RAND Journal of Economics
Vol. 35, No. 1, Spring 2004
pp. 85–110

On the "receiver-pays" principle

Doh-Shin Jeon*

Jean-Jacques Laffont**

and

Jean Tirole***

This article extends the theory of network competition by allowing receivers to derive a surplus from receiving calls and to affect the volume of communications by hanging up. We investigate how receiver charges affect internalization of the call externality. When the receiver charge and the termination charge are both regulated, there exists an efficient equilibrium. When reception charges are market determined, each network finds it optimal to set the prices for calling and reception at its off-net costs. The symmetric equilibrium is efficient for a proper choice of termination charge. Last, network-based price discrimination creates strong incentives for connectivity breakdowns.

1. Introduction

■ **Motivation.** The deregulation of telecommunications has led to new forms of competition for retail customers through sophisticated and discriminatory pricing. The literature on this new competitive environment[1] has neglected the facts that subscribers care about the number of calls they receive (call externality), that networks may charge their subscribers when they receive calls (receiver-pays principle), and that subscribers can affect volume by hanging up (receiver sovereignty).

Reception charges play an increasingly important role in the case of mobile telephony. The receiver-pays principle is, for example, applied to mobile phone reception in the United States, Canada, and Hong Kong, as well as for international roaming on GSM mobile networks or domestic roaming in the United States.[2] Reception charges similarly play a key role in the new

* Universitat Pompeu Fabra and CREA; doh-shin.jeon@upf.edu.

** University of Toulouse (IDEI) and University of Southern California; laffont@cict.fr.

*** University of Toulouse (IDEI), CERAS, and MIT; tirole@cict.fr.

We thank both the Korea Institute for Industrial Economics and Trade, and France Telecom for financial support, and Jean-Pierre Florens, Joseph Harrington, Patrick Rey, Yossi Spiegel, Nicolas Vayatis, and two anonymous referees for helpful comments. Doh-Shin Jeon gratefully acknowledges financial support from DGES under BEC 2000-1026.

[1] See Armstrong (1998, 2002), Carter and Wright (1999, 2000), Cherdron (2000), Dessein (1999a, 2003), Gans and King (2000, 2001), Hahn (2000), Laffont, Rey, and Tirole (1998a, 1998b), and Wright (2002). See Armstrong (2002) for an excellent overview of the literature.

[2] For example, a typical tariff (Airtouch for Georgia on September 19, 2001) involves a per-minute rate of 40 cents for both outgoing and incoming calls (after the monthly allowance of 125 minutes to which the monthly access charge of $25 entitles). The domestic roaming charge is 65 cents per minute.

Copyright © 2004, RAND.

Internet economy, as both sides of the markets (e.g., dial-up customers and websites) are charged for the capacity and usage of their connection with internet service providers or backbones.[3]

The relevance of consumer sovereignty depends on the level of reception charges. For the moment, receivers on fixed-link networks are not charged for receiving calls.[4] Their primary incentive to abbreviate a conversation is the opportunity cost of their time when they receive the call.[5] We would thus expect the impact of receiver sovereignty to be minor. By contrast, mobile phone subscribers often give their number to a restricted set of people and keep conversations short; dial-up customers in countries (such as France) that have per-minute reception charges substantially limit the length of their connection to the Internet.

The purpose of this article is to extend our understanding of network competition to environments with call externalities and receiver sovereignty in which firms can charge customers for receiving calls. Enriching the existing analysis to account for the existence of receiver surplus and sovereignty serves more than a descriptive purpose, though. On the positive side, reception charges and receiver surplus both alter the operators' competitive strategies. The (outbound) call charge exerts an externality on the receiver, whose level depends on the latter's reception charge. Thus, the pricing of the call service to a subscriber affects not only the welfare of this subscriber, but also that of other subscribers on all networks, and similarly for the choice of the reception charge, if the latter has any impact on the determination of volume. We investigate networks' pricing strategies in this environment. On the normative side, the joint determination of communications services by four parties (caller, receiver, and their operators) raises the question of whether proper incentives are in place for the maximization of joint surplus. With complex externalities at work, we would like to know whether reception charges and/or termination charges ought to be regulated or are best left to market forces, and whether network-based price discrimination should be allowed.

□ **Overview of the analysis.** To provide a roadmap for our analysis, it is convenient to introduce some notation. Call externalities can be formalized by assuming that the caller and receiver obtain concave gross surpluses $u(q)$ and $\tilde{u}(q)$ from a representative call of length q. For specific results, we will for purely technical reasons find it convenient to assume that these surpluses are proportional: $\tilde{u}(q) = \beta u(q)$; more generally, we can let β denote the ratio of marginal surpluses. Note that these surpluses are deterministic. We later make them state-contingent by adding some noise to marginal utilities.

Because "it takes two to tango," we make explicit how the caller and the receiver determine the volume. This article considers the polar case in which the two parties act noncooperatively; volume is then determined by the minimum of the two demands. The extension to a more cooperative determination of volume would be worth pursuing, as we discuss in the conclusion; in particular, parties who are involved in a long-term relationship may come closer to the maximization of their joint surplus.[6]

Two symmetrically differentiated networks compete in nonlinear prices for subscribers. A network i subscriber with outgoing volume q and incoming volume $\tilde{q}$ is charged (without loss of generality in our model)

$$T_i(q, \tilde{q}) = F_i + p_i q + r_i \tilde{q},$$

where F_i is the monthly subscriber charge, p_i is the per-unit (outbound) call price, and r_i is the per-unit reception charge. We let c_0 denote a network's marginal cost of terminating calls, a the

[3] Websites are charged for sending material through both capacity and usage charges (which are equivalent if the volume of outgoing traffic is perfectly foreseen, but not otherwise). Thus, there is a strong component of variable pricing for outgoing traffic in the websites' connection expenses.

[4] With the exceptions of 800 numbers and collect calls.

[5] Politeness also plays a role. This social norm could be viewed, in the context of our analysis, as forcing the receiver to partly internalize the caller's surplus. Note also that, like other social norms, the concept of politeness evolves with the environment. The receiver's abbreviating phone calls may no longer be perceived as aggressive if the caller knows that the receiver has to pay for the call.

[6] As discussed in Littlechild (1977).

© RAND 2004.

reciprocal access charge paid by the originating network to the terminating network, and $c > c_0$ the total (origination plus termination) marginal cost of communications. Last, we assume that the market is "covered." That is, all consumers subscribe to a network. Let α_1 and α_2 denote the two networks' market shares ($\alpha_1 + \alpha_2 = 1$). We assume that α_1 and α_2 are determined by the difference $w_1 - w_2$ in the net surpluses w_i (defined below) offered by the two networks.[7]

Benchmark: social optimum or monopoly network.[8] Let us begin with the nonstrategic context of a benevolent social planner or a monopoly network (the two differ in the monthly fee charged to subscribers, but not in the usage prices). Both aim at creating the highest *total* surplus from a call (net of cost) so as to maximize the average subscriber utility (and raise the monthly fee in the case of an unregulated monopolist). To achieve this, the network can set a calling price p^* that induces the caller to internalize the receiver's surplus, and any reception charge r low enough so that the receiver is not tempted to hang up. For the purposes of this roadmap, suppose that the caller's utility in making a call of length q is $u(q)$, whereas the receiver's is $\beta u(q)$ for receiving a call of that length. If the marginal per-minute cost is c, a call should last q such that $(1 + \beta)\, u'(q) = c$. However, if the caller chooses the length and pays p per minute, the length will satisfy $u'(q) = p$. Thus, length optimality requires that $p^* = c/\,(1 + \beta)$.

Now suppose that the receiver also faces a per-unit charge of r for receiving a call. The receiver will only accept length q if $\beta u'(q) \geq r$. The optimum obtains if $r \leq r^* = c\beta/\,(1 + \beta)$. More generally, because the completion of a call requires both a call to be made and not hung up, an optimal calling pattern will occur so long as $p \leq p^*$ or $r \leq r^*$ with an equality for at least one price. This is indeed a critical feature of this type of network situation. Either the caller's "calling constraint" or the receiver's "receiving constraint" will bind.

Part 1: regulated or contractually determined reception charges. If calls can be made to an alternative network, new issues arise. Suppose initially that the reception charges (r_1, r_2) are regulated or else contractually agreed upon by the two operators before they wage competition in monthly fees and calling charges.

We first conduct the following thought experiment: Suppose the caller always determines the volume of communication. That is, either the receiver is not sovereign (i.e., is not allowed to hang up) or the reception charge is sufficiently low (relative to the receiver's call externality) that the receiver does not find it advantageous to hang up. The inelastic demand for call reception implies that a network's reception charge has no impact on call volume. And so, from the operator-subscriber pair's viewpoint, only the "effective fixed fee" $\{F_i + r_i\tilde{q}\}$ matters, not its composition.

When the reception charge is regulated, networks compete solely through their monthly subscriber charge (F_i) and their calling charge (p_i). They choose tariffs (F_i, p_i) simultaneously. Calling charge p_i determines the calling volume of network i's customers. The fixed fee F_i has no impact on this volume, but it determines network i's market share α_i for a given competitive offer (F_j, p_j).

Up to a constant, the net surplus of a network i customer is

$$w_i = \max_q \{u(q) - p_i q\} + \alpha_i \left[\tilde{u}\left(q\left(p_i\right)\right) - r_i q\left(p_i\right)\right] + \alpha_j \left[\tilde{u}\left(q\left(p_j\right)\right) - r_i q\left(p_j\right)\right] - F_i.$$

Because network i's market share strictly increases with w_i (actually with $w_i - w_j$), there is a one-to-one relationship between F_i and α_i. Thus, network i's choosing a tariff (F_i, p_i) is

[7] When market size is not constant (and so $\alpha_1 + \alpha_2$ are coverage fractions rather than market shares, with $\alpha_1 + \alpha_2 < 1$), then α_i increases with both $w_i - w_j$ (relative attractiveness) and w_i (absolute attractiveness). We will discuss incomplete market coverage later. A simple incomplete-coverage model in which the analysis of full coverage applies directly goes as follows: *Stage 1*: Consumers must spend a cost s to learn about the existence of the product, familiarize themselves with the technology, and learn their location in the product space. Consumers are heterogeneous and distributed according to cumulative distribution function $H(s)$ on $[0, \infty)$. *Stage 2*: Networks select (F_i, p_i) noncooperatively. Consumers who have incurred the cost select their network.

[8] We are grateful to a referee for prompting us to start the roadmap with this benchmark and for various other suggestions on this roadmap.

© RAND 2004.

equivalent to choosing a calling price p_i and a market share α_i, for any given competitive offer by network j; in the latter interpretation, the impact of a change in p_i for a given α_i must be accompanied by a change in F_i that keeps the relative attractiveness, $w_i - w_j$, of the two networks and therefore the market share constant. This alternative formulation is perhaps economically less intuitive than the choice of a tariff, but it turns out to be very convenient technically.

Our first insight is that although a network's reception charge does not affect its profit given its rival's competitive offer, reception charges do matter. Indeed, we show that network i's equilibrium usage (calling) charge is equal to its "strategic marginal cost," namely

$$p_i = [c + \alpha_j(a - c_0)] - \alpha_i r_j,$$

where, recall, c is the (industry's) marginal cost of a call, α_j is network j's market share, $(a - c_0)$ is the access charge markup (the difference between the access charge and the termination cost), and r_j is network j's reception charge. An average call originating on network i costs $[c + \alpha_j(a - c_0)]$ when the access markup (or discount) on the fraction α_j of calls that terminate off-net is accounted for. The externalities created by this extra unit on the receivers are equal to $\tilde{u}'(q) - r_i$ for network i receivers and $\tilde{u}'(q) - r_j$ for network j receivers (see the expression of w_i above). Hence, to keep the relative attractiveness $w_i - w_j$ of the two networks' offers constant, network i must lower its fixed fee by $r_i - r_j$. But since it receives r_i from network i receivers and only the effective fixed fee matters, the net cost is $\alpha_i[(r_i - r_j) - r_i] = -\alpha_i r_j$. So to obtain network i's perceived or strategic marginal cost, one subtracts the increase r_j in the monthly fee of the fraction α_i of consumers who subscribe to network i, which network i can afford to implement without losing market share.

