of charging the receiver raises interesting questions about whether a network originating a message should pay an interconnection fee to the network terminating the message as payment

when delay is infeasible, so the parties are less likely to play it, and moreover, sender-pays pricing creates something akin to the menu effects that arose under nonstrategic bilateral calling.

[35] If marginal cost is *not* constant, then demand must also be considered in determining the efficient price, which of course is equal to the equilibrium value of marginal cost.

[36] A model that allowed for set-up costs that would be incurred even if a sent message were refused would put the parties in somewhat asymmetric positions but would not alter the fundamental point.

for a wholesale service or should collect a fee from the terminating network in return for providing that network's subscribers a valuable service. Recently, some authors have begun to address issues such as this one (see, for example, DeGraba, 2003; Hermalin and Katz, 2001; Jeon, Laffont, and Tirole, 2004; and Kim and Lim, 2001). However, much of this work has made rather restrictive assumptions about the distributions of signals and message exchange values, and the results of the present article suggest that such restrictions can strongly influence the results obtained.

Appendix

■ Proofs of Propositions 1, 3, 6, and 9, and Lemmas 2–4 follow.

Proof of Proposition 1. The first part of the proposition follows from the discussion of Figure 1 in the text. To prove the second and third parts, fix a candidate pair of prices. If either $p \geq \bar{\omega}_A$ or $r \geq \bar{\omega}_B$, then expected welfare is zero (where, as above, $\bar{\omega}_i$ (alternatively, $\underline{\omega}_i$) denotes the supremum (alternatively, infimum) of the distribution of ω_i). By Assumption 2, this cannot be a welfare-maximizing outcome. If either $p < \underline{\omega}_A$ or $r < \underline{\omega}_B$, then expected welfare is unaffected by increasing p to $\underline{\omega}_A$ or r to $\underline{\omega}_B$. In summary, there is no loss of generality looking only for optimal prices satisfying

$$\underline{\omega}_A \leq p < \bar{\omega}_A \quad \text{and} \quad \underline{\omega}_B \leq r < \bar{\omega}_B.$$

Given prices such that $p + r = m + \pi$, we wish to show that these prices are suboptimal if (i) $\pi \geq 0$ and the network does not face a profitability constraint, or (ii) $\pi > 0$ and the firm faces the profitability constraint $p + r \geq m$. Because $\underline{\omega}_A + \underline{\omega}_B < m$ (Assumption 2), $p + r \geq m$ implies that $p > \underline{\omega}_A$, $r > \underline{\omega}_B$, or both. Assume that $p > \underline{\omega}_A$ (the case in which only $r > \underline{\omega}_B$ is handled analogously). Expected welfare under prices (p, r) is

$$\int_{m+\pi-r}^{\bar{\omega}_A} \left(\int_r^{\bar{\omega}_B} (\omega_A + \omega_B - m)\psi_B(\omega_B) d\omega_B \right) \psi_A(\omega_A) d\omega_A. \tag{A1}$$

Differentiating (A1) with respect to π yields

$$-\psi_A(m + \pi - r) \int_r^{\bar{\omega}_B} \Big(\omega_B + \pi - r\Big)\psi_B(\omega_B) d\omega_B < 0 \qquad \forall \pi \geq 0.$$

Q.E.D.

Lemma A1. If the survival function $S(\cdot)$ is log concave, then $\int_p^\infty S(\omega)d\omega$ is log concave.

Proof. By hypothesis,

$$\frac{S'(p)}{S(p)} \geq \frac{S'(\omega)}{S(\omega)} \tag{A2}$$

for all $\omega \geq p$. By definition, $\int_p^\infty S(\omega)d\omega$ is log concave if

$$\frac{-S(p)}{\int_p^\infty S(\omega)d\omega}$$

is a decreasing function. The derivative of this last expression has the same sign as

$$-S'(p) \int_p^\infty S(\omega)d\omega - S(p)^2.$$

This expression has the same sign as

$$\int_p^\infty S(\omega) \frac{-S'(p)}{S(p)} d\omega - S(p). \tag{A3}$$

By (A2), (A3) is less than or equal to

$$\int_p^\infty S(\omega) \frac{-S'(\omega)}{S(\omega)} d\omega - S(p) = -\int_p^\infty S'(\omega)d\omega - S(p) = 0.$$

Q.E.D.

