

Internet interconnection and the off-net-cost pricing principle

Jean-Jacques Laffont*

Scott Marcus**

Patrick Rey***

and

Jean Tirole****

We develop a framework for Internet backbone competition. In the absence of direct payments between websites and consumers, the access charge allocates communication costs between websites and consumers and affects the volume of traffic. We analyze the impact of the access charge on competitive strategies in an unregulated retail environment. In a remarkably broad range of environments, operators set prices for their customers as if their customers' traffic were entirely off-net. We then compare the socially optimal access charge with the privately desirable one. Finally, when websites charge micropayments, or sell goods and services, the impact of the access charge on welfare is reduced; in particular, the access charge is neutral in a range of circumstances.

1. Introduction

■ Long an emanation of voluntarist public policies, the Internet has moved in recent years to a market paradigm. While still partly run on the basis of legacy agreements, the Internet industry is actively searching for a business model that will increase Internet usage and facilitate the evolution to enhanced offerings based on differentiated classes of services. A key feature of the Internet is that each computer connected to it can communicate with every other connected computer. In a deregulated environment, this universal connectivity can be achieved only if competing connectivity providers cooperatively reach agreements governing the price and quality of their interconnection.

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

** U.S. Federal Communications Commission; scott@scottmarcus.com.

*** University of Toulouse (IDEI and GREMAQ); prey@cict.fr.

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

We are grateful to Mike Riordan, Aaron Schiff, and Julian Wright, and to the participants of the IDEI conference on the Economics of the Software and Internet Industries (January 18–20, 2001) for helpful reactions on an earlier draft. Concerning Scott Marcus, the opinions expressed do not necessarily reflect the views of the FCC or any of its commissioners. We also thank the Editor, Jennifer Reinganum, and two referees for their comments.

The interconnection charges, also called “access charges,” “settlements,” or “termination charges,” could be vital for enabling efficient use of the Internet. Incentives must be provided for a widespread usage of bandwidth by dial-up, broadband, and dedicated access consumers, and for the posting of content by the websites. Quality-of-service (QoS) agreements between operators can reduce delays and packet losses for marked traffic and thereby enable the development of new and advanced Internet services such as IP telephony and videoconferencing. Competition for end users is a necessary condition for an efficient functioning of the industry, but it will fall short of accomplishing even its most modest goals in the absence of proper interconnection agreements.

The purpose of this article is to develop a framework for modelling the competition among interconnected Internet “backbone operators” or “networks.” In this framework, the “end users” or “customers” are heterogeneous in several respects. First, their patterns of traffic imbalance differ. Consumers receive much more traffic than they send, primarily due to the downloads they request; websites, in contrast, originate much of their traffic, even though they do not request it. Second, different end users generate different value to other end users and thus to the Internet. Third, end users may differ in the cost that their traffic imposes on the operators.

The backbone operators vie for the various types of traffic. In particular, each competes on the two sides of the market (consumers and websites). The competitive analysis offers two sets of insights:

Competitive strategies. On the positive side, we analyze pricing strategies in this interconnected environment. The first key insight of the article is that in a wide range of situations, backbones set their price on each business segment *as if* they had no other customer. That is, they set charges to consumers and websites as if their connections were entirely off-net. We call this the “off-net-cost pricing principle.” We first demonstrate this principle in the simplest perfectly competitive environment with a reciprocal access charge. This simple principle turns out to be remarkably robust to generalizations of the model: mixed traffic patterns, variable demand, QoS agreements, backbone differentiation, installed bases, multihoming, and customer cost heterogeneity.

Impact of the access charge on welfare and profit. The access charge affects the backbones’ marginal cost of incoming and outgoing off-net traffic. It therefore determines how backbones distribute communication costs between websites and consumers. *Ceteris paribus*, a higher access charge penalizes end users, such as websites, with an outgoing-traffic bias, and it benefits end users, such as consumers, with the opposite bias. Network externalities considerations, though, complicate end users’ preferences over access charges, as they want the other side of the market to expand.

We first consider the case where there is no direct payment between websites and consumers. This case is most relevant when there are no micropayments and no other financial transaction resulting from consumers’ visits to the websites. In that case, the access charge should promote economic efficiency by alleviating the burden on those end users (i) whose demand is highly elastic and (ii) who create value for other end users. More generally, we shall argue that the access charge cannot by itself induce all the price differentiation that would be required for an efficient allocation in the Internet. Furthermore, if backbones have market power, they do not necessarily choose the socially optimal access charge.

Also, individual end users’ elasticities will be affected by a more widespread use of micropayments between end users, which partly reallocate costs endogenously. Indeed, we consider more briefly the case where consumers pay a price to the websites for their visits (this price can be a micropayment charged by the website, or part of a transaction resulting from their visit). This financial transaction provides an additional channel for allocating the cost of the communication, which lowers the allocative impact of the access charge.

On the positive side, we analyze the access charge’s impact on profits. There may be no such impact, for example when an increase in the access charge is competed away by the backbones’ offering very low prices to consumers. If backbones have market power, however, profits are affected by the access charge, and backbones will tend to subsidize the more profitable segment.

The article is organized as follows. Section 2 constructs a model of perfect (Bertrand) backbone competition for consumers and websites, assuming that both sides of the market are supplied, i.e., demands are locally inelastic. Section 3 demonstrates the robustness of the off-net-cost pricing principle. Section 4 analyzes the socially optimal access charge. Section 5 discusses some limits of the off-net-cost pricing principle. Section 6 introduces micropayments between customers and websites. Section 7 concludes.

Our article is related to the literature on two-way access in telecommunications, e.g., Armstrong (1998) and Laffont, Rey, and Tirole (1998a 1998b).¹ This literature assumes that while consumers both send and receive traffic, receivers get no surplus from and are not charged for receiving calls. When instead receivers derive some utility from receiving calls, an externality must be internalized for efficiency. The fact that users are not charged for receiving traffic has several implications. First, operators' marginal charge for outgoing traffic is equal to the on-net cost augmented by the *average* termination markup rather than to the off-net cost. Second, it creates some instability in competition if the networks are close substitutes and the termination charge is not in the vicinity of the termination cost; in contrast, Section 2 establishes that no such instability occurs when consumers are charged for receiving calls.

The articles most related to ours are Jeon, Laffont, and Tirole (2001) and Hermalin and Katz (2001). Jeon, Laffont, and Tirole analyze the off-net cost pricing principle in a telecommunications environment where the volume of traffic between each sender and receiver is endogenously determined by the party with the lower marginal willingness to communicate. In particular, this formulation allows one to tell apart monthly (subscription) fees and usage fees (for receiving and sending traffic). That article also considers the case of regulated reception charges, and it stresses furthermore that network-based price discrimination is conducive to connectivity breakdowns. In contrast, in most of this article we suppose that there is a fixed volume of transactions for each consumer-website match. This makes nonlinear tariffs irrelevant (no distinction between fixed and usage fees); we use this simpler formulation to study several additional aspects, such as the impact of multihoming, market power, asymmetric access charges, and micropayments between consumers and websites. Hermalin and Katz also focus on fixed transactions but allow for stochastic (and possibly correlated) gains from communication. They show that double marginalization increases when networks specialize in offering services to senders or receivers and also study asymmetric Bertrand competition, where some operators are more efficient than others.

2. A simple benchmark

■ Although our theory allows for general traffic imbalances, it is useful for expository purposes to distinguish two types of customers: *websites* and *consumers*. Consumers exchange traffic (e.g., emails), browse web pages, download files, and so forth; websites post pages and files, which can be browsed and downloaded by consumers. There is little traffic between websites, and furthermore, the traffic between consumers (such as email exchanges) or from consumers to websites (the requests for pages or file downloads) is much smaller than the traffic from websites to consumers (the actual downloading of web pages and files). To capture this traffic pattern in its simplest form, we neglect the traffic between consumers or between websites, as well as the traffic from consumers to websites, and focus instead on the traffic from websites to consumers.

Most of the article uses the following assumptions:

Balanced calling pattern. We assume that consumers' interest in a website is unrelated to the website's network choice: a consumer is as likely to request a page from a given website belonging to her network and another given website belonging to a rival network.² In the absence of

¹ See also Carter and Wright (1999a, 1999b), Cherdrone (2000), Dessein (forthcoming), Gans and King (2001), and Hahn (2000).

² This assumption ought to be refined in specific instances. For example, regional or international specialization of backbones together with other factors, such as language affinity, may induce some violations of this hypothesis (Chinese consumers may be more likely to browse U.S. websites than U.S. customers to browse Chinese websites).

origination-based price discrimination (that is, if a consumer pays the same price for receiving traffic, regardless of the identity of the originating website's backbone), the percentage of traffic originating on network i and completed on network j is therefore proportional both to the fraction of websites on network j and to the fraction of consumers subscribing to network i .