To see this, consider an increase in the volume of calls from network i to network j. This increase has two opposite effects on the utility of network j's consumers. Their surplus increases because they receive more calls, and they pay for receiving these extra calls. We call the first effect a *direct* externality and the second a *pecuniary* externality. Only pecuniary externalities matter for determining the perceived marginal cost: since the volume of calls received by consumers increases by the same amount regardless of their network affiliation, direct externalities are the same for all consumers. The fact that consumers of network i pay more in reception charges from the increase in the volume does not affect the marginal cost perceived by network i, since any impact on the effective fixed fee is internalized by network i. However, because consumers of network j pay more in reception charges, network i can increase its fixed fee by $\alpha_i r_j$ and still maintain its market share. Pecuniary externalities thus result in a decrease in the perceived marginal cost $c + \alpha_j(a - c_0) - \alpha_i r_j$. That is, an increase in network j's reception charge makes it more desirable for network i to expand output.

In a symmetric equilibrium ($\alpha_1 = \alpha_2 = 1/2$) under symmetric reception charges ($r_i = r_j = r$), $p = c + (a - c_0 - r)/2$. Thus a regulator who would control both the termination and the reception charges could *a priori* avail herself of two instruments to reach one target ($p^* = c/(1 + \beta)$). However, as was shown by the earlier literature, the existence of an equilibrium cannot be taken for granted if networks offer close substitutes. The logic (which generalizes the earlier work) is that if the equilibrium price p differs substantially from the level, $c - r$, that is optimal for $\alpha_i = 1$, a network can gain from internalizing all traffic by making a first-order gain through rebalancing the usage price to $c - r$. Thus existence in the case of close substitutes requires that $p \simeq c - r$, or $r \simeq c_0 - a$. Hence,

$$r = r^* = c_0 - a^* \equiv \frac{\beta c}{1 + \beta}.$$

That is, termination is priced at a discount, and this discount is steeper, the higher the receiver's marginal utility from receiving calls.

Last, the fiction of caller-determined volume (our thought experiment), while open to the criticism that receivers are actually sovereign, turns out to be very useful, as we provide conditions under which it is still valid when receivers are allowed to hang up.

© RAND 2004.

Part 2: market-determined reception charges. Next, we assume that reception charges, just like monthly fees and calling charges, are set noncooperatively by the operators instead of being determined contractually or chosen by a regulator. In the absence of uncertainty about marginal utilities, a potential indeterminacy of equilibria arises: Over the range of parameters for which the volume is determined by the caller, only $\{F_i + r_i\tilde{q}\}$ matters, not its composition, as we have seen. And so, there may be a range of (nonequivalent) equilibria. We make the model more realistic and actually simplify it by letting the marginal utilities of communications be random (the utility of an extra minute of communication is state or time-of-day contingent). With wide enough supports for marginal utilities, both the calling and the reception charges affect volume and therefore are determinate.

To build intuition for our result, first note that call reception is the mirror image of call issuance. While the gross surplus functions $u(\cdot)$ and $\tilde{u}(\cdot)$ in general differ, there is no qualitative difference between calling and being called. Hence the formulae giving the strategic marginal cost for call reception is identical to that for call issuance, after adjusting for the difference in average cost ($\alpha_i c + \alpha_j(c_0 - a)$ instead of $c + \alpha_j(a - c_0)$) and replacing network j's receiver charge r_j by network j's calling charge p_j. Thus the net cost of encouraging an additional call to be made is $c + \alpha_j(a - c_0) - \alpha_i r_j$ while, by symmetry, the net cost of encouraging an additional call on the reception side is $\alpha_i c + \alpha_j(c_0 - a) - \alpha_i p_j$. The noise in the marginal utilities implies that sometimes the caller determines the volume and sometimes the receiver does (this is what the introduction of noise buys us). For this reason, networks charge their net (or strategic) cost of an additional call on both the dialing and receiving sides. The corresponding two equations, whether or not market shares are equal (i.e., regardless of the values of α_i and α_j), imply that the usage and reception charges are set equal to the "off-net cost" of call and call reception, respectively:

$$p = c + (a - c_0)$$

and

$$r = (c_0 - a).$$

That is, each network sets prices for a subscriber's outgoing and incoming traffic at the marginal cost that it would incur if *all* other subscribers belonged to the rival network. Furthermore, $r = c_0 - a$; so if $a = a^*$ $(= c_0 - r^* = c_0 - (\beta c/(1 + \beta)))$, then the market implements the social optimum when the randomness in the marginal utilities vanishes.

Part 3: network-based price discrimination. Last, we allow networks to differentiate their calling and reception charges according to whether the communication is on- or off-net. In the presence of network-based discrimination, we need to separate the market for on-net calls from the market for off-net calls. In the former market, regardless of the introduction of reception charges, each network fully internalizes the externalities on callers and receivers. By contrast, in the latter market, the off-net calling and reception charges affect the welfare of consumers on the rival network and are therefore subject to strategic manipulation. In particular, we find that the direct externalities on the rival network's consumers can create strong incentives for connectivity breakdown, where the latter is defined as a situation in which high reception or calling charges choke off-net traffic.

Intuitively, network i's profitability is determined by the attractiveness of its offer *relative* to that of its competitor. The substantial degradation of the *absolute* quality offered by networks that choke off-net traffic through high off-net prices has no impact on profitability as long as the market is covered.[9]

We first show that connectivity breakdown can occur in the absence of reception charges: each network charges an off-net call price going to infinity when the receivers on the other network

[9] More generally, the connectivity breakdowns would occur as long as the competitive threat of the rival network is perceived as much more intense than that of "outside options." Note also the less drastic version of our result: Competition for market share leads to suboptimal connectivity under network-based price discrimination.

© RAND 2004.

benefit from communications almost as much as (β converges to 1) or more than ($\beta > 1$) the callers. Intuitively, a network bears a positive cost of off-net calls (provided the total cost c exceeds the termination discount $c_0 - a$) and does not gain in relative attractiveness by allowing off-net calls; so it might as well shut them off.

Introducing reception charges, *a priori,* should reduce the connectivity breakdown problem by lowering the sender's network's strategic marginal cost. However, it provides a second instrument for connectivity breakdown: the receiver's network can induce its consumers to hang up. We show, in fact, that in the presence of reception charges, network-based price discrimination often leads to connectivity breakdown through infinite calling or reception charge. The logic of the connectivity breakdown here is unrelated to the standard one emphasized in the network externalities literature,[10] which is driven by the incentive of a dominant player (characterized by a large installed base or a cost superiority) to reinforce dominance by reducing connectivity. Here, connectivity breakdown occurs even with symmetric operators.

The article is organized as follows. Section 2 describes the framework, introduces the relevant concepts and notation, and derives the social optimum benchmark. Section 3 analyzes competition in nonlinear prices in the absence of network-based price discrimination and for regulated reception charges, and then Section 4 studies market-determined reception charges. Section 5 performs similar analysis with network-based price discrimination. Section 6 concludes.

□ **Related literature.** After investigating the British telephone industry, Oftel (1998) concluded that the price of calls from fixed networks to mobile phones was too high and envisioned price regulation. Indeed, with the caller-pays principle (CPP), there might be little competitive pressure on the termination charge that a mobile phone operator can demand for terminating calls originating in the fixed network, and the resulting high charges hurt the consumers of the fixed network. Opposing Oftel's proposed price regulation, Doyle and Smith (1998) offer instead to apply the receiver-pays principle (RPP), by which mobile operators charge their own customers for receiving calls while callers pay a uniform per-minute charge, e.g., the local call charge, regardless of where the calls terminate. They study a model with a monopoly fixed-link network and a duopoly of mobile operators. They first study CPP, assuming that each mobile network sets its own termination charge to be paid by the fixed network. This situation leads to a strong form of double marginalization on fixed-to-mobile calls (all the more that receivers are assumed not to derive surplus from being called). Turning to RPP, they show that mobile operators compete on reception charges to attract customers, which leads to a lower total charge of a fixed-to-mobile call as well as increased usage.

Kim and Lim (2001), as we do, address the question of how RPP may help with the internalization of the call externalities when subscribers derive utility from receiving calls. They consider two models without regulation of the access charge and the reception charge and with linear, unregulated pricing of calls. Their first model is a monopoly model. The introduction of a linear receiver charge decreases the perceived marginal cost of a call for the network and therefore leads to a lower price of a call. However, the effect on the total price (call price plus reception charge) and on welfare depends on how the price elasticity of demand varies with price. In the second model, they introduce call externalities in the Laffont, Rey, and Tirole (1998a) model and assume that the access charge is set cooperatively by the networks before competing in linear prices. A network operator charges reception to all consumers (both his own subscribers and the subscribers of the other network) for calls initiated on his own network. In contrast, we assume that each operator sets reception charges for his own customers. Again, they show that the calling price decreases with RPP but that the access charge is higher, with ambiguous results on welfare.

DeGraba (2000) argues that the bill-and-keep policy (zero access charges) leads to efficient pricing in a symmetric case where the origination and termination costs are equal and the called party and the calling party benefit equally from a phone call, but does not build a formal model.[11]

[10] See, e.g., Katz and Shapiro (1985) and Crémer, Rey, and Tirole (2000).

[11] See Atkinson and Barnekov (2000), who also argue in favor of a bill-and-keep policy but don't build a formal analysis.

© RAND 2004.

The articles most closely related to ours are Berger (2001), Hermalin and Katz (2001), and Laffont et al. (2003). Laffont et al. analyzes the impact of termination charges in the Internet. Like our article, Laffont et al. looks at the impact of inbound and outbound charges in an interconnected-networks environment and derives the off-net-cost pricing principle. The objectives of the two articles are quite different, though. Laffont et al. is primarily an application-driven work that assumes a very specific environment in order to address a variety of (primarily) Internet-specific questions. Namely, it assumes that there is a fixed volume of transactions (say, one) for each dial-up consumer (receiver)–website (caller) match and that all connections deliver the same gross surplus to a given end user. Thus nonlinear tariffs are irrelevant; that is, the distinction between monthly fee and usage fees (call or reception charge) that figures prominently in our work plays no role in Laffont et al. Similarly, there is no issue of joint determination of volume. Our article thus provides a much more general determination of the off-net-cost pricing principle. In particular, it shows that when volume is variable, the marginal cost perceived by each network is affected by the externalities on the rival network's subscribers, and this drives networks to charge prices equal to off-net costs while, when volume is fixed, there are no such externalities. Furthermore, it shows that network-based price discrimination is conducive to connectivity breakdowns, a question that is not investigated in Laffont et al. By contrast, Laffont et al. studies the impact of multihoming, website compression strategies, payments between end users, and asymmetric charges. Hermalin and Katz (2001) similarly focus on unit transactions and linear pricing; they first show that double marginalization arises when networks specialize in offering services to senders or receivers and have market power on their respective segments. They then study Bertrand competition between two undifferentiated networks when end users' valuations are independent (as in Laffont et al.) or perfectly correlated. Last, Berger (2001) extends the analysis of Laffont, Rey, and Tirole (1998b) and Gans and King (2001) to environments in which the receivers derive utility from receiving calls. Berger's analysis complements that of Section 5 in that (i) it does not study connectivity breakdowns but rather focuses on characterizing the equilibrium and privately and socially optimal access charges when networks are sufficiently differentiated and (ii) it focuses on the caller-pays regime.

2. Framework

■ **The model.** *Demand side.* We extend the analysis of network competition in Laffont, Rey, and Tirole (1998a, 1998b) in two respects: Receivers obtain positive utility from receiving calls, and firms can charge receivers for reception.