Lemma A2. Let $\psi(\cdot)$ be weakly log concave, and let (ω_A, ω_B) and (ω'_A, ω'_B) be such that $\omega_A + \omega_B = \omega'_A + \omega'_B = \kappa$, for some constant κ. If $\mid \omega_A - \omega_B \mid < \mid \omega'_A - \omega'_B \mid$, then $\psi(\omega_A)\psi(\omega_B) \geq \psi(\omega'_A)\psi(\omega'_B)$.

Proof. To prove the lemma, it is sufficient to show that

$$\psi(\Delta)\psi(\kappa - \Delta) \tag{A4}$$

is nondecreasing in Δ for $\Delta < \kappa/2$ and nonincreasing in Δ for $\Delta > \kappa/2$. The sign of the derivative of (A4) is the same as the sign of

$$\frac{\psi'(\Delta)}{\psi(\Delta)} - \frac{\psi'(\kappa - \Delta)}{\psi(\kappa - \Delta)}. \tag{A5}$$

Because $\psi(\cdot)$ is weakly log concave, $\psi'(\omega)/\psi(\omega)$ is nonincreasing in ω. Hence, (A5) is nonnegative for $\Delta < \kappa/2$ and nonpositive for $\Delta > \kappa/2$. *Q.E.D.*

Define $Z(\omega) \equiv \Psi_A(\omega) - \Psi_B(\omega)$.

Lemma A3. $Z(\omega) \leq 0$ for all ω. (Observe that this is equivalent to Ψ_A dominating Ψ_B in the sense of first-order stochastic dominance.)

Proof. Let $\bar{\omega} = \sup \operatorname{supp}\{\omega_A\}$ (note that $\bar{\omega}$ could be ∞). Because $\psi_B(\cdot)$ crosses $\psi_A(\cdot)$ once, from above, it follows that $\Psi_B(\bar{\omega}) = 1$; if not, then there would exist $\omega > \bar{\omega}$ such that $\psi_B(\omega) > \psi_A(\omega) = 0$, contradicting single crossing. Because $\psi_A(\omega) < \psi_B(\omega)$ for all $\omega < \hat{\omega}$, $Z(\omega) < 0$ for all $\omega \leq \hat{\omega}$. $Z(\omega)$ is increasing in ω for $\omega \in (\hat{\omega}, \bar{\omega})$ and $Z(\omega) = 0$ for all $\omega \geq \bar{\omega}$. Consequently, by continuity, $Z(\omega) \leq 0$ for all ω. *Q.E.D.*

Proof of Proposition 3. Let $i = A$ and $j = B$ (the proof when these are reversed is a straightforward variant of the one presented here).

The derivative of (1), the expression for expected welfare, has the same sign as

$$\frac{S'_A(p)}{\int_p^\infty S_A(\omega)d\omega} - \frac{S'_B(m-p)}{\int_{m-p}^\infty S_B(\omega)d\omega}. \tag{A6}$$

Because the hazard rates are nondecreasing, (A6) is nonincreasing in p (recall Lemma A1 and the discussion surrounding expression (4)). Hence, if marginal expected welfare is positive for a given p, it is also positive for smaller p.

Observe that $S_B \equiv S_A + Z$. Substituting $S_A + Z$ for S_B in (1) yields the following expression for expected welfare:

$$\begin{aligned} S_A(m-p)\int_p^\infty S_A(\omega)d\omega + S_A(p)\int_{m-p}^\infty S_A(\omega)d\omega \\ + Z(m-p)\int_p^\infty S_A(\omega)d\omega + S_A(p)\int_{m-p}^\infty Z(\omega)d\omega. \end{aligned} \tag{A7}$$

Define $\Lambda(p)$ to equal the first line of (A7). Differentiating (A7) with respect to p yields

$$\Lambda'(p) - Z'(m-p)\int_p^\infty S_A(\omega)d\omega + S'_A(p)\int_{m-p}^\infty Z(\omega)d\omega. \tag{A8}$$

