

Internet Interconnection and the Off-Net-Cost Pricing Principle

Jean-Jacques Laffont *

Scott Marcus **

Patrick Rey ***

Jean Tirole ****

April 23, 2001

(*) IDEI, GREMAQ, and Institut Universitaire de France, Toulouse.

(**) Genuity.

(***) CREST, Paris, and IDEI, GREMAQ, University Toulouse I, Toulouse.

(****) IDEI and GREMAQ (UMR 5603 CNRS), Toulouse, CERAS (URA 2036 CNRS) Paris, and MIT.

We are grateful to Mike Riordan and to the participants of the IDEI conference on the Economics of the Software and Internet Industries (January 18-20, 2001) for helpful comments.

Abstract

The paper develops a framework for Internet backbone competition. It analyzes the impact of access charges on competitive strategies in an unregulated retail environment in which operators compete for consumers and websites. The paper compares socially optimal access charges with freely negotiated ones.

The paper derives three main insights. First, in a remarkably broad range of environments, operators set prices for their customers *as if* the customers' traffic were entirely off-net. Second, in the absence of direct payments between websites and consumers, the access charge determines the allocation of communication costs between websites and consumers and affects the level of traffic. We characterize for this environment the socially optimal access charge. When backbone operators have market power, however, they do not have in general the incentives to choose the socially optimal access charge. In contrast, when websites charge micropayments, or when websites 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. Third, the paper studies the impact of alternative contractual choices (peer vs customer relationships) in the Internet.

Keywords: Internet, Networks, Interconnection, Competition Policy.

JEL numbers: D4, K21, L41,43, L51, L96.

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 will 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 competitive deregulated environment, this universal connectivity can only be achieved if competing connectivity providers cooperatively reach agreements governing the price and quality of their interconnection.

The interconnection charges, also called “access charges,” “settlements” or “termination charges”, could be vital for enabling an efficient use of the Internet. Incentives must be provided for a widespread usage of bandwidth by dial-up and dedicated access consumers, and for the posting of high quality, low download time 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 paper is to develop a framework for modeling 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 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 three sets of insights:

Competitive strategies: On the positive side, we analyze pricing strategies in this

interconnected environment. The first key insight of the paper 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, customer cost heterogeneity, and network-based price discrimination. *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 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 a) whose demand is highly elastic, and b) who create value for other end users. Demand elasticity must be understood properly, though: For example, social welfare requires that lots of content be created and posted and also that data be properly compressed. The former (latter) objective calls for a low (high) access charge. More generally, the paper argues that the access charge cannot by itself induce all the price differentiation that would be required for an efficient allocation in the Internet. Furthermore, when backbones have market power, they do not have the incentives to 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 be 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 impact of the access charge 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. But we also identify environments in which profits are affected by the access charge. *Peering and customer relationship*: Industry participants are debating the strategic impact of being a "peer" rather than a "customer". In a customer contract, the hierarchical superior connects the customer with the rest of the Internet. In contrast, a peering agreement requires both parties to terminate traffic destined to their own customers, but does not involve any transit to third parties. We show that the contracting choice is irrelevant in a perfectly competitive industry. When backbones have market power on some segment of the industry, some customers on the opposite side may gain from moving to the status of peers with all backbones as that status allows them to threaten the other peers with a lack of connectivity and possibly to obtain below-cost termination.

The paper proceeds 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 extends the analysis to backbone market power. Section 6 introduces micropayments between customers and websites. Section 7 looks at the impact of being a peer rather than a customer.

Our paper is related to the literature on two-way access in telecommunications, e.g., Armstrong (1998) and Laffont-Rey-Tirole (1998a,b).¹ This literature assumes that while consumers both send and receive traffic, receivers get no surplus from and are not charged for receiving calls. 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 mark-up rather than to the off-net cost. Second, it creates some in-

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

stability in competition if the networks are close substitutes and the termination charge is not in the vicinity of the termination cost, while no such instability occurs here. We explain why this is so in Section 2. Last, when receivers get some utility for receiving calls, an externality must be internalized for efficiency. The availability of reception charges restores the existence of an equilibrium for arbitrary values of the access charge, but efficiency requires a specific value for this access charge (see Jeon-Laffont-Tirole 2000). In our paper too, there are two usage prices, one for sending traffic and one for receiving traffic.