Reciprocal access pricing. We assume that there is no asymmetry in the interconnection charge: A network pays as much for having its traffic terminated on a rival network ("off-net traffic") as it receives for terminating traffic originating on a rival network. This assumption will be relaxed in Section 5, but it is worth noting that there have been calls for regulators to impose reciprocal access charges.³ (At the moment, most interconnection agreements between the top-level backbones take the form of "bill and keep" peering agreements, with zero (and thus reciprocal) termination charges; however, this situation is likely to evolve in the future—some backbones have already introduced positive termination charges in their agreements with certain other backbones.)

Let us now be more specific about the model:

Cost structure. Two full-coverage⁴ "networks," or "backbones" or "operators," have the same cost structure. For notational simplicity, we ignore request traffic, so that the only costs are those incurred to bring traffic from websites to consumers. We also do not include any fixed network cost. It is straightforward to add both types of costs.⁵

We let c denote the total marginal cost of traffic. When traffic is handed over from one backbone to the other, we let c_o and c_t denote the originating and terminating backbones' marginal costs associated with this traffic ($c_o + c_t = c$).

Although the exact expressions of c_o and c_t are irrelevant for the theory, it is useful for concreteness to discuss the nature of these costs in the current Internet environment. For example, suppose that backbones incur a marginal cost c' per unit of traffic at the originating and terminating ends and a marginal cost c'' in between, which may stand for the routing costs and the marginal cost of trunk lines used for transportation. The total marginal cost of traffic is thus

$$c \equiv 2c' + c''.$$

In practice, top-level backbone operators have multiple interconnection points and have an incentive to pass on off-net traffic as soon as possible. A consequence of this "hot-potato" pattern⁶ is that most of the transportation cost c'' is borne by the receiving backbone.⁷ For off-net traffic, the sending network thus incurs the marginal cost of origination, c' , while the receiving network incurs both the transportation cost c'' and the marginal cost of termination, c' . The total marginal cost of traffic is thus shared by the sending and receiving networks according to

$$c_o \equiv c' \quad \text{and} \quad c_t \equiv c' + c''.$$

Demand structure. We first assume that the networks are perfect substitutes: that consumers and websites have inelastic demand for and supply of web pages. To be sure, consumers and websites

³ See Marcus (1999) and Gao (2000) for overviews of the Internet's hierarchical organization.

⁴ "Full coverage" means that the backbones have a global geographical presence and thus are able to serve all customers.

⁵ The next section considers mixed traffic patterns. For simplicity, we also ignore the impact on the cost structure of caching, replication, and other content-delivery network schemes.

⁶ For a description of hot-potato routing, see Marcus (1999).

⁷ Our analysis would, however, apply to any other method of sharing the cost of off-net traffic. We assume here that the access charge is, as is currently the case, independent of the "distance" between the point at which the traffic is handed over and the location of the receiver. Our analysis would still apply if there were differentiated access charges, as long as differences in access charges reflected differences in termination costs. The white paper NRIC (2002) provides a detailed overview of current interconnection agreements and the issues they raise.

are more likely to use the web if they are charged lower prices; we will thus relax these assumptions later on.

There is a continuum of consumers, of mass 1, and a continuum of websites, of mass 1 as well. Each consumer generates one unit of traffic from each website connected to either backbone. Each unit of traffic from a website to a consumer yields a value v to the consumer and a value \tilde{v} to the website. We will assume that the market is viable, that is,

$$v + \tilde{v} > c.$$

Until Section 6, we assume away “micropayments” between consumers and websites and posit that websites do not charge differentiated prices to consumers depending on whether their connection is on- or off-net. Furthermore, backbones are perfect substitutes on both sides of the market, and so each side chooses the lowest price that it is offered.

We will initially assume that prices are low enough that all consumers or websites connect to a backbone. The volume of traffic associated with each customer is then fixed, and there is thus no point distinguishing between subscription and usage prices or linear and nonlinear prices: consumers’ subscription decisions are based on the prices p_1 and p_2 charged by the two backbones for receiving traffic, while websites’ subscription decisions are based on the prices \tilde{p}_1 and \tilde{p}_2 charged for sending traffic.⁸ Note that the backbones need not be able to tell consumers and websites apart directly. It suffices that inflows and outflows be priced differently.

Denoting by α_i backbone i ’s market share for consumers and by $\tilde{\alpha}_i$ its market share for websites, and assuming that the two operators charge each other the same interconnection charge a for terminating traffic, backbone i ’s profit is given by (for $i \neq j = 1, 2$):

$$\pi_i = \alpha_i \tilde{\alpha}_i (p_i + \tilde{p}_i - c) + \alpha_i \tilde{\alpha}_j (p_i - (c_i - a)) + \alpha_j \tilde{\alpha}_i (\tilde{p}_i - (c_o + a))$$

or

$$\pi_i = \alpha_i \tilde{\alpha} [p_i - (c_i - a)] + \tilde{\alpha}_i \alpha [\tilde{p}_i - (c_o + a)], \quad (1)$$

where $\alpha = \alpha_1 + \alpha_2$ and $\tilde{\alpha} = \tilde{\alpha}_1 + \tilde{\alpha}_2$ denote, respectively, the total numbers of connected consumers and of connected websites. If all potential customers are connected as we assume in this section (that is, $\alpha = \tilde{\alpha} = 1$), this expression reduces to

$$\pi_i = \alpha_i [p_i - (c_i - a)] + \tilde{\alpha}_i [\tilde{p}_i - (c_o + a)]. \quad (2)$$

That is, as long as prices do not exceed customers’ reservation values, the profit of each backbone can be decomposed into two independent components: one for the consumer business, and another one for the website business. The perfect-substitutability assumption ensures furthermore that, in each line of business, all customers go to the cheapest operator whenever their prices differ.

The timing is as follows: (1) the access charge a is determined (through a bilateral agreement or by regulation), (2) the backbones set their prices, and (3) end users select their backbones. As is usual, we solve for a subgame-perfect equilibrium of the game.

Proposition 1 (off-net-cost pricing principle). Assume $v \geq c_i - a$ and $\tilde{v} \geq c_o + a$; then, there exists a unique price equilibrium.⁹ This equilibrium is symmetric and satisfies

$$\begin{aligned} p_1 &= p_2 = p^* = c_i - a, \\ \tilde{p}_1 &= \tilde{p}_2 = \tilde{p}^* = c_o + a, \end{aligned}$$

⁸ The consumer prices p_1 and p_2 can be indifferently interpreted as volume-based prices for receiving traffic, or as subscription prices—if the total number of websites were not normalized to 1, these would be subscription prices per website reached. Similarly, websites’ prices \tilde{p}_1 and \tilde{p}_2 can be interpreted as subscription prices (per consumer reached).

⁹ Market shares are undetermined.

$$\pi_1 = \pi_2 = \pi^* = 0.$$

Proof. The standard Bertrand argument applies to each business segment. The only caveat is that the number of connected customers in one segment affects the market demand in the other segment; however, as long as prices remain below reservation values, all customers are connected (to one or the other network) and, in each segment, the market demand is thus independent of the actual price levels. *Q.E.D.*

For each customer, the price is competitively set equal to the opportunity cost of servicing this customer, rather than letting the customer subscribe to the other network. Suppose for example that backbone 1 “steals” a consumer away from backbone 2. Then, the traffic from backbone 2’s websites to that consumer, which was previously internal to backbone 2, now costs backbone 1 an amount c_t to terminate but generates a marginal termination revenue a ; the opportunity cost of that traffic is thus $c_t - a$. And the traffic from backbone 1’s websites, which initially costs c_o for origination and a for termination on backbone 2, is now internal to backbone 1 and thus costs $c = c_o + c_t$; therefore, for that traffic too, the opportunity cost of stealing the consumer away from its rival is $c - (c_o + a) = c_t - a$. A similar reasoning shows that stealing a website away from the rival backbone generates, for each connected consumer, a net cost $c_o + a$: attracting a website increases originating traffic, which costs c_o , and also means sending more traffic from its own websites to the other backbone’s end users, as well as receiving less traffic from the other backbone (since the traffic originated by the stolen backbone is now on-net); in both cases, a termination revenue a is lost.