Consider $p \in [m - \hat{\omega}, m/2]$. By part (i) of Proposition 2, the first term of (A8) is nonnegative if $p < m/2$ and zero if $p = m/2$. From $m - p \leq \hat{\omega}$, it follows that $-Z'(m-p) > 0$ and $Z(m-p) < 0$. Thus, the second and third terms are positive. Therefore, (A8) is positive if $p \in [m - \hat{\omega}, m/2]$. Coupled with the analysis of (A6), it follows that (A8) is also positive for $p < m - \hat{\omega}$. Hence, the welfare-maximizing price must exceed $m/2$. That is, A must pay a higher price than B. *Q.E.D.*

Proof of Proposition 6. Consider replacing a single pair of prices, $(m/2, m/2)$, with two pairs, $(m/2 + \varepsilon, m/2 - \varepsilon)$ and $(m/2 - \varepsilon, m/2 + \varepsilon)$. The expected gain in welfare is areas I and II in Figure 4 appropriately weighted by the probabilities of drawing ωs in them. The expected loss in welfare is area III in Figure 4 appropriately weighted. The net expected gain is

$$\begin{aligned} \int_{\frac{m}{2}+\varepsilon}^\infty \int_{\frac{m}{2}-\varepsilon}^{\frac{m}{2}} (\omega_A + \omega_B - m)\psi(\omega_A, \omega_B)d\omega_B d\omega_A + \int_{\frac{m}{2}+\varepsilon}^\infty \int_{\frac{m}{2}-\varepsilon}^{\frac{m}{2}} (\omega_A + \omega_B - m)\psi(\omega_A, \omega_B)d\omega_A d\omega_B \\ - \int_{\frac{m}{2}}^{\frac{m}{2}+\varepsilon} \int_{\frac{m}{2}}^{\frac{m}{2}+\varepsilon} (\omega_A + \omega_B - m)\psi(\omega_A, \omega_B)d\omega_B d\omega_A. \end{aligned} \tag{A9}$$

Differentiating (A9) with respect to ε yields

$$
\begin{aligned}
&-\int_{\frac{m}{2}-\varepsilon}^{\frac{m}{2}} \left(\frac{m}{2}+\varepsilon+\omega_B-m\right)\psi\left(\frac{m}{2}+\varepsilon,\omega_B\right)d\omega_B+\int_{\frac{m}{2}+\varepsilon}^{\infty}\left(\omega_A+\frac{m}{2}-\varepsilon-m\right)\psi\left(\omega_A,\frac{m}{2}-\varepsilon\right)d\omega_A\\
&\quad-\int_{\frac{m}{2}-\varepsilon}^{\frac{m}{2}}\left(\omega_A+\frac{m}{2}+\varepsilon-m\right)\psi\left(\omega_A,\frac{m}{2}+\varepsilon\right)d\omega_A+\int_{\frac{m}{2}+\varepsilon}^{\infty}\left(\frac{m}{2}-\varepsilon+\omega_B-m\right)\psi\left(\frac{m}{2}-\varepsilon,\omega_B\right)d\omega_B\\
&\quad-\int_{\frac{m}{2}}^{\frac{m}{2}+\varepsilon}\left(\frac{m}{2}+\varepsilon+\omega_B-m\right)\psi\left(\frac{m}{2}+\varepsilon,\omega_B\right)d\omega_B-\int_{\frac{m}{2}}^{\frac{m}{2}+\varepsilon}\left(\omega_A+\frac{m}{2}+\varepsilon-m\right)\psi\left(\omega_A,\frac{m}{2}+\varepsilon\right)d\omega_A.
\end{aligned}
$$

Simplifying this last expression and evaluating it at $\varepsilon = 0$ yields

$$\int_{\frac{m}{2}}^{\infty}\left(\omega_A-\frac{m}{2}\right)\psi\left(\omega_A,\frac{m}{2}\right)d\omega_A+\int_{\frac{m}{2}}^{\infty}\left(\omega_B-\frac{m}{2}\right)\psi\left(\frac{m}{2},\omega_B\right)d\omega_B>0.$$