There are two operators (suppliers, networks), $i = 1, 2$, located at the two extremes of a Hotelling line of length one ($x_1 = 0, x_2 = 1$). Consumers are differentiated along the Hotelling line. A consumer located at x and selecting network i incurs "transportation cost" $t|x - x_i|$.

The utility of a consumer with income y located at x from joining network i is given by

$$y + v_0 - t|x - x_i| + u(q) + \tilde{u}(\tilde{q}),$$

where $u(q)$ is the utility from calls placed by the consumer and $\tilde{u}(\tilde{q})$ represents the utility from received calls.[12] We assume that these utility functions $u(\cdot)$ and $\tilde{u}(\cdot)$ are twice continuously differentiable, with $u' > 0, u'' < 0; \tilde{u}' > 0, \tilde{u}'' < 0$, which implies that demand functions are differentiable. $u(q)$ and $\tilde{u}(q)$ can be thought of as the caller's and receiver's surpluses attached to a representative call (lasting q minutes). We assume that the receiver's marginal surplus from receiving a call is nonnegative.

We consider four different cases depending on whether network-based (on-net/off-net) price

[12] The constant v_0 in the utility function for a consumer who has joined a network ensures that all consumers will always choose to join one of the two networks if v_0 is high enough. One can think of $v_0 - t|x - x_i|$ as the net utility of having access to basic services such as being able to call (or be called by) family or doctors in case of emergency. Such phone services have high utilities relative to their cost to the consumers and are demanded rather inelastically, contrary to those depicted by the functions $u(\cdot)$ and $\tilde{u}(\cdot)$.

© RAND 2004.

discrimination is allowed, and on whether the volume of calls is determined only by callers or jointly by callers and receivers.

(i) *Price discrimination.* In the absence of network-based price discrimination (Sections 3 and 4), network i offers a three-part tariff $\{F_i, p_i, r_i\}$. F_i is the monthly subscriber charge, p_i is the (caller's) usage price, and r_i represents the per-unit price that network i's consumers pay for the calls received. Under network-based discrimination (Section 5), network i offers a five-part tariff, $\{F_i, p_i, \hat{p}_i, r_i, \hat{r}_i\}$. Here, $\hat{r}_i$ represents the price that a network-i consumer pays for receiving calls originating on network $j \neq i$, while r_i is the reception charge for on-net calls. Similarly, p_i and $\hat{p}_i$ refer to per-unit charges for calls that terminate on- and off-net, respectively.

(ii) *Demand function.* We first make the standard assumption that callers determine the volume. Letting $q(\cdot)$ denote the caller's demand function, given by $u'(q(p)) = p$, the volume of calls placed by a customer of network i is given by $q(p_i)$ in the absence of discrimination, and by $q(p_i)$ and $q(\hat{p}_i)$ in the case of discrimination. Let $v(p)$ be the indirect utility function, i.e.,

$$v(p) = \max_q \{u(q) - pq\}.$$

We then consider the case in which receivers are sovereign, i.e., callers and receivers jointly determine the volume. Consider a representative caller-receiver pair. In the absence of discrimination, the volume of calls from (the caller's) network i to (the receiver's) network j is given by $\min\{q(p_i), \tilde{q}(r_j)\}$, where the receiver's demand function is given by $\tilde{u}'(\tilde{q}(r)) = r$. Under discrimination, the volume of calls if both belong to network i is given by $\min\{q(p_i), \tilde{q}(r_i)\}$, and the volume of calls if the caller belongs to network i and the receiver to network $j \neq i$ is given by $\min\{q(\hat{p}_i), \tilde{q}(\hat{r}_j)\}$.

Supply side. The local loop cost is decomposed into a traffic-sensitive marginal cost, c_0 per unit of volume, and a traffic-insensitive component. The traffic-insensitive part is composed of a per-consumer connection component f, plus possibly some cost that is joint and common to all subscribers. For notational simplicity, we will ignore the latter (introducing a joint and common cost would just require an overall upward adjustment of price levels keeping the price structure as given, if firm viability is an issue). The long distance (trunk) marginal cost is equal to c_1. So, the total marginal cost of a minute of a call involving two local loops and long distance is

$$c \equiv 2c_0 + c_1.$$

We let a denote the access charge or termination charge. The marginal cost of an off-net call is therefore $c + (a - c_0)$ for the caller's network and $(c_0 - a)$ for the receiver's network.

□ **Ramsey benchmark.** For future reference, we derive the social optimum. Consider an idealized situation in which a benevolent regulator would choose the market shares and the volume of calls. In our symmetric setup, equal market division ($\alpha = 1/2$) minimizes the average consumer's disutility from not being able to consume his preferred service. The benevolent regulator would choose the volume of calls so as to maximize

$$u(q) + \tilde{u}(q) - cq.$$

The optimal volume q^* is determined by

$$u'(q^*) + \tilde{u}'(q^*) = c.$$

To implement the optimal outcome, the benevolent regulator can use symmetric tariffs so as to implement equal market division ($\alpha = 1/2$). When the volume is determined by callers, the

© RAND 2004.

optimal volume is obtained by choosing $p_1 = p_2 = p^*$, where

$$p^* = c - \tilde{u}'(q^*).$$

The regulator selects the fixed fee F (or r) in order to satisfy the industry's break-even constraint.

When the volume is jointly determined by callers and receivers, the regulator can still achieve the efficient outcome by choosing $(p = p^*, r = 0)$. The regulator then uses F to satisfy the industry break-even constraint. More generally, any $\{F, r\}$ combination yielding the same level of $F + rq^*$ and such that $r \leq \tilde{u}'(q^*)$ achieves the Ramsey outcome. Thus, reception charges are not needed. The same is true for a monopoly network. It is in the latter's interest to set $\{p = p^*, r = 0)\}$ in order to generate maximal surplus from its consumers. The only monopoly distortion is on the level of the monthly charge F.

3. Regulated or contractually determined reception charges

■ This section studies competition in two-part tariffs without discrimination (F_i, p_i) (when the reception charges are exogenously regulated at some levels $\{r_i\}_{i=1,2}$). It first assumes that callers determine the volume of calls, and then finds sufficient conditions for this to be the case.

Ignoring the "transportation cost," the net surplus of a network-i consumer is

$$w_i = v(p_i) + \alpha_i \tilde{u}(q(p_i)) + \alpha_j \tilde{u}(q(p_j)) - r_i[\alpha_i q(p_i) + \alpha_j q(p_j)] - F_i. \quad (1)$$

Because subscribers are uniformly distributed on the Hotelling interval, network i's market share is given by

$$\alpha_i = \frac{1}{2} + \sigma(w_i - w_j), \quad (2)$$

where $\sigma = 1/(2t)$ measures network substitutability, and thus the intensity of competition. Equivalently,

$$\alpha_i = \frac{1}{2} + \sigma[v(p_i) - v(p_j) - (F_i - F_j) - (r_i - r_j)(\alpha_i q(p_i) + \alpha_j q(p_j))]. \quad (3)$$

Network i's profit is given by

$$\pi_i = \alpha_i\{[p_i - c - \alpha_j(a - c_0) + \alpha_i r_i]q(p_i) + (r_i + a - c_0)\alpha_j q(p_j) + F_i - f\}.$$

□ **Pricing usage at the strategic marginal cost.** As explained in the overview, given p_i and network j's tariff, F_i and α_i are one-to-one. Maximizing network i's profit with respect to (F_i, p_i) is therefore equivalent to optimizing with respect to (α_i, p_i). So we will perform our analysis in two steps. First, we will maximize π_i with respect to p_i given α_i, yielding price $p_i^{**}(\alpha_i)$. This will allow us to define $\Pi_i(\alpha_i) \equiv \pi_i(\alpha_i, p_i^{**}(\alpha_i))$. Second, we will maximize $\Pi_i(\alpha_i)$ with respect to α_i. Note that all choices are *simultaneous*. Our two-step analysis is convenient because the maximization with respect to the usage price keeping market share constant directly yields the choice of the usage price, while the optimization with respect to market share in a sense yields the networks' overall price level. We are interested more in the former (which affects efficiency) than in the latter (which does not, given that the market is assumed to be covered).

Thus, we first study the program of maximizing π_i keeping market share α_i constant. Let

$$\tilde{F}_i = F_i + r_i(\alpha_i q(p_i) + \alpha_j q(p_j)).$$

Intuitively, $F_i + r_i(\alpha_i q(p_i) + \alpha_j q(p_j))$ is a generalized fixed fee. Network i's consumers care only about this sum, not about its composition.

© RAND 2004.

Market shares are determined by the net surplus differential:

$$w_i - w_j = v(p_i) - v(p_j) - \tilde{F}_i + \tilde{F}_j. \tag{4}$$

We have

$$\pi_i \equiv \alpha_i\{(p_i - c)q(p_i) - (a - c_0)(1 - \alpha_i)(q(p_i) - q(p_j)) + \tilde{F}_i - f\}.$$

Using (2) and (4) we have

$$\tilde{F}_i = \tilde{F}_j + v(p_i) - v(p_j) + \frac{1}{\sigma}\left(\frac{1}{2} - \alpha_i\right).$$

After substitution of $\tilde{F}_i$ into the profit function, we have

$$\pi_i(p_i, \alpha_i) \equiv \alpha_i\{(p_i - c)q(p_i) - (a - c_0)(1 - \alpha_i)(q(p_i) - q(p_j)) + v(p_i) - v(p_j) + \tilde{F}_j + \frac{1}{2\sigma} - \frac{\alpha_i}{\sigma} - f\},$$

where $\tilde{F}_j$ is a function of p_i. Indeed,

$$\frac{\partial \tilde{F}_j}{\partial p_i} = r_j \alpha_i \frac{dq}{dp_i}.$$

The first-order derivative of network i's profit with respect to p_i keeping α_i constant is given by

$$\left.\frac{d\pi_i}{dp_i}\right|_{\alpha_i=\text{constant}} = \alpha_i\left[p_i - c - \alpha_j(a - c_0) + \alpha_i r_j\right]\frac{dq}{dp_i}, \qquad \text{for } p_i > 0.$$

For a given α_i, the profit-maximizing price p_i is therefore given by $p_i^{**}(\alpha_i)$:

$$p_i^{**}(\alpha_i) = c + \alpha_j(a - c_0) - \alpha_i r_j. \tag{5}$$

In fact, we have, for $p_i^{**}(\alpha_i) > 0$,

$$\left.\frac{d\pi_i}{dp_i}\right|_{\alpha_i=\text{constant}} > 0, \qquad \text{for all } 0 < p_i < p_i^{**}(\alpha_i),$$
$$\left.\frac{d\pi_i}{dp_i}\right|_{\alpha_i=\text{constant}} = 0, \qquad \text{for } p_i = p_i^{**}(\alpha_i),$$
$$\left.\frac{d\pi_i}{dp_i}\right|_{\alpha_i=\text{constant}} < 0, \qquad \text{for all } p_i > p_i^{**}(\alpha_i).$$

If $c > r_j$ holds for $a - c_0 + r_j \geq 0$ or if $c + a - c_0 > 0$ holds for $a - c_0 + r_j < 0$, we have $p_i^{**}(\alpha_i) > 0$.

If $r_j = 0$, the usage price is equal to the average marginal cost faced by network i as in Laffont, Rey, and Tirole (1998a). However, in the presence of r_j, an increase in $q(p_i)$ imposes pecuniary externalities on network-j consumers by making them pay more money for the calls received from network i. Hence, to maintain market share α_i constant, network i can charge more money to the consumers of its own network by increasing F_i. In other words, network j's charging for reception results in a decrease in the marginal cost perceived by network i. For example, if the access charge is near or above termination cost, charging receivers is socially desirable because it induces firms to lower calling charges.

© RAND 2004.