In this very simple benchmark case of perfectly substitutable networks and inelastic demand, Bertrand-like competition ensures that profits are set at their competitive level ($\pi^* = 0$); whatever the access charge a , the combined per-unit charge to consumers and websites covers the cost of the traffic:¹⁰

$$\tilde{p}^* + p^* = (c_o + a) + (c_t - a) = c_o + c_t = c.$$

The access charge a thus merely determines how the cost of the traffic is shared between senders (websites) and receivers (consumers)—with a higher access charge leading to a larger burden being placed on the websites. In particular, the access charge has no impact on network profits and on social welfare, defined as the sum of customers’ surpluses, which is equal to its first-best level:

$$\begin{aligned} W &= \sum_i \alpha_i \tilde{\alpha} (v - p_i) + \sum_i \tilde{\alpha}_i \alpha (\tilde{v} - \tilde{p}_i) + \sum_i \pi_i \\ &= W^{FB} \equiv v + \tilde{v} - c. \end{aligned}$$

Finally, let us compare Proposition 1 with the results in Laffont, Rey, and Tirole (1998a) and Armstrong (1998) for interconnection of telephone networks. A key difference with this telecommunications literature is that in the latter there is a missing price: receivers do not pay for receiving calls; that is, in the notation of this article, $p = 0$. The missing price has two important implications:

Pricing. The operators’ optimal usage price reflects their *perceived* marginal cost. But when operators do not charge their customers (here, consumers) for the traffic they receive, operator i ’s perceived marginal cost of outgoing (here, website) traffic is given by

$$c + \alpha_j (a - c_t). \tag{3}$$

That is, the unit cost of traffic is the on-net cost c , augmented by the expected off-net “markup”

¹⁰ This holds as long as customers’ prices remain lower than customers’ reservation values, that is, as long as $c_o + a \leq \tilde{v}$ and $c_t - a \leq v$. If for example $c_o + a > \tilde{v}$, the maximal price that can be charged to websites, $\tilde{p} = \tilde{v}$, does not cover the opportunity cost they generate, $c_o + a$. Thus, no backbone wants to host a website and there is then no traffic at all for such an access charge.

(or discount) $(a - c_t)$ on the fraction α_j of website traffic that terminates off-net. Comparing the two perceived marginal costs of outgoing traffic with and without receiver charge, for given access charge and market shares, the price for sending traffic is higher (lower) than in the presence of reception charges if and only if there is a termination discount (markup).¹¹

Note that if the “missing payment” $\alpha_i p_i$ were subtracted from the right-hand side¹² of (3) and p_i were equal to the off-net cost¹³ $(c_t - a)$, then (3) would be equal to the off-net cost $(c_o + a)$. In sum, the missing payment affects the backbones’ perceived costs, and it reallocates costs between origination and reception.

Stability in competition. When networks are close substitutes, and receivers are not charged, there exists no equilibrium unless the access charge is near the termination cost. The intuition is easily grasped from (3). If there is a substantial termination tax or subsidy, perceived marginal costs (and thus prices) are far from actual costs, thereby introducing a source of inefficiency. But if networks are sufficiently close substitutes, either operator could corner the market with a small reduction in its price, in which case it faces the true costs and can offer a better deal. This issue does not arise when end users pay (or are paid) for receiving traffic. In that case, the sum of the perceived costs for origination and termination always equals the actual cost of communication: $(c_o + a) + (c_t - a) = c$, irrespective of the access charge.

3. Robustness of the off-net-cost pricing principle

■ The off-net-cost pricing principle is robust to various extensions of the perfectly competitive model.

- (i) *Arbitrary number of backbones.* The principle extends trivially to n backbones ($n \geq 2$): it suffices to replace “ α_j ” in equation (1) by “ $\sum_{j \neq i} \alpha_j$ ”.
- (ii) *Mixed traffic patterns.* We have caricatured reality by assuming that websites have only outgoing traffic, and consumers only incoming traffic. All Internet users in fact have a mixed, although often very biased, pattern. It is easily verified that under perfect competition, backbones ask their customers (consumers or websites) to pay

$$T_i(x, y) = (c_t - a)x + (c_o + a)y,$$

where x and y are the customer’s incoming and outgoing traffic volumes.

- (iii) *Multihoming.* Suppose now that each website may choose to locate in both backbones. Websites do not gain or lose from multihoming as long as the backbones charge the competitive tariff $\tilde{p}^* = c_o + a$.¹⁴
- (iv) *Quality of service (QoS).* Proposition 1 extends to multiple qualities of service, as long as costs and access charge refer to the quality of service in question.
- (v) *Customer cost heterogeneity.* Our assumption that all customers impose the same cost on the backbone for incoming or outgoing traffic is more restrictive than needed. Suppose that there are K types of customers, $k = 1, \dots, K$. A customer of type k , whether a consumer or a website, imposes cost c_o^k at origination and c_t^k at termination.¹⁵ The off-net-cost pricing principle still holds as long as backbones can price discriminate.

¹¹ Indeed, $c + \alpha_j(a - c_t) > c_o + a$ is equivalent to $(1 - \alpha_j)(a - c_t) < 0$.

¹² To reflect the fact that the traffic generated by backbone i ’s websites brings reception revenue for the share α_i of the traffic that remains on-net.

¹³ If consumers do not derive any utility from receiving calls ($v = 0$), as in Laffont, Rey, and Tirole (1998a), the price p_i cannot be positive; networks could, however, subsidize receivers.

¹⁴ In practice, however, websites may gain from enhanced reliability or redundancy, at the cost of diseconomies of scale in the interface with the backbones.

¹⁵ For example, European or Australian Internet service providers must be connected to U.S. backbones through costly transoceanic cables that raise both origination and termination costs relative to a U.S.-based customer.

In practice, this cost-based price discrimination may be implemented by setting different charges for local delivery; alternatively, it can be implemented by uniform charges applied at given points of interconnection, together with the requirement of the provision by the end users (or their ISPs) of circuits leading to these points of interconnection.

- (vi) *Installed bases.* Suppose that backbone i has an installed base $\hat{\alpha}_i$ of consumers and an installed base $\hat{\alpha}_i$ of websites that are, for example, engaged in long-term contracts. Let \hat{p}_i and $\hat{\tilde{p}}_i$ denote the predetermined prices charged to installed base consumers and websites by network i . The operators' profits become

$$\pi_i = \alpha_i [p_i - (c_t - a)] + \tilde{\alpha}_i [\tilde{p}_i - (c_o + a)] + \hat{\alpha}_i [\hat{p}_i - (c_t - a)] + \hat{\alpha}_i [\hat{\tilde{p}}_i - (c_o + a)].$$

Consequently, the equilibrium prices are unchanged: new customers are charged the off-net-cost prices and operator i 's equilibrium profit $\pi_i^*(a)$ is equal to

$$\frac{d\pi_i^*}{da} = \hat{\alpha}_i - \hat{\tilde{\alpha}}_i.$$

Two simple implications can be drawn from this observation. First, web-hosting backbones prefer a low termination charge, while backbones that are stronger on the dial-up side, say, prefer a high termination charge. Second, if the termination charge is determined in a private negotiation, two backbones tend to have conflicting interests if one leans much more heavily to one side of the market than does the other. However, their interests do not necessarily conflict (even if one carries far more traffic than the other) if, say, one segment of the market has (for both backbones) developed more quickly than the other segment.

4. Ramsey access charges

■ By focusing on inelastic demands, the benchmark model of Section 2 and the various extensions performed in Section 3 sidestepped welfare issues. This section maintains the perfect-competition assumption but allows for elastic demands. Perfect competition implies that backbones' budgets are always balanced, whatever the access charge. But through its allocation of costs between end users, the access charge plays a central role in achieving economic efficiency. We show below that the Ramsey access charges, i.e., the access charges that maximize social welfare, must take into account not only the demand elasticities of the two segments, but also the externality that each side exerts on the other.¹⁶

Suppose for example that a consumer derives surplus v , drawn from a distribution $F(v)$, from being connected with websites; similarly, a website derives a surplus \tilde{v} , drawn from a distribution $\tilde{F}(\tilde{v})$, from being connected with consumers. Consumers' and websites' demands are thus given by $q = D(p) = 1 - F(p)$ and $\tilde{q} = \tilde{D}(\tilde{p}) = 1 - \tilde{F}(\tilde{p})$.¹⁷ Furthermore, consumers' and websites' net surpluses are given by $S(p) = \int_p^{+\infty} (v - p)dF(v)$ and $\tilde{S}(\tilde{p}) = \int_{\tilde{p}}^{+\infty} (\tilde{v} - \tilde{p})d\tilde{F}(\tilde{v})$. Then,

¹⁶ Similar conclusions hold for the credit card industry, in which the "backbones" are "banks," the "websites" and "consumers" are the "merchants" and "cardholders," and the "access charge" is the "interchange fee." See Rochet and Tirole (2002), Schmalensee (2002), Schwartz and Vincent (2000), and Wright (2000). Related insights apply to B2B—see Caillaud and Jullien (2001).

¹⁷ There again, prices can be interpreted as pure traffic-based prices or as subscription prices (per website reached or per consumer reached). For example, if backbones simply charge a subscription price T for receiving traffic, the relevant consumer price is $p = T/\tilde{D}$, where \tilde{D} denotes the number of connected websites, and a consumer with a valuation v subscribes again if $p \leq v$.