The positive derivative implies that a single, equal-price pair cannot be optimal. *Q.E.D.*

Proof of Lemma 2. Define $\zeta_i(t) = F_i(t) - \lim_{\tau\uparrow t} F(\tau)$. Note that $F_i(\cdot)$ is continuous at t if and only if $\zeta_i(t) = 0$. If $\zeta_i(t) > 0$, then $F_i(\cdot)$ has an atom at t. Suppose $\zeta_i(t_0) > 0$ at some $t_0 \geq 0$. We first show that it cannot be a best response for j to initiate at t_0 and, if $t_0 > 0$, it cannot be a best response for j to initiate in the interval $(t_0 - \eta, t_0]$, for some $\eta > 0$. Suppose that j initiated at t_1 in this interval when $\omega_j = \hat{\omega}$, and consider the deviation to waiting to send a message until $t_0 + \varepsilon$, ε an arbitrarily small but positive number. Conditional on having an expected message value of $\hat{\omega}$, this deviation lowers expected expenditures by at least $\zeta_i(t_0)(p - r)/2$ and reduces benefits by $\hat{\omega}(e^{-\delta t_1} - e^{-\delta t_0})$ for messages j subsequently receives at t_0. The reduction in benefits goes to zero as η goes to zero (i.e., as $t_1 \uparrow t_0$), while the cost savings are bounded away from zero. For a type of player i who is willing to accept, but not send, a message at t_0, j suffers the cost of delay and a potential reduction in the probability that i answers, but both of these effects are vanishingly small for small ε and η. Hence, $F_j(t_0) = F_j(t_0 - \eta)$ for some $\eta > 0$.[37]

Now suppose $\zeta_i(t_0) > 0$ at some $t_0 > 0$. Let ω_i be a type of i who sends a message at t_0. By the previous reasoning, i is never sent a message in the interval $(t_0 - \eta, t_0]$ for some $\eta > 0$. Let $t_2 \in (t_0 - \eta, t_0)$. Suppose ω_i deviates to playing a strategy that puts all the weight she would have put on sending at t_0 to sending at t_2. Because she wouldn't receive any messages in $[t_2, t_0]$, she doesn't risk forgoing an opportunity to be a receiver and pay r rather than p to complete an exchange. The probability that j accepts is at least as great at t_2 as at t_0, and the loss of message value due to decay is strictly less at t_2 than t_0. Hence, she strictly gains by this deviation. Therefore, by contradiction, $\zeta_i(t_0) = 0$ for all $t_0 > 0$ and $F_i(\cdot)$ is continuous on $(0, T]$. *Q.E.D.*

Proof of Lemma 3. Suppose not, so $F_i(t_1) = F_i(t_0)$. We first show it cannot be a best response for j to send a message between t_0 and t_1. To see this, suppose j were to send a message at $\tau \in (t_0, t_1)$ and consider a deviation to $\tau' \in (t_0, \tau)$. Because the probability i sends in (t_0, t_1) is zero, j doesn't forgo an opportunity to be a receiver by deviating. The probability that i accepts at τ' is at least as great as it is at τ, and the loss of message value due to decay is strictly less at τ' than τ. Hence, j would gain by this deviation, a contradiction. Because $F_j(\cdot)$ is continuous (right continuous at zero), it follows that $F_j(t_0) = F_j(t_1)$; i.e., as claimed, party j also does not send in $[t_0, t_1]$. Fix $\varepsilon > 0$, then there is an ω_i who sends a message in $(t_1, t_1 + \varepsilon)$ (otherwise $F_i(\cdot)$ wouldn't be increasing on $[t_1, t_2]$). Suppose that ω_i deviated from sending at $\hat{\tau} \in (t_1, t_1 + \varepsilon)$ to sending at $t_0 + \eta$, $\eta > 0$ but arbitrarily small. Define $A_j^0(t) \equiv \Pr\{j \text{ would accept at } t \mid j \text{ hasn't sent by } t_0\}$. Observe that $A_j^0(\cdot)$ is a nonincreasing function of t and, by assumption, is positive at $t = t_2$. Therefore, $A_j^0(t) > 0\ \forall t \in [t_0, t_2]$. The benefit of deviating from $\hat{\tau}$ to $t_0 + \eta$ is

$$\left(\omega_i e^{-\delta(t_0+\eta)} - p\right) A_j^0(t_0+\eta) - \left(\omega_i e^{-\delta\hat{\tau}} - p\right) A_j^0(\hat{\tau}) > 0 \tag{A10}$$