The profit-maximizing price p_i is uniquely given by $p_i^{**}(\alpha_i)$, and we can define $\Pi_i(\alpha_i)$ by

$$\begin{aligned}\Pi_i(\alpha_i) &\equiv \pi_i(p_i^{**}(\alpha_i), \alpha_i)\\ &= \alpha_i \Bigg\{ (p_i^{**} - c)q(p_i^{**}) - (a - c_0)(1 - \alpha_i)(q(p_i^{**}) - q(p_j))\\ &\qquad + v(p_i^{**}) - v(p_j) + \tilde{F}_j(p_i^{**}, \alpha_i) + \frac{1}{2\sigma} - \frac{\alpha_i}{\sigma} - f \Bigg\}.\end{aligned}$$

Proposition 1 (strategic marginal cost).

(i) When volume is determined solely by the caller, only the sum of the monthly subscriber charge and the subscriber's total reception charge matters, given the rival network's competitive offering.

(ii) In the absence of network-based price discrimination, a network's marginal cost and calling charge decreases with the other network's reception charge:

$$p_i^{**}(\alpha_i) = c + \alpha_j(a - c_0) - \alpha_i r_j.$$

As explained in Section 2, the socially optimal marginal price is

$$c - \tilde{u}'(q^*),$$

which could be obtained by pricing access at its marginal cost $a = c_0$ and by subsidizing calls at a rate

$$\tau = \tilde{u}'(q^*)$$

in the absence of reception charges.

In the absence of subsidization, two instruments—the access charge and the reception charge—can be used to induce the optimal marginal price. Because network i internalizes only the positive externality on its own consumers, the reception charge would have to be (in a symmetric equilibrium) twice the marginal externality if the termination charge were set equal to termination cost ($a = c_0$). But this would induce receivers to hang up. The fact that the reception charge cannot exceed the marginal externality mandates an access charge *below* the termination cost.

We now maximize $\Pi_i(\alpha_i)$ with respect to α_i. Since we will focus on symmetric equilibria, we assume $r_j = r$ and study the maximization of $\Pi_i(\alpha_i)$ given that network j optimally sets $p_j = c + (1/2)(a - c_0 - r)$. This yields the following first-order condition:

$$(p_i - c)q_i - (a - c_0)(q_i - q_j)(1 - 2\alpha_i) + v_i - v_j - f + \frac{1}{2\sigma} - \frac{2\alpha_i}{\sigma} + F_j + r(2\alpha_i q_i + (1 - 2\alpha_i)q_j) = 0. \quad (6)$$

□ **Symmetric equilibria.** *Characterization*. For a given reception charge r, equilibria (p_i, F_i, α_i) are characterized by (3), (5), and (6). We are interested in symmetric equilibria: $(p, F, \alpha = 1/2)$ for a given value of r. From (5),

$$p = c + \frac{1}{2}(a - c_0 - r). \quad (7)$$

Equation (6) then yields

$$F = f + \frac{1}{2\sigma} - (p + r - c)q(p). \quad (8)$$

After some computations, we obtain $\pi = 1/(4\sigma)$ when (F, p) satisfy (7) and (8). Thus, the profit is always equal to the Hotelling profit with unit demand, as in Laffont, Rey, and Tirole (1998a). Provided that symmetric equilibria exist, the access charge and the reception charge have

© RAND 2004.

no impact on profit. This particular result is an artifact of our full-coverage assumption. When consumers have alternative choices, then a and r affect surplus and the networks' market shares vis-à-vis these outside alternatives.[13]

Summarizing, we have the following proposition.

Proposition 2 (characterization). The symmetric equilibrium is characterized by

(i) $p = c + (1/2)(a - c_0 - r)$, and $F = f + (1/2\sigma) - (p + r - c)q(p)$.

(ii) Each firm's profit is equal to $1/(4\sigma)$.

Existence. The study of existence of an equilibrium generalizes and closely parallels that in Laffont, Rey, and Tirole (1998a). We therefore only provide the results (proofs can be found in our discussion paper, Jeon, Laffont, and Tirole (2001)). To show the existence of a symmetric equilibrium, let us restrict the analysis to meaningful values of the access charge and the reception charge: $\infty > a > c_0 - c$ and $c > r > -\infty$. Then we have the following lemma.

Lemma 1. $\Pi_i(\alpha_i) \equiv \pi_i(p_i^{**}(\alpha_i), \alpha_i)$ is well defined and continuous. Furthermore, if σ is small enough or $|a - c_0 + r|$ is small enough, it is concave.

Proposition 3 (existence).

(i) If $\Pi_i(\alpha_i)$ is concave, a symmetric equilibrium $(p, F, \alpha = 1/2)$ exists.

(ii) No cornered-market equilibrium exists.

(iii) For any $\varepsilon > 0$, if $|a - c_0 + r| > \varepsilon$, no equilibrium exists for σ large enough.

The case of a small substitutability σ is relevant only if the networks are specialized in the sense of fitting geographic or technological niches.[14] For large substitutability, a small amount of undercutting (say, a small reduction in the monthly fee) allows a network to corner the market. Rewrite network i's optimal price as

$$p_i = c - r + \alpha_j(a - c_0 + r).$$

Now, if $|a - c_0 + r|$ weren't small, the price p_i charged in a candidate equilibrium in which, say, the networks share the market would be quite different from the one, $c - r$, that is privately optimal when cornering the market ($\alpha_j = 0$). Hence, cornering the market would involve a small fixed fee reduction (for large substitutability) and a sizeable gain from price rebalancing. This suggests why the reception charge and the termination discount must be nearly equal in order to obtain existence.

□ **Receiver sovereignty.** The above analysis has not taken into account the fact that receivers may want to hang up. However, the above existence and optimality results are readily extended to the case of receiver sovereignty as long as the regulated reception charge does not exceed the marginal utility of reception; $r \leq \tilde{u}'(q(p))$.

Proposition 4. A symmetric equilibrium $(p^e, F^e, \alpha = 1/2)$ in the absence of receiver sovereignty (where p^e and F^e are given in Proposition 2) is still an equilibrium under receiver sovereignty as long as

$$r \leq \tilde{u}'(q(p^e)).$$

Proof. Consider an equilibrium in the absence of receiver sovereignty. In particular, for each network i, p^e maximizes profit (given market share) over the domain $P = \{p_i : r \leq \tilde{u}'(q(p_i))\}$.

[13] See Laffont et al. (2003) for an illustration of this point in a Hotelling model with hinterlands. Another reason why the neutrality result may break down, unveiled in Gans and King (2001), arises in the presence of network-based price discrimination; the termination discount can boost profits by giving networks an incentive to stay small.

[14] It may also be relevant to study the competition between one fixed and one mobile network.

© RAND 2004.

Because under receiver sovereignty network i's profit becomes insensitive to p_i for p_i such that $r > \tilde{u}'(q(p_i))$ (since demand is then determined by r), p^e still maximizes profit in the extended domain $P' = \{p_i : p_i \geq 0\}$. *Q.E.D.*

□ **Social-welfare-maximizing equilibrium.** Ignoring for a moment receiver sovereignty, efficiency requires that

$$u'(q(p^*)) + \tilde{u}'(q(p^*)) = c,$$

and so, using $u'(q(p^*)) = p^*$ and condition (7),

$$\tilde{u}'(q(p^*)) = \frac{1}{2}(r + c_0 - a). \tag{9}$$

An efficient choice of termination and reception charges must therefore satisfy (9). This gives us one degree of freedom,[15] which we can use in the following way: choose the highest reception charge that is consistent with receiver sovereignty,

$$r^* = \tilde{u}'(q(p^*)),$$

or, from (9),

$$r^* = c_0 - a.$$

Then, from Proposition 3, an equilibrium exists and it implements the social optimum.

Proposition 5 (efficiency). Let the termination and reception charges satisfy

(i) $r^* = c_0 - a^*$ (reception charge = termination discount) and

(ii) $r^* = \tilde{u}'(q(p^*))$ (equalization of sender's and receiver's demands).

Then a symmetric equilibrium exists and is efficient.

Remark. Suppose that reception charges are prohibited. Then, ignoring existence problems (which do not arise if the networks are not too substitutable—say they are fixed and mobile networks), the termination charge must bear the brunt of the burden to achieve efficiency. The termination discount, $c_0 - a$, should now be $2\tilde{u}'(q(p^*))$, which is *twice* the level in Proposition 5.[16]

4. Market-determined reception charges

■ **Heuristics.** As we discussed in the Introduction, the absence of uncertainty about marginal utilities makes each operator locally indifferent as to the composition of his "effective fixed fee" between the fixed fee F_i and the reception charge r_i in the range of parameters in which the caller's net marginal utility strictly exceeds the receiver's net marginal utility. This is unrealistic and complicates the analysis as well. In reality, even assuming that receivers value calls less than senders on average, the receiver's utility, for example, may be subject to noise: for example, one is less eager to stay long on the phone when a visitor is in one's office or when watching over young children. When reception charges are market determined, it turns out to be convenient to allow for such state-contingent marginal utilities, since then both the calling and the reception charges affect demand.

When the receiver's demand has (at least a small) probability of determining overall traffic, the symmetry in the principle guiding the choice of the two usage prices is reestablished. Recall from Proposition 1 that (outbound) calls are priced at their strategic marginal cost:

$$p_i = \left[c + \alpha_j(a - c_0)\right] - \alpha_i r_j. \tag{10}$$

[15] For example, if $\tilde{u}(q) = \beta u(q)$, then $r + c_0 - a = 2\beta c/(1 + \beta)$.

[16] Intuitively, as $p^* = c + ((a - c_0)/2) - (r^*/2)$, if r^* is equal to zero rather than to $c_0 - a$, then the termination discount must be doubled to maintain volume efficiency.

© RAND 2004.

Intuitively, the networks also optimally charge their strategic marginal cost for inbound calls:

$$r_i = \left[\alpha_i c + \alpha_j (c_0 - a)\right] - \alpha_i p_j. \tag{11}$$

To understand (11), note that a fraction α_i of calls received by a network i subscriber are on-net and cost c, and the fraction α_j of terminated off-net calls cost the termination discount $(c_0 - a)$. From this average cost must be subtracted the reduction p_j in network-i subscribers' fixed fee needed to keep market share constant; this term corresponds to the pecuniary externality and is the counterpart of r_j in (10).

For any market division, equations (10) and (11) admit the symmetric solution

$$p = c + (a - c_0) \tag{12}$$

and

$$r = c_0 - a. \tag{13}$$

That is, the networks price calls and call receptions at their *off-net* costs; they choose their prices *as if* they had a single consumer and so all inbound and outbound traffic were off-net.

Besides this remarkably simple pricing rule, the analysis unveils two interesting properties:

(i) While our symmetric model predicts symmetric market shares ($\alpha_1 = \alpha_2 = 1/2$), the off-net-cost pricing principle is broader. For example, network 1 might have a larger installed base and therefore a higher market share than network 2 ($\alpha_1 > \alpha_2$). Equations (10) and (11) would still yield off-net-cost pricing.

(ii) When the demand uncertainty vanishes (keeping a large support, though, so as to make the reception charge determinate), the reception charge (given by (13)) corresponds to the regulated value that guarantees existence of an equilibrium for high network substitutability.

□ **Full analysis.** Suppose that the marginal utility that a receiver derives from receiving a call is subject to a noise ε.[17] The receiver's utility is

$$\tilde{u}(q) + \varepsilon q.$$

We assume that ε follows the distribution function $F(\cdot)$, with wide enough support $[\underline{\varepsilon}, \overline{\varepsilon}]$, zero mean, and density $f(\cdot)$, which is strictly positive for all ε in $[\underline{\varepsilon}, \overline{\varepsilon}]$; and that the noise ε is identically and independently distributed for each caller-receiver pair.