Proposition 2. When both consumers' and websites' demands are elastic, the Lindahl (first-best) prices are given by

$$p^{FB} + \tilde{p}^{FB} = c - \frac{S(p^{FB})}{D(p^{FB})} = c - \frac{\tilde{S}(\tilde{p}^{FB})}{\tilde{D}(\tilde{p}^{FB})},$$

whereas the Ramsey (second-best) prices and access charge are characterized by $p^{SB} = c_t - a^{SB}$, $\tilde{p}^{SB} = c_o + a^{SB}$, and

$$\frac{S(p^{SB})}{D'(p^{SB})} = \frac{\tilde{S}(\tilde{p}^{SB})}{\tilde{D}'(\tilde{p}^{SB})}. \tag{4}$$

Proof. Social welfare is equal to

$$W = S(p) \tilde{D}(\tilde{p}) + D(p) \tilde{S}(\tilde{p}) + (p + \tilde{p} - c) D(p) \tilde{D}(\tilde{p});$$

its first-best level is thus characterized by

$$p + \tilde{p} = c - \frac{S(p)}{D(p)} = c - \frac{\tilde{S}(\tilde{p})}{\tilde{D}(\tilde{p})}.$$

Ramsey prices maximize W subject to the budget constraint

$$(p + \tilde{p} - c) D(p) \tilde{D}(\tilde{p}) \geq 0.$$

Denoting by λ the Lagrangian multiplier associated with this budget constraint, and using $p + \tilde{p} = c$, the first-order conditions boil down to

$$-\lambda = D'(p) \tilde{S}(\tilde{p}) = \tilde{D}'(\tilde{p}) S(p).$$

Q.E.D.

From a first-best perspective, each segment is charged a price equal to the marginal cost, minus a discount that reflects the positive externality exerted on the other segment. For example, an extra website generates an additional gross consumer surplus $S + pD$, so that the (per-consumer) price \tilde{p} charged to websites must be decreased by an amount equal to the (per-capita, or average) consumer surplus $v^e \equiv p + S/D$:

$$\tilde{p} = c - v^e.$$

Similarly, the (per-website) price charged to consumers must be discounted for the average surplus $\tilde{v}^e \equiv \tilde{p} + \tilde{S}/\tilde{D}$ that consumers bring to websites: $p = c - \tilde{v}^e$. Since average surpluses exceed prices ($v^e > p$, $\tilde{v}^e > \tilde{p}$), the total price charged to the two segments, $p + \tilde{p}$, must be lower than the cost c ; the subsidy must reflect the positive externality that each segment exerts on the other:

$$c - (p + \tilde{p}) = \frac{S(p)}{D(p)} = \frac{\tilde{S}(\tilde{p})}{\tilde{D}(\tilde{p})},$$

which in particular implies that, at the optimum, these two externalities must be equalized.

In a second-best world, the budget constraint rules out outside subsidies. Prices must therefore be increased so as to fully cover the cost c , according to standard Ramsey principles: the departure from first-best prices should be inversely related to the magnitude of demand elasticities:

$$\frac{p - (c - v^e)}{p} = \frac{\lambda}{\eta}, \quad \frac{\tilde{p} - (c - v^e)}{\tilde{p}} = \frac{\lambda}{\tilde{\eta}},$$

where η and $\tilde{\eta}$ denote the demand elasticities and λ the Lagrangian multiplier associated with the budget constraint. In the absence of fixed cost, the budget constraint is simply that the total price $p + \tilde{p}$ must cover the joint cost c and the above Ramsey formulas boil down to (4), which can be interpreted as follows. Increasing the consumer price discourages some consumers, which reduces website surplus; the corresponding welfare loss is thus given by $D'(p)\tilde{S}(\tilde{p})$. Similarly, increasing the website price discourages some websites, which reduces consumer surplus, thereby generating a welfare loss $\tilde{D}'(\tilde{p})S(p)$. The optimal tradeoff thus depends on how many end users are discouraged on one side, as well as on the net surplus lost on the other side, and balances the two types of welfare losses: $D'(p)\tilde{S}(\tilde{p}) = \tilde{D}'(\tilde{p})S(p)$. A special case occurs when one side of the market is inelastic as in Section 2; then, the access charge shifts the burden as much as possible to the inelastic segment.

Remark. In sharp contrast with the recommendations usually derived from standard Ramsey pricing formulas, the tradeoff just described can lead, in equilibrium, to a higher price for the segment with the higher elasticity. To see this, note that condition (4) can be rewritten as (letting $\eta_S = -pS'/S$ and $\tilde{\eta}_{\tilde{S}} = -\tilde{p}\tilde{S}'/\tilde{S}$)

$$\left(\frac{p}{\tilde{p}}\right)^2 = \frac{\eta \eta_S}{\tilde{\eta} \tilde{\eta}_{\tilde{S}}}.$$

That is, prices in the two segments should covary with their respective demand elasticities (η or $\tilde{\eta}$) (and with the related surplus elasticities, η_S and $\tilde{\eta}_{\tilde{S}}$).

Under perfect competition, firms make zero profit; they are thus indifferent as to the level of the access charge and should not resist a regulation of the access charge that implements the second-best optimum. In practice, backbones have historically opted for “bill and keep” ($a = 0$), which minimizes transaction costs. Bill and keep is favorable to websites,¹⁸ which might have been a good idea to promote the development of Internet-based services. Now that many web services are available, and the emphasis is more on encouraging consumers to connect and use these services, absent significant transaction costs, bill and keep is unlikely to be close to optimal.

5. Amending the off-net-cost pricing principle

■ **Variable demand and two-part tariffs.** Let us extend the model to allow for variable demand functions and connection costs for consumers—sticking to the same formulation as before for websites. It is then natural to also allow backbones to charge two-part tariffs to consumers. Because of the connection costs, the off-net costs no longer predict average retail prices; however, they still define the relevant marginal usage prices if backbones compete in nonlinear tariffs. To see this, for $i = 1, 2$, let p_i denote the volume-based fee and F_i the fixed fee charged by backbone i , and $D(p_i)$ the demand of a representative consumer who subscribes, with $S(p_i)$ the associated net surplus (but gross of the fixed fee F_i). A consumer thus subscribes to backbone i if

$$S(p_i) - F_i > S(p_j) - F_j.$$

Backbone i 's profit is then given by

$$\begin{aligned} \pi_i &= \alpha_i (F_i - f) + \alpha_i \tilde{\alpha}_i D(p_i) (p_i + \tilde{p}_i - c) + \alpha_i \tilde{\alpha}_j D(p_j) (p_i - c_i + a) \\ &\quad + \alpha_j \tilde{\alpha}_i D(p_j) (\tilde{p}_i - c_o - a) \\ &= \alpha_i (F_i - f) + \alpha_i (\tilde{\alpha}_1 + \tilde{\alpha}_2) D(p_i) [p_i - (c_i - a)] \\ &\quad + \tilde{\alpha}_i [\alpha_1 D(p_1) + \alpha_2 D(p_2)] [\tilde{p}_i - (c_o + a)]. \end{aligned}$$

¹⁸ When $a = 0$, consumers pay the entire termination cost, which, as noted above, is in practice the larger part of the cost due to “hot-potato” routing.

The opportunity cost of stealing a website away from the rival network is

$$\alpha_i D(p_i) [c - (c_t - a)] + \alpha_j D(p_j) [(c_o + a) - 0] = (c_o + a) q,$$

where $q = \alpha_1 D(p_1) + \alpha_2 D(p_2)$ denotes the volume of traffic generated by each website. The opportunity cost of stealing a website, per unit of traffic, is thus again $c_o + a$; therefore, in equilibrium, $\tilde{p}_1 = \tilde{p}_2 = c_o + a$.

Also, if $p_1 = p_2 = p$, then the opportunity cost of stealing a consumer away from the rival network is

$$\tilde{\alpha}_i D(p) [c - (c_o + a)] + \tilde{\alpha}_j D(p) [(c_t - a) - 0] + f - F_i = (c_t - a) D(p) + f - F_i;$$

furthermore, if $\tilde{p}_i = c_o + a$, then the opportunity cost of inducing its own consumers to generate one more unit of traffic is similarly given by

$$\tilde{\alpha}_i D(p) [c - \tilde{p}_i] + \tilde{\alpha}_j D(p) [(c_t - a) - 0] = (c_t - a) D(p).$$

Therefore, the off-net-cost pricing principle still applies, although now *only* in equilibrium. We thus have the following:

Proposition 3. When $\tilde{v} \geq c_o + a$ and $S(c_t - a) \geq f$, there exists a unique two-part-tariff equilibrium, given by

$$\begin{aligned} p_1 &= p_2 = p^* = c_t - a, \\ \tilde{p}_1 &= \tilde{p}_2 = \tilde{p}^* = c_o + a, \\ F_1 &= F_2 = f, \\ \pi_1 &= \pi_2 = \pi^* = 0. \end{aligned}$$

The off-net-cost pricing principle therefore still applies: in equilibrium, the fixed fee is equal to the connection cost, and usage prices for sending and receiving traffic are equal to the off-net costs of outgoing and incoming traffic.