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 webpages, 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 webpages 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 end users.

Most of the paper makes the following assumptions:

Balanced calling pattern: We assume that the consumers are interested in all websites, independently of websites' network choices: 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 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 pro-

²This assumption seems appropriate for the analysis of this paper. It 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. For example Chinese consumers are more likely to browse US websites than US customers to browse Chinese websites.

portional 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 3, but it is worth noting that there have been calls for regulators to impose reciprocal access charges.³

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” pattern is that most of the transportation cost c'' is born by

³At the moment, most interconnection agreements between the top level backbones (see Marcus (1999) and Gao (2000) for an overview of the Internet hierarchical organization) 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 other backbones.

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

⁵We consider mixed traffic patterns below. For simplicity, we also ignore the impact on the cost structure of caching, replication and other content delivery network schemes.

the receiving backbone.⁶ 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 webpages. To be sure, consumers and websites are more likely to use the web if they are charged lower prices; furthermore, websites can use compression techniques that lower the volume of traffic for a given content. Consequently, we will 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, or more generally 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 for the two 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 attached to each customer is then fixed, and there is thus no point distinguishing between 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

⁶Our analysis would however apply to any other way of sharing the cost of off-net traffic. For a description of hot potato routing, see Marcus (1999, Chapter 14). We here assume 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 (a contingent access charge would require much cooperation between the networks and would be harder to manage); our analysis would still apply if there were differentiated access charges, as long as differences in access charges reflected differences in termination costs.

decisions are based on the prices \tilde{p}_1 and \tilde{p}_2 charged for sending traffic. 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 $i \neq j = 1, 2$):

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

or

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

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

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

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 $v \geq c_t - a$ and $\tilde{v} \geq c_o + a$; then, there exists a unique price equilibrium. This equilibrium is symmetric and satisfies:*

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

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.

■

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 costs initially 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:⁷

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

The access charge a thus merely determines how the cost of the traffic is

⁷This 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. (Formally, there still exists an equilibrium where both backbones charge \tilde{p}^* and p^* , and no traffic is generated because no website subscribes.)

shared between senders (websites) and receivers (consumers) – a higher access charge leading to a larger burden being placed on the websites. In particular, the access charge has no impact on social welfare, defined as the sum of customers’ surpluses and network profits, which is always 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-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 paper, $p = 0$. The missing price has two important implications.

a) Pricing. The operators’ optimal usage price reflects their *perceived* marginal cost. But when operators do not charge their customers (i.e., here, consumers) for the traffic they receive, operator i ’s perceived marginal cost of outgoing (i.e. 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” (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).⁸

Note that if the “missing payment” $\alpha_i p_i$ were subtracted from the right-hand side⁹ of (3) and p_i were equal to the off-net cost¹⁰ $(c_t - a)$, then (3) would be equal to the off-net cost $(c_o + a)$. In sum, the missing payment affects the backbones’ perceived costs, and reallocates costs between origination and reception.

⁸Indeed, $c + \alpha_j (a - c_t) > c_o + a$ is equivalent to $(1 - \alpha_j)(a - c_t) < 0$.

⁹To 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.

¹⁰If consumers do not derive any utility from receiving calls ($v = 0$), as in Laffont-Rey-Tirole (1998a), the price p_i cannot be positive; networks could however subsidize receivers.

b) 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.

- **Arbitrary number of backbones.** The principle extends trivially to n backbones ($n \geq 2$): It suffices to replace in equation (1) “ α_j ” by “ $\sum_{j \neq i} \alpha_j$ ”.
- **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 charge their customers (consumers or websites) tariff:

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

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

- **Variable demand.** We have so far assumed that the volume of traffic was fixed, but the pricing principle applies as well when consumer demand is variable. For example, suppose that a website located in one network generates an additional unit of traffic. That is, the network’s new traffic is *created* rather than *stolen* from the other network, unlike in the intuition described in Section 2. If this traffic goes to the other network (“off-net”), the opportunity cost is again $c_o + a$; and if this traffic remains in the network (“on-net”) then its opportunity cost, *given*

that consumers are charged $p = c_t - a$ for receiving traffic, is still $c - p = c_o + a$. Similarly, if a consumer generates an additional unit of traffic, its opportunity cost is given by $c_t - a$, whether this traffic comes from a website located off-net or on-net, as long as websites are charged $c_o + a$ for sending traffic.