For technical simplicity only, we further assume in this section that

$$\tilde{u}(q) = \beta u(q) \qquad \text{with } \beta > 0.$$

We first study how the volume is determined given (p_i, r_j) and a realized value ε of the random variable. Unless the caller interrupts the conversation first, the receiver with noise ε will equate his marginal utility $\tilde{u}' + \varepsilon$ to the reception charge r_j. Hence, the volume of calls is given by $q(\max(p_i, (r_j - \varepsilon)/\beta))$. Therefore, the volume of calls from network i to network j is given by

$$\alpha_i \alpha_j D(p_i, r_j),$$

with

$$D(p_i, r_j) \equiv \left[1 - F(r_j - \beta p_i)\right] q(p_i) + \int_{\underline{\varepsilon}}^{r_j - \beta p_i} q\left(\frac{r_j - \varepsilon}{\beta}\right) f(\varepsilon) d\varepsilon.$$

[17] The caller's marginal utility could also be subject to a noise without any change in the results. The important feature of the following analysis is that both the caller and the receiver have positive probability of hanging up first.

© RAND 2004.

Similarly, the utility that a network-i consumer derives by making calls to network-j consumers is given by

$$\alpha_j U(p_i, r_j),$$

with

$$U(p_i, r_j) \equiv \left[1 - F(r_j - \beta p_i)\right] u(q(p_i)) + \int_{\underline{\varepsilon}}^{r_j - \beta p_i} u\left(q\left(\frac{r_j - \varepsilon}{\beta}\right)\right) f(\varepsilon) d\varepsilon.$$

The utility that a network-j consumer derives from receiving calls from network-i consumers is given by

$$\alpha_i \tilde{U}(p_i, r_j),$$

with

$$\tilde{U}(p_i, r_j) \equiv \int_{r_j - \beta p_i}^{\overline{\varepsilon}} [\tilde{u}(q(p_i)) + \varepsilon q(p_i)] f(\varepsilon) d\varepsilon + \int_{\underline{\varepsilon}}^{r_j - \beta p_i} \left[\tilde{u}(q\left(\frac{r_j - \varepsilon}{\beta}\right)) + \varepsilon q\left(\frac{r_j - \varepsilon}{\beta}\right)\right] f(\varepsilon) d\varepsilon.$$

Therefore, the net surplus of a network-i consumer is given by

$$w_i = \alpha_i U(p_i, r_i) + \alpha_j U(p_i, r_j) + \alpha_i \tilde{U}(p_i, r_i) + \alpha_j \tilde{U}(p_j, r_i) - p_i \left[\alpha_i D(p_i, r_i) + \alpha_j D(p_i, r_j)\right] - r_i \left[\alpha_i D(p_i, r_i) + \alpha_j D(p_j, r_i)\right] - F_i.$$

And the profit of network i is given by

$$\pi_i \equiv \alpha_i \Big\{\alpha_i (p_i - c) D(p_i, r_i) + \alpha_j \left[p_i - c - (a - c_0)\right] D(p_i, r_j) + \alpha_j (a - c_0) D(p_j, r_i) + r_i \left[\alpha_i D(p_i, r_i) + \alpha_j D(p_j, r_i)\right] + F_i - f\Big\}.$$

We now analyze the first-order conditions. Given market share α_i, the first-order derivative of π_i with respect to p_i is given by

$$\alpha_i \left[1 - F(r_i - \beta p_i)\right] \alpha_i \left[u' + \tilde{u}' + E(\varepsilon \mid \varepsilon \geq r_i - \beta p_i) - c\right] q' + \alpha_i \left[1 - F(r_j - \beta p_i)\right] \left[\alpha_j (u' - c - a + c_0) + \alpha_i \left(r_j - \tilde{u}' - E(\varepsilon \mid \varepsilon \geq r_j - \beta p_i)\right)\right] q'.$$

Intuitively, consider a small decrease in p_i. This increases the volume of on-net calls by $[1 - F(r_i - \beta p_i)]|q'|$ and the volume of off-net calls by $[1 - F(r_j - \beta p_i)]|q'|$. In the market for on-net calls, network-i consumers' utility increases by $u' + \tilde{u}' + E(\varepsilon \mid \cdot)$, which the network can extract by increasing the fixed tariff F_i. In the market for off-net calls, network-i consumers' utility increases by u' and network-j consumers' utility increases by $\tilde{u}' + E(\varepsilon \mid \cdot) - r_j$. As before, $\tilde{u}' + E(\varepsilon \mid \cdot)$ represents the direct externalities and r_j represents the pecuniary externalities on consumers of network j. And an increase in network-j consumers' utility requires a decrease in F_i to keep α_i constant. The reader will check that the strategic-marginal-cost pricing formula of Proposition 1 holds as the noise vanishes and the caller determines volume with probability (close to) one.

When $r = r_i = r_j$, the first-order derivative simplifies to

$$\alpha_i \left[1 - F(r - \beta p_i)\right] \left[p_i - c - \alpha_j (a - c_0) + \alpha_i r\right] q',$$

which gives pricing at strategic marginal cost:

$$p_i = c + \alpha_j (a - c_0) - \alpha_i r.$$

© RAND 2004.

Similarly, the first-order derivative of π_i with respect to r_i is given by[18]

$$\alpha_i F(r_i - \beta p_i)\alpha_i E\left[\left(u' + \tilde{u}' + \varepsilon - c\right) q' \mid \varepsilon \leq r_i - \beta p_i\right] \frac{1}{\beta}$$
$$+ \alpha_i F(r_i - \beta p_j) E\left[\alpha_j(\tilde{u}' + \varepsilon + a - c_0)q' + \alpha_i(p_j - u')q' \mid \varepsilon \leq r_i - \beta p_j\right] \frac{1}{\beta}.$$

When $p = p_i = p_j$, the first-order derivative simplifies to

$$\alpha_i F(r_i - \beta p) E\left[(r_i - \alpha_i c + \alpha_j(a - c_0) + \alpha_i p)q' \mid \varepsilon \leq r_i - \beta p_j\right] \frac{1}{\beta},$$

which again gives pricing at strategic marginal cost:

$$r_i = \alpha_i c - \alpha_j(a - c_0) - \alpha_i p.$$

Consider now a symmetric equilibrium with $\alpha_i = 1/2$. From the two first-order conditions, we have

$$p = c + \frac{a - c_0 - r}{2}, \quad \text{and} \quad r = \frac{c - (a - c_0) - p}{2}.$$

These two conditions yield $p = c + (a - c_0)$ and $r = c_0 - a$. Furthermore, we show below that as the noise vanishes (the distribution F converges to a spike at value zero while keeping a wide enough support to confer an incentive role upon the reception charges),[19] the equilibrium in which the volume is determined by callers exists if the access charge markup is larger than $-[\beta/(1 + \beta)]c$. This analysis is summarized in the following proposition.

Proposition 6 (off-net-cost pricing). Suppose that the reception charges are noncooperatively set by the networks and that the marginal utility of call reception is random.

(i) There exists at most one symmetric equilibrium. For this equilibrium, the reception charge is equal to the access charge discount:

$$r = c_0 - a;$$

and the calling charge is

$$p = c + (a - c_0).$$

That is, the networks price calls and call receptions at their off-net cost.

(ii) Furthermore, when $a - c_0 \geq -[\beta/(1 + \beta)]c$ holds, as the noise converges to zero, the candidate equilibrium is an equilibrium, and in this equilibrium volume is determined by callers with probability converging to one.

(iii) As noise converges to zero, the social optimum can be approximated by choosing an

[18] Consider a small decrease in r_i. This will increase the volume of on-net calls by $F(r_i - \beta p_i)E(|q'| \mid \cdot)/\beta$ and the volume of off-net calls received from network j by $F(r_i - \beta p_j)E(|q'| \mid \cdot)/\beta$. In the market for on-net calls, network-i consumers' utility increases by $E(u' + \tilde{u}' + \varepsilon \mid \cdot)$. In the market for off-net calls, network-i consumers' utility increases by $E(\tilde{u}' + \varepsilon \mid \cdot)$ and network-j consumers' utility increases by $E(u' - p_j \mid \cdot)$. u' represents the direct externalities and p_j represents the pecuniary externalities on consumers of network j.

[19] We consider a family of distributions (F_n, f_n) with identical support and satisfying, for all $\varepsilon_0 > 0$,

$$\lim_{n \to \infty} \Pr(|\varepsilon| \leq \varepsilon_0) = 1.$$

© RAND 2004.

access charge a^* such that $a^* - c_0 = -r^* \equiv -\beta c/(1+\beta)$, and so the caller and receiver demand the same length of communications given the off-net-cost prices charged by their network(s).

Proof of (ii) and (iii). See our discussion paper, Jeon, Laffont, and Tirole (2001).

By contrast, efficiency cannot be achieved with a single instrument in the presence of (nonvanishing) noise. To see this, note that $u'(q) \geq p = c + (a - c_0)$ and $\tilde{u}'(q) + \varepsilon \geq r = c_0 - a$ with almost surely one strict inequality. Thus, $u'(q) + \tilde{u}'(q) + \varepsilon > c$, and so the sum of the marginal utilities always exceeds c: there is always *underprovision* of communications.[20]

□ **Incomplete coverage.** Our derivation of the off-net-cost pricing principle relies on the full-market-coverage assumption. More generally, if α_i denotes the fraction of potential consumers who subscribe to network i ($\alpha_1 + \alpha_2 \leq 1$ and α_i is an increasing function of w_i and $w_i - w_j$), we assumed that α_i was a function of $w_i - w_j$ only (which is indeed the case with fixed market size).

With variable market size, the technique used in this article still applies (maximization with respect to p_i keeping α_i constant). However, the expressions are changed if a change in p_i at equilibrium changes w_i (and therefore α_i). But a change in price p_i accompanied by a change in F_i that keeps $w_i - w_j$ constant also keeps w_i constant (and therefore α_i constant) if it keeps w_j constant, i.e., if (in the absence of noise)

$$\frac{dw_j}{dp_i} = \alpha_i \left[\tilde{u}'(q(p_i)) - r_j\right] \frac{dq}{dp_i} = 0.$$

Thus, heuristically (the exact proof is available upon request from the authors), the off-net-cost pricing principle still holds if

$$\tilde{u}'(q(p_i)) = r_j$$

at the off-net-cost prices. This (and the symmetric condition for the reception charge) holds if the access charge is equal to the socially optimal access charge a^*.

Proposition 7 (incomplete coverage). Suppose that the coverage is incomplete and the networks compete between themselves and with a fixed outside option. Then if the termination charge is equal to a^* (such that $a^* - c_0 = -\beta c/(1+\beta)$), as the noise vanishes, then in a symmetric equilibrium (which necessarily is unique), networks price calls and call receptions at their off-net cost.

For nonoptimal access charges, marginal prices on one side (calling or receiving) create externalities on the other side, and the opportunity cost reasoning underlying the off-net-cost pricing principle must be amended to account for the competition with outside options. Suppose for example that $a > a^*$, and so at the off-net prices, callers keep conversations too short from the point of view of receivers:

$$u'(q(c + (a - c_0))) - (c + (a - c_0)) = 0 < \tilde{u}'(q(c + (a - c_0))) - (c_0 - a).$$

A small reduction in p_i compensated, as is now familiar, by an increase in the monthly subscriber charge so as to keep the relative attractiveness $w_i - w_j$ constant, does nothing to attract (or lose) consumers from (to) network j, but it increases the absolute attractiveness w_i of network i (it increases the absolute attractiveness w_j of network j and, since F_i is adjusted so as to keep $w_i - w_j$ constant, increases w_i as well), and so helps network i attract new consumers who were not consuming telecommunications services.

[20] Efficiency could not be achieved with more instruments either: the difference in marginal utilities is random, and so price instruments that are not contingent on the realization of this difference cannot equalize the two marginal utilities (which is necessary for efficiency).

© RAND 2004.