□ **Market power.** Section 3 has demonstrated the remarkable robustness of the off-net-cost pricing principle in a competitive industry. We now investigate how the principle must be amended if the backbones have some market power, which supposes that they provide differentiated services. Intuitively, the relevant marginal cost remains the off-net-cost, but a markup should be added because of market power. We will say that an access charge is “constrained Ramsey optimal” if it is optimal when the only regulatory instrument is the access charge (that is, when market power cannot be directly addressed).

Let us maintain the assumption that the backbones are perfectly substitutable on the consumer segment but introduce some differentiation on the website segment. Websites’ surplus from being connected with consumers then depends on the network to which they subscribe.

Backbones can engineer their networks to provide different levels of quality, in several dimensions. They can, for example, ensure that capacity is significantly higher than offered load,¹⁹ and can do so with a focus on either the mean or the variance of delay, which may be of some interest for different users.²⁰ Backbones can also invest in reliability.²¹ Higher speed, lower

¹⁹ By spending more to build out more capacity, but also by doing a better job of predicting demand or by using caching or replication to reduce the amount of traffic they haul around the network. The straightness of the fiber runs and the number of router hops can also play a role.

²⁰ For example, real-time voice requires a low average delay and low variability, whereas email users might not care much about variability.

²¹ E.g., by ensuring redundancy. Simplicity can also contribute to reliability, as well as good connectivity to other networks.

delay variability, and higher reliability have an impact on the benefit that a website can derive from connecting to alternative backbones, and this impact depends on the website’s business model.

A different kind of engineering enhances reliability. Ensuring redundancy, with no single point of failure, can be important. Simplicity can also contribute to reliability, but paradoxically is often somewhat at odds with extensive redundancy. Good connectivity to other networks can also be a differentiator. This tends to play a larger role in periods where, for one reason or another, connectivity between networks is in general less than ideal.

These reliability and quality differences have an impact on the benefit that a website can derive from connecting to the alternative backbones, and this impact depends on the website’s business model. Letting \tilde{v}_1 and \tilde{v}_2 denote the surpluses that a website derives when subscribing respectively to backbone 1 or 2, the website subscribes to network i if (for $i \neq j = 1, 2$)²²

$$\tilde{v}_i - \tilde{p}_i \geq \max \{0, \tilde{v}_j - \tilde{p}_j\}.$$

The values \tilde{v}_1 and \tilde{v}_2 are distributed among websites according to $\tilde{F}(\tilde{v}_1, \tilde{v}_2)$ on \mathfrak{R}_+^2 , which determines the number $\tilde{D}_i(\tilde{p}_i, \tilde{p}_j)$ of websites choosing to subscribe to network i ; by construction, an increase in one operator’s price increases the demand for the other operator but reduces the total number of connected websites: $\partial \tilde{D}_j / \partial \tilde{p}_i > 0$ and $\partial \tilde{D}_i / \partial \tilde{p}_i + \partial \tilde{D}_j / \partial \tilde{p}_i < 0$. We will furthermore maintain the following assumptions:

- (i) $\tilde{D}_1(\tilde{p}_1, \tilde{p}_2) = \tilde{D}_2(\tilde{p}_2, \tilde{p}_1)$. Therefore $\tilde{D}(\tilde{p}) \equiv \tilde{D}_i(\tilde{p}, \tilde{p}) = [1 - \tilde{G}(\tilde{p})]/2$, where $\tilde{G}(\tilde{v})$ is the cumulative distribution of $\tilde{v} \equiv \max\{v_1, v_2\}$, and decreases when \tilde{p} increases: $\tilde{D}' < 0$.
- (ii) For every \tilde{c} there exists a unique price $\tilde{p}(\tilde{c})$ such that $\tilde{p} = \arg \max_{\tilde{p}'} (\tilde{p}' - \tilde{c}) \tilde{D}_i(\tilde{p}', \tilde{p})$ and $\tilde{\pi}(\tilde{c}) \equiv [\tilde{p}(\tilde{c}) - \tilde{c}] \tilde{D}[\tilde{p}(\tilde{c})]$ decreases when \tilde{c} increases: $\tilde{\pi}' < 0$.

The price \tilde{p} and profit $\tilde{\pi}$ can be interpreted as benchmark equilibrium price and (per-backbone) profit in a hypothetical economy where the two backbones face demands \tilde{D}_1 and \tilde{D}_2 and compete with the same cost \tilde{c} . The assumptions on the demand and profits are plausible: an increase in either backbone’s price reduces the number of connected websites, and an increase in the industry marginal cost reduces equilibrium profits.²³

Inelastic consumer demand. Suppose first that all consumers get the same value v from being connected to a website. Then, normalizing to 1 the population of consumers, we have the following:

Proposition 4. Assume that consumer demand is inelastic and $v \geq c_t - a$.²⁴ Then there exists a unique equilibrium. This equilibrium is symmetric and satisfies

$$p_i = p^* = c_t - a, \quad \tilde{p}_i = \tilde{p}(c_o + a), \quad \pi_i = \tilde{\pi}(c_o + a).$$

Increasing the access charge raises the equilibrium website price but reduces both the number of connected websites and the equilibrium profits. Backbones favor the lowest admissible access charge, $a_\Pi = c_t - v$, which fully extracts the surplus from consumers and subsidizes the profitable website segment; consumer demand being inelastic, this access charge is constrained Ramsey optimal.

²² As before, we assume that each consumer gets the same surplus from connecting to every website and, similarly, that each website gets the same surplus from connecting to every consumer. This allows for a simpler analysis of consumers’ and websites’ subscription decisions—and in particular eliminates any interdependence between these decisions.

²³ This condition is, for example, always satisfied when backbones absorb part of the cost increase ($0 < \tilde{p}' < 1$), since then both the demand and the margin decrease when the cost increases.

²⁴ For $v < c_t - a$, there is no symmetric equilibrium in pure strategies—but an asymmetric equilibrium may exist, in which one backbone serves all consumers at a price $p = v$.

Proof. Denoting by α_i backbone i 's share of consumers, the profit of that backbone is given by

$$\pi_i = \alpha_i (\tilde{D}_1 + \tilde{D}_2) [p_i - (c_i - a)] + \tilde{D}_i [\tilde{p}_i - (c_o + a)].$$

The standard Bertrand argument thus applies to the inelastic consumer segment. As long as it remains below v , the price paid by consumers does not affect their demand and thus has no effect on the website business. Therefore, $p_1 = p_2 = c_i - a$, which by assumption does not exceed v , and network i 's profit boils down to

$$\pi_i = \tilde{D}_i (\tilde{p}_i, \tilde{p}_j) [\tilde{p}_i - (c_o + a)],$$

which leads at a symmetric equilibrium to $\tilde{p}_i = \tilde{p}(c_o + a)$ and $\pi_i = \tilde{\pi}(c_o + a)$. *Q.E.D.*

Differentiation weakens the intensity of price competition for websites and allows backbones to earn a positive profit on that segment; website prices are thus higher than the off-net cost $c_o + a$, but prices and profits are “as if” backbones were competing for websites with a cost equal to their off-net cost. A lower access charge increases the opportunity cost of servicing consumers, leading the networks to raise the consumer price. On the other side, the reduction in the access charge lowers the opportunity cost of servicing the websites, which generates more profit on websites. In other words, decreasing the access charge allows the networks to extract more rents from the surplus that consumers derive from receiving traffic; part of those rents are passed on to websites, which benefit from a reduction in the price they are charged for sending traffic to consumers, and part of it serves to increase networks' profits.

Backbones thus favor the lowest access charge a , subject to consumers' participation constraint ($v \geq c_i - a$). Since consumer demand is inelastic, backbones' interest is here in line with the social interest, which also requires extracting consumer rents in order to subsidize and attract more websites.

Elastic consumer demand. Suppose now that consumers' surplus v is drawn from a distribution $F(v)$. Consumers' demand is thus $D(p) = 1 - F(p)$. The following assumption guarantees that profit functions are “well behaved” and rules out the possibility that a backbone might desire making losses on consumers in order to enhance profits on the website segment.

Assumption 1. There exists $k > 0$ such that

- (i) $\forall \tilde{c}, \quad \tilde{p}(\tilde{c}) - \tilde{c} < k;$
- (ii) $\forall p, \quad D(p) + kD'(p) \geq 0.$

Assumption 1 ensures that prices and profits are again “as if” backbones were (imperfectly) competing on each segment with a cost equal to their off-net cost:

Proposition 5. With an elastic consumer demand, there exists a unique symmetric equilibrium under Assumption 1. This equilibrium satisfies

$$p_i = p^* = c_i - a, \quad \tilde{p}_i = \tilde{p}(c_o + a), \quad \pi_i = \tilde{\pi}(c_o + a) D(c_i - a).$$

An increase in the access charge raises the equilibrium website price and thus reduces the number of connected websites, but it decreases the equilibrium consumer price and thus attracts more consumers.