This implies in particular that the off-net-cost pricing principle still applies when consumers' demand for traffic is elastic. Denoting by $D(p)$ the demand for traffic, backbone i 's profit is then given by:

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

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; therefore, *per unit of traffic*, the opportunity cost of stealing a website is again $c_o + a$. Furthermore, whenever $p_1 = p_2 = p$, the opportunity cost of stealing a consumer away from the rival network is

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

so that, *per unit of traffic*, the opportunity cost is again $c_t - a$, which precisely coincides with the opportunity of inducing an additional unit of traffic.

Assuming $\tilde{v} \geq c_o + a$, there exists a unique price equilibrium. This equilibrium is symmetric and satisfies the off-net-cost pricing principle:

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

• **Multihoming.** Suppose now that each website may choose to locate in both backbones. Websites do not gain nor lose from multihoming as long as the backbones charge the competitive tariff $\tilde{p}^* = c_o + a$.¹¹

¹¹More generally, websites do not benefit from multihoming even if the backbones charge

• **Quality of Service (QoS):** A key challenge for the industry is the development of standards and pricing rules that will allow the supply of services, such as interactive services, that do not tolerate delays (in mean or variance) or packet losses. To the current quality, which is fine for a number of familiar applications, will need to be added a superior service, for which the carriers give Service Level Assurances to their counterparts, who can then offer the new applications to end users. The corresponding traffic will then be marked at the edge and prioritized throughout routers and fiber optic transmission. It will of course be subject to a specific access charge.

Let c_o^+ and c_t^+ denote the marginal costs of this enhanced service, where $c_o^+ > c_o$ and $c_t^+ > c_t$ due to the high cost of the on-peak router and fiber-optic capacity required to guarantee the service. Similarly, let a^+ , p^+ and \tilde{p}^+ denote the access charge, and the end users' prices for the enhanced service.

The off-net-cost pricing principle extends trivially to the enhanced service; and so

$$p^+ = c_t^+ - a^+ \quad \text{and} \quad \tilde{p}^+ = c_o^+ + a^+.$$

• **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 origina-

different prices to multihomers. Suppose that a website initially connected to backbone j decides to multihome on backbone i ($i \neq j = 1, 2$). The traffic to backbone j 's consumers remains internal to backbone j ; and the traffic to backbone i 's consumers now costs c to backbone i , whereas it previously costed $c_t - a$; the cost generated by the multihoming website is thus

$$\alpha_i [c - (c_t - a)] = \alpha_i (c_o + a).$$

Now, let \tilde{p}_i be the per unit price paid by the website to backbone i . Backbone i requests a per-unit payment greater than the cost incurred, i.e.:

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

The total payment for the website is

$$\alpha_1 \tilde{p}_1 + \alpha_2 \tilde{p}_2 \geq c_o + a.$$

A website will find this option attractive only if $\alpha_1 \tilde{p}_1 + \alpha_2 \tilde{p}_2$ is less than the offer of a backbone for a single homing website which is $c_o + a$. Hence these two inequalities must hold with equality in an equilibrium with multihoming.

tion and c_t^k at termination. Let us assume that backbones can tell the different groups apart and therefore can price discriminate.

A concrete example is provided by the US backbone market, with a fraction of the customers located abroad. European or Australian Internet Service Providers must be connected to US backbones through costly transoceanic cables that raise both origination and termination costs by the same amount (the cost of capacity of the transoceanic line) relative to a US based customer.