($F_j(\cdot)$ is continuous, so it is a measure-zero event that j also sends at $\hat{\tau}$ and we may ignore a possible tie). The cost is the forgone chance to be a receiver, but, given that j doesn't send between t_0 and t_1, that is no greater than $(\omega_i e^{-\delta t_1} - r) \times [F_j(\hat{\tau}) - F_j(t_1)]$, which, by the continuity of $F_j(\cdot)$, can be made arbitrarily small by choosing the appropriate ε. Hence, there is an interval $(t_1, t_1 + \varepsilon)$ such that i doesn't call, contradicting the hypothesis that $F_i(\cdot)$ is strictly increasing on (t_1, t_2). *Q.E.D.*

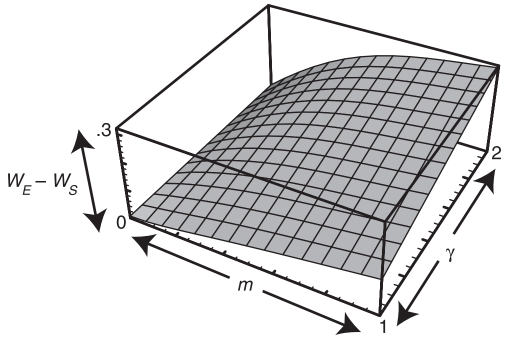
Proof of Lemma 4. Define $A_j(t) \equiv \Pr\{j \text{ would accept at } t \mid j \text{ hasn't sent by } t\}$. And let $R_i(t)$ denote the expected value of $e^{-\delta\tau}$, where τ is the actual exchange time, when i plans to send a message at t if not sent a message before t. Note that the expectation is calculated taking into account both the distribution over ω_j and party j's strategy. Expression (7) can be rewritten as

$$\omega_i R_i(t) - pQ_j(t)A_j(t) - r\Big(1 - Q_j(t)\Big). \tag{A11}$$

[37] By construction $F_j(-\eta) = 0$, so this also establishes the "moreover" part of the lemma.

FIGURE A1

DEMONSTRATION THAT $W_E - W_S > 0$ FOR $m \in (0, 1)$ WHEN $\gamma \in [1, 2]$



Observe that $R_i(\tau) \geq R_i(\tau')$ for all $\tau < \tau'$ because (i) messages sent by j at or before τ are unaffected by i's shift from τ to τ', and (ii) messages that were sent by i and accepted by j at τ are either accepted later (i.e., at $\tau' > \tau$), so decay is greater, or they are never accepted, which also makes expected decay greater. Moreover, if a positive measure of j will accept at τ given j's strategy, then $R_i(\tau)$ is *strictly* decreasing at τ for reason (ii).

The rest of the proof is by contradiction. Suppose, counterfactually, that the best-response sending times satisfy $t_i^*(\omega_1) \equiv t_1 < t_2 \equiv t_i^*(\omega_2)$ and j accepts at t_1 with positive probability. Then $R_i(\cdot)$ is strictly decreasing at t_1, and

$$R_i(t_1) > R_i(t_2). \tag{A12}$$

The optimality of t_1 and t_2 implies, from (A11), that

$$\omega_1[R_i(t_1) - R_i(t_2)] \geq p[Q_j(t_1)A_j(t_1) - Q_j(t_2)A_j(t_2)] - r[Q_j(t_1) - Q_j(t_2)]$$

and

$$p[Q_j(t_1)A_j(t_1) - Q_j(t_2)A_j(t_2)] - r[Q_j(t_1) - Q_j(t_2)] \geq \omega_2[R_i(t_1) - R_i(t_2)].$$