Proof. Fix $p_j = p = c_i - a$ and $\tilde{p}_j = \tilde{p} = \tilde{p}(c_o + a)$. By raising its price p_i above $p = c_i - a$, backbone i gets $[\tilde{p}_i - (c_o + a)]\tilde{D}_i(\tilde{p}_i, \tilde{p})D(p)$ and thus cannot earn more than (and actually earns exactly) $\tilde{\pi}(c_o + a)D(c_i - a)$. By reducing its price p_i below $c_i - a$, backbone i 's profit is of the form $\pi_i(p_i, \tilde{p}_i) = D(p_i)\hat{\pi}_i(p_i, \tilde{p}_i)$, where

$$\hat{\pi}_i(p_i, \tilde{p}_i) = [p_i - (c_i - a)] [\tilde{D}_i(\tilde{p}_i, \tilde{p}) + \tilde{D}_j(\tilde{p}, \tilde{p}_i)] + [\tilde{p}_i - (c_o + a)] \tilde{D}_i(\tilde{p}_i, \tilde{p})$$

is maximal for some $\tilde{p}_i(p_i) > \tilde{p}$: the loss on the consumer segment gives an incentive to reduce the traffic and thus the number of connected websites, $\tilde{D}_i + \tilde{D}_j$. Using the envelope theorem, the impact of p_i on backbone i 's maximal profit is given by

$$\begin{aligned} \frac{d\pi_i(p_i, \tilde{p}_i(p_i))}{dp_i} &= D(p_i) [\tilde{D}_i(\tilde{p}_i(p_i), \tilde{p}) + \tilde{D}_j(\tilde{p}, \tilde{p}_i(p_i))] + D'(p_i) \hat{\pi}_i(p_i, \tilde{p}_i(p_i)) \\ &\geq D(p_i) \tilde{D}(\tilde{p}) + D'(p_i) [\tilde{p} - (c_o + a)] \tilde{D}(\tilde{p}), \end{aligned}$$

where the inequality stems from (using $\tilde{p}_i \geq \tilde{p}$)

$$\tilde{D}_i(\tilde{p}_i, \tilde{p}) + \tilde{D}_j(\tilde{p}, \tilde{p}_i) \geq \tilde{D}_j(\tilde{p}, \tilde{p}_i) \geq \tilde{D}_j(\tilde{p}, \tilde{p}) = \tilde{D}(\tilde{p})$$

and (using $p_i \leq p$)

$$\hat{\pi}_i(p_i, \tilde{p}_i(p_i)) \leq \hat{\pi}_i(p, \tilde{p}_i(p)) = [\tilde{p} - (c_o + a)] \tilde{D}(\tilde{p}).$$

Assumption 1 then ensures that $d\pi_i(p_i, \tilde{p}_i(p_i))/dp_i \geq 0$ for all $p_i < c_i - a$, implying that backbones cannot gain from subsidizing consumers. *Q.E.D.*

The off-net-cost pricing principle still applies: in each segment, backbones' equilibrium prices are as if backbones' marginal cost were equal to the off-net cost; they are exactly equal to the off-net cost of receiving traffic in the competitive consumer segment and correspond to the oligopolistic price $\tilde{p}(c_o + a)$ in the website segment. Let us now assume that backbones partially pass through cost increases to websites:

Assumption 2. $0 < \tilde{p}' < 1$.

Under this reasonable assumption, by adjusting the access charge the backbones move both prices, p and \tilde{p} , in such a way that $d\tilde{p}/dp = -\tilde{p}'$ lies between -1 and 0 .

Backbones' preferred access charge maximizes the per-backbone profit

$$\Pi = (p + \tilde{p} - c) D(p) \tilde{D}(\tilde{p}).$$

The privately optimal access charge trades off the impact on the two prices p and \tilde{p} and satisfies

$$\Pi' = \left(1 + \frac{d\tilde{p}}{dp}\right) D(p) \tilde{D}(\tilde{p}) + (p + \tilde{p} - c) \left[D'(p) \tilde{D}(\tilde{p}) + D(p) \tilde{D}'(\tilde{p}) \frac{d\tilde{p}}{dp} \right] = 0.$$

Given the partial pass-through assumption, $1 + \frac{d\tilde{p}}{dp} > 0$, moving some of the communication costs from the website segment to the consumer segment increases the competitive consumer price more than it reduces the less competitive website price—in that segment, backbones keep part of the reduction in the form of increased margins. Therefore, the first term is negative and gives backbones an incentive to use the consumer segment for subsidizing the website segment. The other term reflects a profit motivation for a large volume of communication, and it increases in magnitude, the larger the total margin $p + \tilde{p} - c$.

Backbones' preferred access charge can be compared with the constrained Ramsey optimal one, which maximizes total welfare given market power on the website segment. The latter is given by

$$\max_{\{p, \tilde{p}\}} W = (v^e + \tilde{v}^e - c) D(p) \tilde{D}(\tilde{p}),$$

where v^e and \tilde{v}^e represent consumers' and websites' average surplus:

$$v^e = \frac{\int_p^{+\infty} v dF(v)}{D(p)}, \quad \tilde{v}^e = \frac{\int_{\tilde{p}}^{+\infty} \tilde{v} d\tilde{G}(\tilde{v})}{\tilde{D}(\tilde{p})}.$$

We thus have²⁵

$$\begin{aligned}
 W' &= (p + \tilde{v}^e - c) D'(p) \tilde{D}(\tilde{p}) + (v^e + \tilde{p} - c) D(p) \tilde{D}'(\tilde{p}) \frac{d\tilde{p}}{dp} \\
 &= \Pi' + \left(1 + \frac{d\tilde{p}}{dp}\right) D(p) \tilde{D}(\tilde{p}) + (\tilde{v}^e - \tilde{p}) D'(p) \tilde{D}(\tilde{p}) + (v^e - p) D(p) \tilde{D}'(\tilde{p}) \frac{d\tilde{p}}{dp}.
 \end{aligned}$$

As compared with backbones' preferred access charge, for which $\Pi' = 0$, the first additional term corresponds to a social gain from lowering consumer prices at the expense of websites. This stems from the fact that, as noted above, backbones have an excessive incentive to shift communication costs toward the more competitive consumer segment, since this cost is then passed through to consumers; backbones then benefit from lower costs in the less competitive website segment, where they pocket part of the cost reduction. The other two additional terms derive from the fact that backbones do not fully appropriate their customers' surplus. Thus, for example, attracting one more consumer gives websites an additional surplus $(\tilde{v}^e - \tilde{p})\tilde{D}(\tilde{p})$ that backbones cannot grab.

Note that, for the access charge a_{Π} that maximizes profits, we have

$$W' = \left(1 + \frac{\tilde{v}^e - \tilde{p}}{p + \tilde{p} - c}\right) \left(1 + \frac{d\tilde{p}}{dp}\right) D(p) \tilde{D}(\tilde{p}) + [(v^e - p) - (\tilde{v}^e - \tilde{p})] D(p) \tilde{D}'(\tilde{p}) \frac{d\tilde{p}}{dp}.$$

Therefore we have the following:

Proposition 6. With an elastic consumer demand and under Assumptions 1 and 2, backbones prefer an access charge that is lower than the socially optimal one, thereby favoring websites, whenever either (i) backbones leave more surplus to customers on the consumer segment ($v^e - p > \tilde{v}^e - \tilde{p}$ for $a = a_{\Pi}$) or (ii) backbones appropriate most of the cost reduction on the website segment ($d\tilde{p}/dp$ close to zero).

□ **Asymmetric access charges.** While the basic insight of the benchmark model has very broad applicability, the symmetric-access-charge assumption is crucial. We now demonstrate that asymmetric access charges are a factor of instability. Consider first the competitive backbone industry of Section 2, and now let a_i denote the access charge paid by backbone $j \neq i$ to backbone i for terminating backbone j 's off-net traffic. Without loss of generality, let us assume that

$$a_1 > a_2.$$

A first intuition is that the high-access-charge backbone 1 has a comparative advantage for both consumers (since receiving traffic is particularly attractive to this network) and websites (since terminating traffic on the rival backbone is cheaper for backbone 1). This reasoning, however, fails to account for opportunity costs. For example, if network 1 makes much money when its consumers download from network 2's websites, for the same reason network 2 finds it costly to leave consumers to network 1.

A second observation is that backbone 1 has an incentive to focus on one side of the market so as to generate off-net traffic, whereas backbone 2 has an incentive to be present on both markets so as to avoid off-net traffic. To see this, note that backbone i 's profit can be written as

$$\begin{aligned}
 \pi_i &= \alpha_i \tilde{\alpha}_j (p_i + \tilde{p}_i - c) + \alpha_i \tilde{\alpha}_j [p_i - (c_t - a_i)] + \alpha_j \tilde{\alpha}_i [\tilde{p}_i - (c_o + a_j)] \\
 &= \alpha_i [p_i - (c_t - a_i)] + \tilde{\alpha}_i [\tilde{p}_i - (c_o + a_j)] + \alpha_i \tilde{\alpha}_i (a_j - a_i),
 \end{aligned} \tag{5}$$

if all potential end users are connected.