5. Network-based discrimination

■ This section studies competition in two-part tariffs in the presence of network-based discrimination. Network i's tariff is characterized now by a five-tuple $\{F_i, p_i, \hat{p}_i, r_i, \hat{r}_i\}$, where hats refer to off-net communications. Again, we can distinguish the case in which the reception charges $(r_i, \hat{r}_i)$ are regulated from the case in which they are not.

We show below that network-based discrimination is a mixed blessing. While it induces network i to choose the on-net price p_i so as to induce callers to fully internalize the externalities on its receivers, it also allows networks to implement *selective* connectivity breakdown by charging very high or even infinite prices ($\hat{p}_i$ or $\hat{r}_i$) for off-net calls, which results in a de facto connectivity breakdown.

□ **Heuristics.** The reason why connectivity is prone to break down under network-based price discrimination is that profit opportunities are driven by the networks' packages' *relative* standing.[21] We identify two different reasons why network competition results in connectivity breakdown.

Absence of reception charges. We show below that each network's equilibrium off-net calling charge tends to infinity as the receiver's utility converges toward the caller's (that is, β converges to one). The intuition for this result is that a receiver on the rival network (who, recall, does not pay any reception charge) fully enjoys her surplus from the call. In contrast, the caller-network pair perceives only the net surplus (caller surplus minus calling cost). And so, off-net calls make the rival network relatively more attractive for β large, which leads to a connectivity breakdown.

Let us generalize the strategic-marginal-cost formula in order to sharpen this intuition:

$$\alpha_j \hat{p}_i = \alpha_j \left[c + (a - c_0)\right] - \alpha_i \left[-\tilde{u}'\left(q_{ij}\right)\right]. \tag{14}$$

To understand (14), consider a unit increase in the call volume $q_{ij} \equiv q(\hat{p}_i)$ from network i to network j. Viewed from network i, this increase in call volume concerns only a fraction α_j of subscribers (whereas it concerned the full fraction, 1, of subscribers in the absence of price discrimination: see (10)). Network i's margin on an off-net minute is $[\hat{p}_i - (c + (a - c_0))]$. Because $\hat{r}_j = 0$, this extra minute generates an extra net surplus $\tilde{u}'(q_{ij})$ for network j's subscribers that must be offset by an equal reduction in network i's fixed fee F_i to keep market share constant.

In a symmetric equilibrium, and using $u'(q_{ij}) = \hat{p}_i$, (14) becomes

$$u'\left(q_{ij}\right) = c + (a - c_0) + \tilde{u}'\left(q_{ij}\right). \tag{15}$$

Condition (15) formalizes our earlier intuition that the sender's network internalizes in its pricing decision the difference in marginal surpluses as well as its own marginal cost (but not the receiver's network's marginal cost). Provided that the off-net cost is positive, it cannot hold if the receiver's marginal utility is close to or larger than the sender's. Hence, $\hat{p}_i$ goes to infinity and q_{ij} to zero.[22]

Market-determined reception charges. The introduction of reception charges should *a priori* reduce the previous incentive for connectivity breakdown by lowering the sender's network's strategic marginal cost. Reception charges, however, provide a second instrument for implementing selective connectivity breakdown: Each network can induce the receiver to hang up off-net

[21] This unfettered competition is exacerbated by our assumption that the market is covered and so networks at the margin compete entirely among themselves and not against outside opportunities.

[22] For example, if $\tilde{u}(q) = \beta u(q)$, then

$$\hat{p}_i = \frac{(1 - \alpha_i)(c + a - c_0)}{1 - (1 + \beta)\alpha_i} = \frac{c + a - c_0}{1 - \beta}$$

for symmetric market shares. So $\hat{p}_i = \infty$ for $\beta \geq 1$ and $\hat{p}_i \to \infty$ as β converges to one from below.

© RAND 2004.

calls. We show that any symmetric equilibrium exhibits de facto connectivity breakdown.[23] To avoid price indeterminacy, we assume, as in Section 4, that the receiver's marginal utility is subject to noise: $\tilde{u}'(q) + \varepsilon$. This enables us to determine both calling and reception charges.

As earlier, we are primarily interested in the situation in which the noise becomes vanishingly small (while keeping a large support so as to avoid price indeterminacy). We impose a mild regularity condition.

Definition. A sequence of distributions $F_n(\varepsilon)$ of the random variable ε with mean zero is regular if

$$\lim_{n\to\infty} E_n\left(\varepsilon \mid \varepsilon \geq \varepsilon_0\right) = \varepsilon_0 \qquad \text{for } \varepsilon_0 > 0$$

and

$$\lim_{n\to\infty} E_n\left(\varepsilon \mid \varepsilon \leq \varepsilon_0\right) = \varepsilon_0 \qquad \text{for } \varepsilon_0 < 0.$$

For example, normal distributions with variance σ_n^2 going to zero are regular. Any single-peaked distribution with exploding hazard rates $(f_n(\varepsilon)/(1 - F_n(\varepsilon)) \to \infty$ for $\varepsilon > 0$ and $f_n(\varepsilon)/F_n(\varepsilon) \to \infty$ for $\varepsilon < 0)$ is also regular.

Consider for example off-net calls originating on network i and the case where most of the volume is determined by callers $\left(\hat{r}_j < \beta \hat{p}_i\right)$. Receivers determine the volume only when the noise ε takes low values. Therefore, the reception charge $\hat{r}_j$ is determined by the strategic-marginal-cost formula, which takes the form of an expectation conditional on the receivers determining the volume. When the distribution of ε is regular, the events in which the volume is determined by receivers can be approximated by the event in which both callers and receivers determine the volume ($\varepsilon = \hat{r}_j - \beta \hat{p}_i$). Then, in a symmetric equilibrium, the formula can be simplified as follows:

$$\left[\hat{r}_j - (c_0 - a)\right] + \left[\hat{p}_i - u'\left(q_{ij}\right)\right] = 0. \tag{16}$$

The first bracketed term represents the markup with respect to the off-net cost and the second bracketed term represents the difference between the cost and the marginal utility that an incremental increase in the volume inflicts on network i's consumers. Since $\hat{p}_i = u'(q_{ij})$ holds when both callers and receivers determine the volume at the same time, the two externalities have the same value. Therefore, $\hat{r}_j$ is equal to $(c_0 - a)$ and the off-net-cost pricing principle applies.

Turning now to the choice of $\hat{p}_i$ by network i, note that the externality on network-j receivers of an extra unit of phone calls is almost surely $\tilde{u}'(q'_{ij}) - \hat{r}_j$ as the noise vanishes. Hence, $\hat{p}_i$ is given by the following strategic-marginal-cost formula:

$$[\hat{p}_i - (c + a - c_0)] + \left[\hat{r}_j - \tilde{u}'\left(q_{ij}\right)\right] = 0. \tag{17}$$

After replacing $\hat{p}_i$ with $u'(q_{ij})$ and $\hat{r}_j$ with $(c_0 - a)$, we obtain

$$u'\left(q_{ij}\right) = c + 2(a - c_0) + \tilde{u}'\left(q_{ij}\right),$$

which is similar to (15). Therefore, provided that $c + 2(a - c_0) > 0$, connectivity breakdown occurs for $\beta \geq 1$ as in the absence of reception charge: the pecuniary externalities created by $\hat{r}_j$ are too small to counterbalance the direct externalities.

Remark. The two reasons why connectivity might break down are of a different nature than that studied by Wright (2002) and Gans and King (2000). These articles analyze interconnection charges between mobile networks and a fixed-link network. Each mobile network sets the termination charge to be paid by the fixed-link network, and there is no competition between fixed and mobile telephony. Each mobile network can impose a markup on termination, for which it has

[23] By de facto connectivity breakdown we mean that the volume-determining price(s) goes or is equal to infinity.

© RAND 2004.

monopoly power regardless of its size. This incentive to set high termination charges is magnified if off-net calls originating on the fixed-link network are charged the same price regardless of the identity of the terminating network (because of regulation or because consumer ignorance[24] makes it hard to price discriminate according to the terminating network), as the termination charge increase goes entirely to the mobile network and its impact on traffic is borne by all mobile networks. Indeed, Wright shows how a mobile network's high access charge raises its rivals' cost by lowering their termination revenue when the fixed network must practice uniform pricing. Gans and King also study competition between two mobile networks and show that connectivity breakdowns occur because each network's profit is determined by its standing relative to the other network. Connectivity breakdown in Wright and Gans and King is linked to the ability to tax *another* network's consumers through an access charge. Our connectivity breakdowns occur even for a fixed access charge and involve networks taxing *their own* consumers through high usage prices.

□ **Connectivity breakdown in the absence of reception charge.** In this subsection and the next, we assume for technical simplicity that $\tilde{u}(q) = \beta u(q)$. As the heuristics above suggest, the results do not hinge on this assumption, but they are simpler to state (we will here take β converging to one or $\beta \geq 1$) and to prove under it. This subsection shows that in the absence of a reception charge, the equilibrium price for off-net calls $\hat{p}$ goes to infinity as β goes to one. Furthermore, for $\beta \geq 1$, we show that network competition always results in a complete connectivity breakdown ($\hat{p} = \infty$).

In the absence of a reception charge, the volume is automatically determined by callers. The net surplus of a network-i consumer is given by

$$w_i = \alpha_i v(p_i) + \alpha_j v(\hat{p}_i) + \alpha_i \tilde{u}(q(p_i)) + \alpha_j \tilde{u}(q(\hat{p}_j)) - F_i.$$

Network i's market share is given by

$$\alpha_i = \frac{1}{2} + \sigma(w_i - w_j).$$

Equivalently,

$$\alpha_i = \frac{1}{2} + \sigma \left[\alpha_i v(p_i) + \alpha_j v(\hat{p}_i) - \alpha_j v(p_j) - \alpha_i v(\hat{p}_j) + \alpha_i \tilde{u}(q(p_i)) + \alpha_j \tilde{u}(q(\hat{p}_j)) - \alpha_j \tilde{u}(q(p_j)) - \alpha_i \tilde{u}(q(\hat{p}_i)) - F_i + F_j\right].$$

Network i's profit is given by

$$\pi_i = \alpha_i \left\{(p_i - c)\alpha_i q(p_i) + (\hat{p}_i - c - (a - c_0))\alpha_j q(\hat{p}_i) + (a - c_0)\alpha_j q(\hat{p}_j) + F_i - f\right\}.$$

As in the no-discrimination case, we will perform our analysis in two steps. First, we maximize π_i with respect to p_i and $\hat{p}_i$ keeping market share α_i constant. Second, we perform the maximization with respect to the market share.

For all $\alpha_i > 0$, the profit-maximizing price p_i is equal to the social-welfare-maximizing price p^*:

$$p_i^{**}(\alpha_i) = p^* = c - \tilde{u}'(q(p^*)). \quad (18)$$

For on-net calls, network i fully internalizes the externalities on receivers. Since network i is a monopoly in the market for on-net calls, under a two-part tariff, it maximizes the size of the pie. Hence, both networks choose the same price p^* regardless of the market shares.

[24] "Consumer ignorance" refers to the fact that when a consumer makes a fixed-to-mobile call, he is unaware of the identity of the mobile network to which the call is made.

© RAND 2004.

The first-order derivative of profit with respect to $\hat{p}_i$ is given by[25]

$$\alpha_i \left\{\alpha_j \left[\hat{p}_i - c - (a - c_0)\right] - \alpha_i \tilde{u}'(q(\hat{p}_i))\right\} \frac{dq}{d\hat{p}_i}, \qquad \text{for } \hat{p}_i > 0.$$

Since $\tilde{u}'(q(\hat{p}_i)) = \beta \hat{p}_i$, the optimal price $\hat{p}_i^{**}$ depends upon the market share as follows:

$$\hat{p}_i^{**}(\alpha_i) = \begin{cases} \dfrac{(1 - \alpha_i)(c + a - c_0)}{1 - (1 + \beta)\alpha_i} & \text{if } \alpha_i < \dfrac{1}{1 + \beta}, \\ \infty & \text{otherwise,} \end{cases}$$

where we assume that the off-net cost is positive ($c + a - c_0 > 0$).