Given that $\omega_1 < \omega_2$, these two inequalities hold only if $R_i(t_1) \leq R_i(t_2)$. But this contradicts inequality (A12). *Q.E.D.*

Proof of Proposition 9 (sketch). The sign of $W_E - W_S$ is equal to the sign of terms in (17) enclosed in the large outside square brackets. Because $\gamma > 1$ and $0 < m < 1$, the expression in these brackets cannot be less than

$$m^\gamma \left(2^\gamma(2+\gamma) - 1\right) + \left\{4^\gamma(\gamma - 1) - 2^{\gamma+1}\gamma\right\}. \tag{A13}$$

The first term of (A13) is positive. The second term (in curly brackets) is nonnegative for $\gamma \geq 2$. This establishes the proposition for $\gamma \geq 2$.

For $1 \leq \gamma < 2$, the "proof" is completed by direct computation. Figure A1 plots expression (17) for $m \in (0, 1)$ and $\gamma \in [1, 2]$. *Q.E.D.*

References

Acton, J.P. and Vogelsang, I. "Telephone Demand over the Atlantic: Evidence from Country-Pair Data." *Journal of Industrial Economics*, Vol. 40 (1992), pp. 305–323.

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

DeGraba, P. "Efficient Intercarrier Compensation of Competing Networks When Customers Share the Value of a Call." *Journal of Economics and Management Strategy*, Vol. 12 (2003), pp. 207–230.

Hahn, J.-H. "Nonlinear Pricing of Telecommunications with Call and Network Externalities." *International Journal of Industrial Organization*, Vol. 21 (2003), pp. 949–967.

Hermalin, B.E. and Katz, M.L. "Network Interconnection with Two-Sided User Benefits." Working Paper, University of California, Berkeley, 2001.

——— and ———. "Retail Telecommunications Pricing in the Presence of External Effects." In G. Madden, ed., *International Handbook on Emerging Telecommunications Networks*, Vol. 2. Northampton, Mass.: Edward Elgar Publishing, 2003.

JEON, D.S., LAFFONT, J.-J., AND TIROLE, J. "On the Receiver-Pays Principle." *RAND Journal of Economics*, Vol. 35 (2004), pp.85–110..

KIM J.-Y. AND LIM, Y. "An Economic Analysis of the Receiver Pay Principle." *Information Economics and Policy*, Vol. 13 (2001), pp. 231–260.

———, BAE, H., AND WON, D. "Dutch Treat versus Oriental Treat." *Journal of Economic Behavior and Organization*, Vol. 48 (2002), pp. 413–422.

LAFFONT, J.-J. AND TIROLE, J. *Competition in Telecommunications*. Cambridge, Mass.: MIT Press, 2000.

LITTLECHILD, S.C. "Two-Part Tariffs and Consumption Externalities." *Bell Journal of Economics*, Vol. 6 (1975), pp. 661–670.

PEČARIĆ, J.E., PROSCHAN, F., AND TONG, Y.L. *Convex Functions, Partial Orderings, and Statistical Applications*. Boston: Academic Press, 1992.

PRÉKOPA, A. "Logarithmic Concave Measures with Applications." *Acta Scientiarum Mathematicarum*, Vol. 32 (1971), pp. 302–316.

RAMSEY, F.P. "A Contribution to the Theory of Taxation." *Economic Journal*, Vol. 37 (1927), pp. 47–61.

ROHLFS, J. "A Theory of Interdependent Demand for a Communications Service." *Bell Journal of Economics and Management Science*, Vol. 5 (1974), pp. 16–37.

SQUIRE, L. "Some Aspects of Optimal Pricing for Telecommunications." *Bell Journal of Economics and Management Science*, Vol. 4 (1973), pp. 515–525.

SRINAGESH, P. AND GONG, J. "A Model of Telecommunications Demand with Call Externality." Working paper, Bellcore, 1996.

TAYLOR, L.D. *Telecommunications Demand in Theory and Practice*. Boston: Kluwer Academic Publishers, 1994.

WILLIG, R.D. "The Theory of Network Access Pricing." In H.M. Trebing, ed., *Issues in Public Utility Regulation*. East Lansing, Mich.: Michigan State University, 1979.