²⁵ See Schmalensee (2002) for a similar derivation in the context of the credit card industry.

Backbone 2’s gain from a simultaneous increase in both α_2 and $\tilde{\alpha}_2$ exceeds the sum of the gains obtained by increasing α_2 or $\tilde{\alpha}_2$ alone. Similarly, backbone 1 gains more when it simultaneously increases its market share on one side and reduces its market share on the other. As a result of these conflicting interests, there is no equilibrium in pure strategies:

Proposition 7. If the backbones charge asymmetric access charges, then there is no pure-strategy equilibrium.

Proof. See the Appendix.

This inexistence problem stems from the fact that, as just noted, asymmetric access charges make the profits nonconcave.²⁶ This nonconcavity problem is somewhat robust. Suppose, for example, that the backbones are horizontally differentiated à la Hotelling on both segments, with differentiation parameters t and \tilde{t} and a unit length on each segment. The profit functions are still as in (5), but market shares are now given by

$$\alpha_i = \frac{1}{2} + \frac{1}{2t} (p_j - p_i),$$

$$\tilde{\alpha}_i = \frac{1}{2} + \frac{1}{2\tilde{t}} (\tilde{p}_j - \tilde{p}_i).$$

Therefore,

$$\frac{\partial^2 \pi_i}{\partial p_i^2} = -\frac{1}{t}, \quad \frac{\partial^2 \pi_i}{\partial \tilde{p}_i^2} = -\frac{1}{\tilde{t}}, \quad \frac{\partial^2 \pi_i}{\partial p_i \partial \tilde{p}_i} = \frac{a_j - a_i}{t\tilde{t}}$$

and thus profits are concave only if

$$t\tilde{t} \geq (a_1 - a_2)^2,$$

that is, backbones must be sufficiently differentiated (and thus not competing too effectively) on both segments—and the required level of differentiation is proportional to the asymmetry in the access charges.²⁷

Proposition 7 seems to call for reciprocal access charges. Reciprocity, however, should be understood in a broad sense, allowing for termination cost differences. For example, suppose that backbone 1 has a more expensive “shortest-exit” policy; backbone 1 then bears a larger proportion of termination transportation cost on off-net traffic: $c_i^1 = c_i^2 + \Delta$ (and thus $c_o^2 = c_o^1 - \Delta$). Then a (pure-strategy) equilibrium exists only when backbones account for this cost asymmetry when setting their access charges, that is, $a_1 = a_2 + \Delta$ (the competitive prices are then $p^* = c_i^1 - a_1 = c_i^2 - a_2$ and $\tilde{p}^* = c_o^1 + a_2 = c_o^2 + a_1$). That is, the backbone that keeps off-net traffic on its own network longer before delivering it to the other should be “rewarded” by being charged a lower termination fee.

6. Micropayments and neutrality

■ An increase in the access charge raises the cost for websites of doing business. Websites then may be tempted to pass through the increased traffic-related cost to the consumers who request the traffic. With some exceptions, such traffic-based “micropayments” do not yet exist. They require putting in place costly billing and end-user micropayment information systems.

²⁶ This nonconcavity—and the issue of the existence of an equilibrium—is reminiscent of the analysis of two-way interconnection between telecom operators, as in Laffont, Rey, and Tirole (1998a). Due to the “missing price,” however, in that case this problem appears even with symmetric access charges.

²⁷ The characterization of an access charge equilibrium in pure strategies would thus be achievable when backbones are sufficiently differentiated.

If websites pass their cost of traffic through to the consumers, the consumers' final demand does not depend on the share of termination cost that they pay directly (through the price p for receiving traffic) but rather on the total price of the communication ($p + \tilde{p}$). This in turn suggests that the way in which the total cost is *a priori* distributed between senders and receivers is irrelevant. Put differently, the access charge, which mainly affects how the cost of traffic is divided between senders and receivers, may have no impact on the consumers' final demand and thus on traffic volume. This section shows that in many contexts, the access charge is indeed neutral, i.e., it has no impact on traffic and efficiency.

We consider four illustrations. In the first, there is perfect competition at both the backbones' and websites' levels, and websites can charge consumers (through "micropayments") for their cost of traffic. As a result, backbones charge senders and receivers according to their perceived opportunity costs, as before, but consumers end up incurring the total cost of traffic regardless of the level of the access charge. The next two illustrations show that the access charge remains neutral when the consumers have an elastic demand for websites' services and when websites are not perfectly competitive. The last illustration considers situations where consumers use websites to buy goods or services. To the extent that the amount of communication is related to the volume of transactions on the goods and services, the price charged for those goods and services can play the role of micropayments. In the case of perfect correlation between communications and transactions, the access charge is again neutral, even if backbones do not perfectly compete for websites.

□ **Perfect competition at the backbone and website levels.** Let us assume that micropayments are feasible and costless. The pricing behavior of the websites depends on the degree of competition between them. Let us start with the case where there are multiple identical websites of each "type." We otherwise assume that the industry is as described in Section 2; in particular, backbones are perfect competitors on both sides, and consumers want to download one unit of traffic from each type of website. The timing goes as follows. After agreeing on an access charge, the backbones set prices (p_i for consumers, \tilde{p}_i for websites). Then, the websites subscribe and choose micropayments (denoted by s) per unit of downloaded volume. Finally, the consumers subscribe and select websites.

The backbones' profits can still be written as

$$\pi_i = \alpha_i [p_i - (c_i - a)] + \tilde{\alpha}_i [\tilde{p}_i - (c_o + a)],$$

where, as before, for each category of end user (consumer or website), the market shares only depend on the prices charged to that category. This is clear for websites, which, by choosing the backbone with the lowest website price, not only minimize the cost of their traffic, but also enhance their competitive situation. But this is also true for consumers: given the micropayment s charged by a website, they face a total price $p_i + s$ if they subscribe to backbone i ; they thus choose the backbone with the lowest consumer price.²⁸ As a result, off-net-cost pricing still prevails:

$$p_i = c_i - a,$$

and

$$\tilde{p}_i = c_o + a.$$

Bertrand websites set micropayments equal to their marginal net cost, which consists of their traffic cost, \tilde{p}_i , decreased by the value \tilde{v} that they derive from consumers' visits. So websites

²⁸ With network-based price discrimination, the subscription decision of one category affects the price paid by the other category; in particular, websites would care about consumers' subscription decisions, since it would affect their competitive situation. Different timings with respect to subscription decisions may then lead to different coordination patterns.

located on backbone i charge²⁹

$$s_i = \tilde{p}_i - \tilde{v} = c_o + a - \tilde{v}.$$

This implies that consumers bear the full cost of web traffic net of the website’s surplus, and that the access charge is neutral as regards the total price paid by consumers.³⁰ For all i ,

$$p_i + s_i = c - \tilde{v}.$$

□ **Elastic demand for websites’ services.** This neutrality result extends to the case where consumers have an elastic demand for websites’ services, of the form $q = D(p + s)$. In that case, each category of end user still selects the backbone with the lowest price for that category, so that the volume of traffic is $\hat{D} = D(\min\{p_1, p_2\} + \min\{\tilde{p}_1, \tilde{p}_2\} - \tilde{v})$. Thus, backbones’ profits can still be written as

$$\pi_i = \alpha_i \hat{D}[p_i - (c_i - a)] + \tilde{\alpha}_i \hat{D}[\tilde{p}_i - (c_o + a)], \tag{6}$$

and the standard Bertrand argument still applies to both categories of end users, so that again, $p = c_i - a$ and $\tilde{p} = c_o + a$. Hence, the volume of traffic is efficient: it is given by

$$D(p + \tilde{p} - \tilde{v}) = D(c - \tilde{v})$$

and is thus independent of the access charge.

□ **Imperfect competition among websites.** The neutrality result remains valid even when websites have market power. Suppose, for example, that there is only one website of each “type,” therefore enjoying a monopoly position for this type. For notational simplicity, suppose also that $\tilde{v} = 0$; that is, the website does not derive any direct reputational or commercial benefit from the visit. As before, each category of end user selects the lowest price offered to that category, so that the relevant prices are $\tilde{p} = \min\{\tilde{p}_1, \tilde{p}_2\}$ and $p = \min\{p_1, p_2\}$. Given those prices, each website will choose s so as to maximize its profit, given by

$$(s - \tilde{p}) D(p + s).$$

This amounts to choosing a “consumer price” $\hat{s} = p + s$ that maximizes $(\hat{s} - p - \tilde{p}) D(\hat{s})$ and thus leads to

$$s = s^M(p + \tilde{p}) - p,$$

where

$$s^M(x) = \arg \max_s (s - x) D(s),$$

thereby generating a traffic

$$\hat{D} = D(s^M(p + \tilde{p})). \tag{7}$$

Therefore, backbones’ profits are still given by (6), with \hat{D} now given by (7). As a result, Bertrand competition between the two backbones leads again to off-net-cost pricing, $p_i = c_i - a$

²⁹ Since websites pass their traffic cost through to consumers, we need not make any assumption on $c_o + a$ and \tilde{v} . As a result, s_i can be either positive or negative, depending on the value of the access charge a .