Proposition 8. In the absence of reception charges, and if a symmetric equilibrium with network-based price discrimination exists,

(i) the price for on-net calls is socially optimal: $p = p^*$.

(ii) Connectivity breakdown:

(a) For $0 \leq \beta < 1$, network competition results in asymptotic connectivity breakdown in that the price for off-net calls goes to infinity as β tends to one:

$$\hat{p} = \frac{c + a - c_0}{1 - \beta}.$$

(b) For $\beta \geq 1$, any symmetric equilibrium exhibits connectivity breakdown: $\hat{p} = \infty$.

Remark. Our working paper (Jeon, Laffont, and Tirole 2001) studies the existence of equilibrium for a constant-elasticity demand function. The second-order derivative for the program of maximizing the profit Π_i with respect to α_i is negative if σ is small enough and $a \simeq c_0$. These are sufficient conditions for a symmetric equilibrium to exist.

□ **Connectivity breakdown with reception charges.** In this subsection we examine how the introduction of reception charges affects connectivity. As discussed earlier, reception charges are a mixed blessing in the context of network-based price discrimination. On the one hand, a positive $\hat{r}_j$, through the pecuniary externalities, reduces the marginal cost perceived by network i, which induces it to internalize externalities on network-j receivers. On the other hand, $\hat{r}_j$ is set strategically by network j, whose private incentive conflicts with social welfare maximization: the gain that consumers of network i derive from placing calls to consumers of network j increases the marginal cost perceived by network j.

Suppose that, as in Section 4, the receiver's utility is subject to a regular noise ε and given by

$$\tilde{u}(q) + \varepsilon q.$$

The net surplus of a network-i consumer is given by

$$\begin{aligned} w_i = {} & \alpha_i U(p_i, r_i) + \alpha_j U(\hat{p}_i, \hat{r}_j) + \alpha_i \tilde{U}(p_i, r_i) + \alpha_j \tilde{U}(\hat{p}_j, \hat{r}_i) \\ & - p_i \left[\alpha_i D(p_i, r_i) + \alpha_j D(\hat{p}_i, \hat{r}_j)\right] - r_i \left[\alpha_i D(p_i, r_i) + \alpha_j D(\hat{p}_j, \hat{r}_i)\right] - F_i, \end{aligned}$$

[25] The first term represents the direct impact on profit and the second term represents the indirect impact arising from the condition of keeping market share constant. The second represents the direct externalities on consumers of network j. Consider an incremental increase in the volume of off-net calls initiated on network i. Then, network-j consumers' utility from reception increases by $\tilde{u}'$. When market share α_i is kept constant, an increase in the utility obtained by network-j consumers implies an increase in the marginal cost perceived by network i.

© RAND 2004.

where $U(\cdot)$, $\tilde{U}(\cdot)$, and $D(\cdot)$ are defined in Section 4. The profit of network i is given by

$$\pi_i \equiv \alpha_i \Big\{ \alpha_i (p_i + r_i - c) D(p_i, r_i) + \alpha_j \left[\hat{p}_i - c - (a - c_0)\right] D(\hat{p}_i, \hat{r}_j) + \alpha_j (\hat{r}_i + a - c_0) D(\hat{p}_j, \hat{r}_i) + F_i - f \Big\}.$$

As in the absence of a reception charge, it is optimal for network i to maximize the size of the pie in the market for on-net calls. In fact, given any market share α_i, the first-order condition with respect to p_i and r_i is given by

$$\frac{d\left[U(p_i, r_i) + \tilde{U}(p_i, r_i) - cD(p_i, r_i)\right]}{dp_i} = 0 = \frac{d\left[U(p_i, r_i) + \tilde{U}(p_i, r_i) - cD(p_i, r_i)\right]}{dr_i}.$$

Consider now the off-net calls initiated on network i. Let D_{ij} denote $D(\hat{p}_i, \hat{r}_j)$: U_{ij} and $\tilde{U}_{ij}$ are similarly defined. For expositional convenience, we introduce the following notation:

$$\pi_i^{\hat{p}}(\hat{p}_i : \alpha_i, \hat{r}_j) \equiv \alpha_i \left\{ \alpha_j \left[U_{ij} - (c + a - c_0) D_{ij}\right] + \alpha_i \left[\hat{r}_j D_{ij} - \tilde{U}_{ij}\right] \right\},$$
$$\pi_j^{\hat{r}}(\hat{r}_j : \alpha_i, \hat{p}_i) \equiv \alpha_j \left\{ \alpha_i \left[\tilde{U}_{ij} + (a - c_0) D_{ij}\right] + \alpha_j \left[\hat{p}_i D_{ij} - U_{ij}\right] \right\}.$$

$\pi_i^{\hat{p}}(\hat{p}_i)$ (respectively, $\pi_j^{\hat{r}}(\hat{r}_j)$) represents the share of π_i (respectively, π_j) that can be affected by $\hat{p}_i$ (respectively, $\hat{r}_j$) when α_i is kept constant. We note that $\pi_i^{\hat{p}}(\infty) = \pi_j^{\hat{r}}(\infty) = 0$. Therefore, each network can obtain at least zero net profit in the market for the off-net calls initiated on network i by implementing selective connectivity breakdown. In particular, if $\pi_j^{\hat{r}}(\hat{r}_j) < 0$ for $\hat{r}_j < \infty$, network j implements connectivity breakdown by setting $\hat{r}_j = \infty$.

From now on, we focus on symmetric equilibria. The first-order derivative of $\pi_i^{\hat{p}}(\hat{p}_i)$ with respect to $\hat{p}_i$ keeping market share constant ($\alpha_i = 1/2$) is given by

$$\frac{1}{4}\left[\frac{\partial U(\hat{p}_i, \hat{r}_j)}{\partial \hat{p}_i} - \left[(c + a - c_0) - \hat{r}_j\right]\frac{\partial D(\hat{p}_i, \hat{r}_j)}{\partial \hat{p}_i} + \frac{\partial \tilde{U}(\hat{p}_i, \hat{r}_j)}{\partial \hat{p}_i}\right],$$

which is equivalently written as follows:

$$\frac{1}{4}\left[1 - F(\hat{r}_j - \beta\hat{p}_i)\right]\left\{\left[\hat{p}_i - (c + a - c_0)\right] + \hat{r}_j - \tilde{u}' - E\left[\varepsilon \mid \varepsilon \geq \hat{r}_j - \beta\hat{p}_i\right]\right\} q'(\hat{p}_i). \qquad (19)$$

The first term in $\{\cdot\}$ represents the markup with respect to the off-net marginal cost and the second term is composed of the pecuniary externalities and the direct externalities on the consumers of the rival network. The direct externalities include the average noise conditional on callers determining the volume. Similarly, the first-order derivative of $\pi_j^{\hat{r}}$ with respect to $\hat{r}_j$ keeping market share constant ($\alpha_j = \frac{1}{2}$) is given by

$$\frac{1}{4} F(\hat{r}_j - \beta\hat{p}_i) E\left\{\left[\left(\hat{r}_j - (c_0 - a)\right) + \left(\hat{p}_i - u'\right)\right] q'\left(\frac{\hat{r}_j - \varepsilon}{\beta}\right) \mid \varepsilon < \hat{r}_j - \beta\hat{p}_i\right\}. \qquad (20)$$

In the next proposition, we ignore the total connectivity breakdown equilibrium based on "weakly dominated strategies" (this equilibrium always exists because $\hat{p}_i = \infty$ is a best response to $\hat{r}_j = \infty$ and vice versa) and show that any symmetric equilibrium exhibits asymptotic or complete connectivity breakdown.

Proposition 9 (connectivity breakdown). Let ε be regularly distributed. As the noise vanishes,

(i) for $0 < \beta < 1$: If $a - c_0 > -\beta c/(1 + \beta)$, the only symmetric candidate equilibrium

© RAND 2004.

without complete connectivity breakdown satisfies

$$\hat{p} = \hat{p}^* \equiv \frac{c + 2(a - c_0)}{1 - \beta}, \hat{r} = c_0 - a,$$

where $\beta \hat{p} > \hat{r}$ holds. If $a - c_0 \leq -\beta c/(1 + \beta)$, any symmetric equilibrium exhibits complete connectivity breakdown, with $\hat{r} = \infty$.

(ii) For $\beta > 1$: If $a - c_0 < -\beta c/(1+\beta)$, the only symmetric candidate equilibrium without connectivity breakdown is given by

$$\hat{r} = \hat{r}^* \equiv \frac{[-2(a - c_0) - c]\beta}{\beta - 1}, \hat{p} = c + a - c_0,$$

where $\beta \hat{p} < \hat{r}$ holds. If $a - c_0 \geq -\beta c/(1 + \beta)$, any symmetric equilibrium exhibits complete connectivity breakdown, with $\hat{p} = \infty$.

Proof. See the Appendix.

□ **Regulation of reception charges.** We just saw that network-based price discrimination often allows each network to implement *selective* connectivity breakdown, that is, breakdown of connectivity in one direction. This selective connectivity breakdown results in a two-way lack of connectivity. This observation calls for some form of regulation (broadly defined), in the same way termination charges cannot just be left to the discretion of the terminating networks. This "regulation" can take the form of a cooperatively determined off-net reception charge. Alternatively, the off-net reception charge may be set by a regulatory agency. Our discussion paper (Jeon, Laffont, and Tirole, 2001) considers a specific regulatory scheme that ties a network's reception charge to the sending network's off-net call charge in the following way:

$$\hat{r}_j = \begin{cases} g(\underline{p}), & \text{if } \hat{p}_i < \underline{p}, \\ g(\hat{p}_i), & \text{if } \hat{p}_i \geq \underline{p}, \end{cases}$$

with

$$g(\hat{p}_i) \equiv \beta \frac{\eta}{\eta - 1} \hat{p}_i - \varepsilon \hat{p}_i^{\eta} \quad \text{and} \quad 0 < \underline{p} < c + a - c_0,$$

where η is the elasticity of demand, which we will assume is constant, and ε is a positive constant such that $g(\underline{p}) = \beta \underline{p}$ and $g(\hat{p}_i) \leq \beta \hat{p}_i$ for all $\hat{p}_i \geq \underline{p}$. Hence, $\hat{r}_j$ is chosen indirectly by network i.

The regulation of the reception charge $\hat{r}_j$ has two main consequences. First, since network j cannot control $\hat{r}_j$, the regulation eliminates its strategic behavior regarding the choice of $\hat{r}_j$. In particular, network j's opportunity profit related to off-net calls from network i to network j can be negative (network j would prefer to set $\hat{r}_j = \infty$ if it could). Second, the regulation is designed so as to ensure that a change in network-i 's off-net caller charge ($\hat{p}_i$) does not affect the welfare of a receiver on network j. This specific scheme is information intensive, but it can be shown to have some nice properties. In particular, the symmetric equilibria are characterized by the following:

(i) $p = p^*$, $\hat{p} = c + (a - c_0)$.

(ii) If the termination charge is chosen equal to $c_0 - r^*$, the symmetric equilibrium maximizes social welfare.

(iii) There exists an equilibrium in which the firms maximize social welfare and obtain the monopoly profit.

© RAND 2004.

6. Conclusion

■ We provided a comprehensive overview of the main insights in the Introduction, so there is no need to reproduce it fully here. Suffice it to reiterate the key lines of the analysis:

(i) From a normative viewpoint, calling charges must lie below the marginal cost of communication when receivers value receiving calls. This "calling subsidy" can be obtained by setting the termination charge below the marginal cost of termination.

(ii) By lowering each network's "strategic marginal cost," reception charges also contribute to an internalization of the externality on receivers. The termination charge and the reception charge can be regulated in such a way that a symmetric equilibrium exists and is efficient.