³⁰ A similar result can be found in Rochet and Tirole (2002) and Gans and King (2003) for the credit card industry. They provide conditions under which the removal of the no-discrimination rule (the rule forcing or inducing merchants (depending on the country) to charge the same price for cash and card payments) leads to a neutrality of the interchange fee.

and $\tilde{p}_i = c_o + a$; the volume of traffic and each website's profit are given by $D(s^M(c))$ and

$$\pi_w^M = (s^M(c) - c) D(s^M(c)),$$

respectively, and are thus independent of the access charge. In addition, because of the websites' market power, consumers pay more than the cost of the web traffic, but the price that they face is not affected by the access charge.

□ **Websites selling goods and services.** This setup is also relevant when there is a transaction associated with the consumer visiting the site (e.g., Amazon.com selling books through its websites). If there is perfect correlation between the bandwidth usage and the size of the transaction, the price of the transaction can play a role similar to the micropayment s . Below we consider the case where consumers buy a commodity at unit price P after having browsed the website. Buying from the website involves two types of communication costs: a search cost (browsing, listening to samples of music, etc.), which websites usually do not charge to consumers, and downloading costs, which websites can recover through the price P of the commodity. We will focus here on the latter cost and assume that downloading requires bandwidth usage q . Accordingly, the demand function for the commodity is $D(P + pq)$.

Denoting by C the unit cost of production of the commodity, the websites' profit is given by

$$[P - C - \tilde{p}q] D(P + pq) = [\hat{P} - C - (p + \tilde{p})q] D(\hat{P}),$$

where $\hat{P} = P + pq$. From this expression we see that the optimal price \hat{P} , and consequently the demand for the commodity and thus the downloading traffic, depends only on the total price $p + \tilde{p}$. With perfect competition between backbones, this total price equals c , and therefore the equilibrium traffic is independent of the access charge. With imperfect competition of the type modelled in Section 5, this total price is higher than c but remains independent of the access charge, and so is the equilibrium traffic.

7. Summing up

■ We have developed a framework for Internet backbone competition, which has allowed us to analyze the impact of access charges on backbones' competitive strategies. As we have seen, in a broad range of environments the operators set prices for their customers *as if* the customers' traffic were entirely off-net. This comes from the fact that the opportunity cost of stealing traffic away from rival operators is indeed equal to the off-net cost of traffic. In addition, the opportunity cost of creating outgoing (incoming) traffic is again equal to the off-net cost of that traffic, provided that the price for receiving (sending) traffic itself reflects its own off-net cost.

Given this off-net-cost pricing principle, in the absence of direct payments between websites and consumers, the access charge determines the allocation of communication costs between senders (mainly websites) and receivers (mainly consumers) and thus affects the level of traffic. The socially optimal access charge takes into account the demand elasticities on the two segments, but also the magnitude of the externality that each segment generates on the other segment. Since perfectly competitive backbones are indifferent to the level of the access charge, they would not object to a regulation of this access charge. In contrast, if they have market power, backbone operators' interests are in general no longer aligned with social welfare, although assessing the bias in their ideal access charge requires detailed knowledge not only of the elasticities of demand and externalities, but also of the operators' relative market power vis-à-vis websites and consumers. Finally, when websites charge micropayments, or when websites sell goods and services, the impact of the access charge on welfare is reduced and is even neutral if websites can perfectly pass through the cost of traffic to their consumers.

Appendix

■ Proof of Proposition 7. In any Bertrand equilibrium, the two backbones must charge the same prices (otherwise, the backbone charging the lower price could profitably raise that price):

$$\begin{aligned} p_1 &= p_2 = p, \\ \tilde{p}_1 &= \tilde{p}_2 = \tilde{p}. \end{aligned}$$

Furthermore, since it could attract all end users by slightly undercutting both p and \tilde{p} , backbone 2 must get at least

$$\pi_2 \geq p + \tilde{p} - c.$$

Similarly, since backbone 1 can decide to attract all consumers or all websites, it must get at least

$$\pi_1 \geq \max \{p - (c_t - a_1), \tilde{p} - (c_o + a_2)\}.$$

However, the two backbones' joint profits cannot exceed $p + \tilde{p} - c$. Hence,

$$p + \tilde{p} - c \geq p + \tilde{p} - c + \max \{p - (c_t - a_1), \tilde{p} - (c_o + a_2)\},$$

which implies

$$\begin{aligned} p &\leq c_t - a_1, \\ \tilde{p} &\leq c_o + a_2. \end{aligned}$$

Backbone 1's profit thus satisfies

$$\begin{aligned} \pi_1 &= \alpha_1 [p - (c_t - a_1)] + \tilde{\alpha}_1 [\tilde{p} - (c_o + a_2)] - \alpha_1 \tilde{\alpha}_1 (a_1 - a_2) \\ &\leq -\alpha_1 \tilde{\alpha}_1 (a_1 - a_2), \end{aligned}$$

and it can be nonnegative only if $\alpha_1 \tilde{\alpha}_1 = 0$, that is, if backbone 2 attracts all end users on at least one side of the market. Backbone 2's profit similarly satisfies

$$\begin{aligned} \pi_2 &= \alpha_2 [p - (c_t - a_2)] + \tilde{\alpha}_2 [\tilde{p} - (c_o + a_1)] + \alpha_2 \tilde{\alpha}_2 (a_1 - a_2) \\ &\leq (\alpha_2 \tilde{\alpha}_2 - \alpha_2 - \tilde{\alpha}_2)(a_1 - a_2) \end{aligned}$$

and is thus nonnegative only if $\alpha_2 = \tilde{\alpha}_2 = 0$, a contradiction. *Q.E.D.*

References

- ARMSTRONG, M. "Network Interconnection in Telecommunications." *Economic Journal*, Vol. 108 (1998), pp. 545–564.
- CAILLAUD, B. AND JULLIEN, B. "Competing Cybermediaries." *European Economic Review*, Vol. 45 (2001), pp. 797–808.
- CARTER, M. AND WRIGHT, J. "Local and Long-Distance Network Competition." Mimeo, Universities of Canterbury and Auckland, 1999a.
- AND —. "Interconnection in Network Industries." *Review of Industrial Organization*, Vol. 14 (1999b), pp. 1–25.
- CHERDRON, M. "Interconnection, Termination-Based Price Discrimination, and Network Competition in a Mature Telecommunications Market." Mimeo, Mannheim University, 2000.
- DESSEIN, W. "Network Competition in Nonlinear Pricing." *RAND Journal of Economics*, forthcoming.
- GANS, J. AND KING, S.P. "Using 'Bill-and-Keep' Interconnect Arrangements to Soften Network Competition." *Economics Letters*, Vol. 71 (2002), pp. 413–420.
- AND —. "The Neutrality of Interchange Fees in Payment Systems." *Topics in Economic Analysis and Policy*, Vol. 3 (2003), available at <http://www.bepress.com/bejeap/topics/vol3/iss1/art1/>.
- GAO, L. "On Inferring Autonomous System Relationships in the Internet." Mimeo, 2000.
- HAHN, J.H. "Network Competition and Interconnection with Heterogenous Subscribers." Mimeo, Oxford University, 2000.
- HERMALIN, B.E. AND KATZ, M.L. "Network Interconnection with Two-Sided User Benefits." Working Paper, Haas School of Business, University of California at Berkeley, 2001.

- 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. *Designing Wide Area Networks and Internetworks: A Practical Guide*. Reading, Mass.: Addison-Wesley, 1999.
- NRIC. "Service Provider Interconnection for Internet Protocol Best Effort Service." White paper of the Network Reliability and Interoperability Council, http://www.nric.org/fg/fg4/ISP_Interconnection.doc, 2002.
- ROCHET, J.-C. AND TIROLE, J. "Cooperation Among Competitors: The Economics of Payment Card Associations." *RAND Journal of Economics*, Vol. 33 (2002), pp. 1–22.
- SCHMALENSEE, R. "Payment Systems and Interchange Fees." *Journal of Industrial Economics*, Vol. 50 (2002), pp. 103–122.
- SCHWARTZ, M. AND VINCENT, D.R. "The No Surcharge Rule in Electronic Payments Markets: A Mitigation of Pricing Distortions?" Mimeo, 2000.
- WRIGHT, J. "An Economic Analysis of a Card Payment Network." Mimeo, University of Auckland, 2000.