(iii) When both calling and reception demands are elastic and reception charges are market determined, it is optimal for each operator to equate the prices for calling and receiving with their off-net costs. Consequently, the equilibrium reception charges decrease with the termination charge, which reinforces the encouragement provided by termination discounts to set low caller charges. For an appropriately chosen termination charge, the symmetric equilibrium is again efficient.

(iv) Last, network-based price discrimination creates strong incentives for connectivity breakdown, even among equal networks.

Because the issues studied here are central to the development of network industries exhibiting externalities between the various sides of the market, we hope that this article will stimulate further research extending the analysis in several important directions, such as alternative descriptions of call externalities. In this respect, our analysis of noncooperative volume setting should be extended to allow for more cooperative behavior. At the other extreme lie the maximization of joint surplus over (i) call length (i.e., the maximization of $\{u(q)+\tilde{u}(q)-(p_i+r_j)q\}$ when a caller on network i calls a receiver on network j), or (ii) callback strategies (maximization of joint surplus as in (i), together with the choice of calling direction so as to minimize $(p_i + r_j)$), or even (iii) *ex ante* coordination on network choices.[26]

Conversely, one could follow Hermalin and Katz (forthcoming) in analyzing other aspects in the absence of sender-receiver cooperation.[27] That article provides an interesting analysis of gaming in such a context. For example, when the identity of the sender is determined endogenously, a war of attrition may add further inefficiencies. Hermalin and Katz show that with symmetric surpluses, cost sharing through the receiver-pays principle dominates the caller-pays regime. Hermalin and Katz also consider the possibility that the sender chooses the allocation of cost c between prices p and r in a menu (as is done through 800 numbers).

Obviously, our article is only a first step in a rather broad research agenda.

Appendix

■ We now prove Proposition 9. The first-order derivative with respect to $\hat{p}_i$ (19) is equivalently written as follows:

$$\frac{1}{4}\left[1 - F(\hat{r}_j - \beta\hat{p}_i)\right]\left\{(1-\beta)\hat{p}_i + \hat{r}_j - (c + a - c_0) - E\left[\varepsilon \mid \varepsilon \geq \hat{r}_j - \beta\hat{p}_i\right]\right\} q'(\hat{p}_i). \tag{A1}$$

The first-order derivative with respect to $\hat{r}_j$ (20) is equivalently written as follows:

$$\frac{1}{4}F(\hat{r}_j - \beta\hat{p}_i)\left\{\left[\hat{p}_i + (1 - \frac{1}{\beta})\hat{r}_j - (c_0 - a) + \frac{\varepsilon}{\beta}\right] q'(\frac{\hat{r}_j - \varepsilon}{\beta}) \mid \varepsilon < \hat{r}_j - \beta\hat{p}_i\right\}. \tag{A2}$$

[26] This third possibility seems less interesting, as it may presuppose quasi-exclusive traffic between the two end users.

[27] Their article does not consider network competition.

© RAND 2004.

For simplicity, from now on we will not write $(1/4)[1 - F(\hat{r}_j - \beta\hat{p}_i)]$ or $(1/4)F(\hat{r}_j - \beta\hat{p}_i)$ in the first-order derivatives. After using the property that the noise distribution is regular (see the definition in the text) to (A1), we obtain the first-order derivative with respect to $\hat{p}_i$ as the noise converges to zero:

$$\begin{cases} \left[(1-\beta)\hat{p}_i + \hat{r}_j - (c + a - c_0)\right] q'(\hat{p}_i), & \text{if } \beta\hat{p}_i \geq \hat{r}_j, \\ \left[(1-\beta)\hat{p}_i + \hat{r}_j - (c + a - c_0) - (\hat{r}_j - \beta\hat{p}_i)\right] q'(\hat{p}_i), & \text{if } \beta\hat{p}_i \leq \hat{r}_j, \end{cases}$$

which is equivalent to

$$\begin{cases} \left[(1-\beta)\hat{p}_i + \hat{r}_j - (c + a - c_0)\right] q'(\hat{p}_i), & \text{if } \beta\hat{p}_i \geq \hat{r}_j, \\ \left[\hat{p}_i - (c + a - c_0)\right] q'(\hat{p}_i), & \text{if } \beta\hat{p}_i \leq \hat{r}_j. \end{cases} \tag{A3}$$

Similarly, as the noise converges to zero, the first-order derivative with respect to $\hat{r}_j$ is given by

$$\begin{cases} \left[\hat{r}_j + a - c_0\right] q'(\hat{p}_i) = 0, & \text{if } \beta\hat{p}_i \geq \hat{r}_j, \\ \left[\hat{p}_i - \dfrac{1-\beta}{\beta}\hat{r}_j + (a - c_0)\right] q'\left(\dfrac{\hat{r}_j}{\beta}\right) & \text{if } \beta\hat{p}_i \leq \hat{r}_j. \end{cases} \tag{A4}$$

In what follows, we distinguish three regimes depending on whether $\beta\hat{p}_i - \hat{r}_j$ is positive, negative, or zero in equilibrium. We here prove the result for $0 < \beta < 1$. The case $\beta > 1$ can be similarly proved. First, consider the case in which most of the volume is determined by callers in equilibrium: $\beta\hat{p}_i > \hat{r}_j$. Then, from the first-order conditions derived from (A3) and (A4), we have the unique equilibrium candidate:

$$\hat{p}_i = \hat{p}^* \equiv \frac{c + 2(a - c_0)}{1 - \beta}, \hat{r}_j = -(a - c_0).$$

Since $\beta\hat{p}^* > -(a - c_0)$ holds if and only if $a - c_0 > -\beta c/(1 + \beta)$ holds, the caller-determined volume case does not exist for $a - c_0 \leq -\beta c/(1 + \beta)$.

Second, consider the case in which most of the volume is determined by receivers in equilibrium: $\beta\hat{p}_i < \hat{r}_j$. Then, from (A3), we have $\hat{p}_i = c + a - c_0$. Inserting $\hat{p}_i = c + a - c_0$ into (A4), the first-order derivative with respect to $\hat{r}_j$ is given by

$$\left[c + a - c_0 - \frac{1-\beta}{\beta}\hat{r}_j + (a - c_0)\right] q'\left(\frac{\hat{r}_j}{\beta}\right).$$

Then, the first-order derivative is strictly negative for $\hat{r}_j < \hat{r}^* \equiv [c + 2(a - c_0)]\beta/(1 - \beta)$ and strictly positive for $\hat{r}_j > \hat{r}^*$. Therefore, the profit is maximized either for $\hat{r}_j = \beta(c + a - c_0)$ or $\hat{r}_j = \infty$. However, $\hat{r}_j = \beta(c + a - c_0) = \beta\hat{p}_i$ is incompatible with the case $\beta\hat{p}_i < \hat{r}_j$. Therefore, $\hat{r}_j$ must be equal to ∞, which leads to complete connectivity breakdown.

Last, consider the intermediate case: $\beta\hat{p}_i = \hat{r}_j$. After substituting $\beta\hat{p}_i = \hat{r}_j$ into the first-order conditions derived from (A3) and (A4), we have a unique candidate: $\hat{p}_i = c + a - c_0$ and $\hat{r}_j = -(a - c_0)$. The candidate satisfies $\beta\hat{p}_i = \hat{r}_j$ if and only if $a - c_0 = -\beta c/(1 + \beta)$. However, when $a - c_0 = -\beta c/(1 + \beta)$ holds, the first-order derivative with respect to $\hat{r}_j$ is strictly positive for $\hat{r}_j > -(a - c_0)$, implying that $\hat{r}_j$ must be equal to ∞. Therefore, there cannot exist any equilibrium with $\beta\hat{p}_i = \hat{r}_j < \infty$. *Q.E.D.*

References

Armstrong, M. "Network Interconnection." *Economic Journal*, Vol. 108 (1998), pp. 545–564.

———. "The Theory of Access Pricing and Interconnection." In M.E. Cave, S.K. Majumdar, and I. Vogelsang, eds., *Handbook of Telecommunications Economics*. Amsterdam: North-Holland, 2002.

Atkinson, J. and Barnekov, C. "A Competitively Neutral to Network Interconnection." OPPWP Series, FCC, 2000.

Berger, U. "Network Competition and the Collusive Role of the Access Charge." Mimeo, University of Vienna, 2001.

Carter, M. and Wright, J. "Interconnection in Network Industries." *Review of Industrial Organization*, Vol. 14 (1999), pp. 1–25.

——— and ———. "Local and Long-Distance Competition." Mimeo, Universities of Canterbury and Auckland, 2000.

Cherdron, M. "Interconnection, Termination-Based Price Discrimination, and Network Competition in a Mature Telecommunications Market." Mimeo, Universität Mannheim, 2000.

Cremer, J., Rey, P., and Tirole, J. "Connectivity in the Commercial Internet." *Journal of Industrial Economics*, Vol. 48 (2000), pp. 433–472.

DeGraba, P. "Efficient Interconnection Regimes for Competing Networks." Mimeo, U.S. Federal Communications Commission, 2000.

Dessein, W. "Network Competition in Nonlinear Pricing." *RAND Journal of Economics*, Vol. 34 (2003), pp. 593–611.

© RAND 2004.

———. "Network Competition with Heterogeneous Customers and Calling Patterns." *Information, Economics and Policy*, forthcoming.

Doyle, C. and Smith, J.C. "Market Structure in Mobile Telecoms: Qualified Indirect Access and the Receiver Pays Principle." *Information Economics and Policy*, Vol. 10 (1998), pp. 471–488.

Gans, J.S. and King, S.P. "Mobile Network Competition, Customer Ignorance and Fixed-to-Mobile Call Prices." *Information Economics and Policy*, Vol. 12 (2000), pp. 301–327.

——— and ———. "Using 'Bill and Keep' Interconnect Arrangements to Soften Network Competition." *Economics Letters*, Vol. 71 (2001), pp. 413–420.

Hahn, J.-H. "Network Competition and Interconnection with Heterogeneous Subscribers." Mimeo, Christ Church College, University of Oxford, 2000.

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

——— and ———. "Sender or Receiver: Who Should Pay to Exchange an Electronic Message?" *RAND Journal of Economics*, forthcoming.

Jeon, D.S., Laffont, J.-J., and Tirole, J. "On the Receiver Pays Principle." Mimeo, IDEI, University of Toulouse, 2001.

Katz, M. and Shapiro, C. "Network Externalities, Competition, and Compatibility." *American Economic Review*, Vol. 75 (1985), pp. 424–440.

Kim, J.-Y. and Lim, Y. "An Economic Analysis of the Receiver Pays Principle." *Information Economics and Policy*, Vol. 13 (2001), pp. 231–260.

Laffont, J.-J., Rey, P., and Tirole, J. "Network Competition I: Overview and Nondiscriminatory Pricing." *RAND Journal of Economics*, Vol. 29 (1998a), pp. 1–37.

———, ———, and ———. "Network Competition II: Price Discrimination." *RAND Journal of Economics,* Vol. 29 (1998b), pp. 38–56.

———, Marcus, S., Rey, P., and Tirole, J. "Internet Interconnection and the Off-Net-Cost Pricing Principle." *RAND Journal of Economics*, Vol. 34 (2003), pp. 370–390.

Littlechild, S. "The Role of Consumption Externalities in the Pricing of Telephone Service." in J.T. Wenders, ed. *Pricing in Regulated Industry: Theory and Application.* Denver, Colo.: Regulatory Matters Division, Mountain States Telephone and Telegraph Co., 1977.

Oftel. "Oftel's Submission to the Monopolies and Mergers Commission Inquiry into Prices of Calls to Mobile Phones." London: 1998.

Wright, J. "Access Pricing Under Competition: An Application to Cellular Networks." *Journal of Industrial Economics*, Vol. 50 (2002), pp. 289–315.

© RAND 2004.