A straightforward extension of the analysis in Section 2 shows that *backbone competition leaves no scope for cost absorption*: Customers of type k pay

$$p^k = c_t^k - a \quad \text{for incoming traffic}$$

and

$$\tilde{p}^k = c_o^k + a \quad \text{for outgoing traffic,}$$

and thus entirely bear the cost differential with other groups of customers. We thus conclude that the off-net-cost pricing principle is robust to customer cost heterogeneity.¹²

In practice, this cost-reflecting 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 provision, by the end users or their ISPs, of circuits leading to these points of interconnection, as is the case for rural or foreign customers in the US.

• **Network-based price discrimination.** Suppose that the networks charge differentiated prices for on-net and off-net traffic, and denote by \hat{p}_i the usage price that network i charges to consumers for receiving off-net traffic, and by $\tilde{\hat{p}}_i$ the unit price that the same backbone charges to websites for sending traffic off-net. Backbone i 's profit is then given by:

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

¹²Under perfect competition, the absence of cost absorption holds regardless of whether type k customers bring substantial value to other types of customers and have an elastic demand – see Section 4.1 for an analysis with elastic demand from consumers. In contrast, a backbone monopolist would cross-subsidize among customers, and also absorb some of the cost of elastic customers whose presence is highly valued by other customers.

This price discrimination introduces network externalities: consumers' choice of backbone depends on where websites are located, and conversely – in particular, for given prices there can be multiple end-user equilibria. We will focus on “stable” end-user equilibria: that is, we will restrict attention to equilibria such that a small reallocation on one side (consumers or websites) does not trigger tipping on the other side. We show in the Appendix that such tipping occurs whenever on-net prices differ from off-net prices:

Proposition 2 *Assume that the backbones can practice network-based price discrimination. Then, in any equilibrium in which the end-user equilibrium is stable, prices are nondiscriminatory and obey the off-net-cost pricing principle:*

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

That is, backbones earn zero profits and set all usage prices equal to the perceived cost of sending or receiving off-net traffic.

Proof. See Appendix A. ■

• **Installed bases.** Suppose that backbone i has an installed base $\hat{\alpha}_i$ of consumers and an installed base $\hat{\tilde{\alpha}}_i$ 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 . Using

$$\tilde{\alpha}_i + \hat{\alpha}_i + \tilde{\alpha}_j + \hat{\alpha}_j = 1$$

and

$$\alpha_i + \hat{\alpha}_i + \alpha_j + \hat{\alpha}_j = 1,$$

simple computations yield the following expression for the operators' profits:

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

Consequently, the equilibrium prices are unchanged. And so new customers are charged the off-net-cost prices. Operator i 's profit is equal to the sum of the

last two terms in the expression of π_i . And so, if $\pi_i^*(a)$ denotes operator i 's equilibrium profit,

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

Two simple implications can be drawn from this observation. First, webhosting 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.¹³

• **Asymmetric access charges.** While the basic insight of the benchmark model has very broad applicability as shown in this section, the symmetric-access-charge assumption is crucial. We now demonstrate that asymmetric access charges are a factor of instability. Suppose that, in the competitive backbone industry of Section 2, access charges are asymmetric. 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, whereas backbone 2 has an incentive to be present on both markets. To see this, note that backbone i 's profit can be written as

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

¹³These two observations still hold if the backbones are differentiated as in Section 5.

if all potential end-users are connected.

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.¹⁴

Proposition 3 *If the backbones charge asymmetric access charges, then there is no pure strategy equilibrium.*

Proof. : See Appendix B. ■

4 Ramsey access charges

By focusing mostly 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 and studies elastic demands. Perfect competition implies that backbones' budget is always balanced, whatever the access charge. Through its allocation of costs between end users, the access charge however 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 nature and magnitude of the externality that each side exerts on the other.¹⁵ We start with simple cases where access charge actually leads to the first-best allocation. We first consider a situation where only one segment has an elastic demand. Section 4.1 first shows that the standard Ramsey recommendation ("tax the inelastic segment") applies when the elastic segment

¹⁴This result 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: $c_t^1 = c_t^2 + \Delta$ (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 $p^* = c_t^1 - a_1 = c_t^2 - a_2$ 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.

¹⁵Under the perfect competition assumption firms make zero profit and are thus indifferent to the level of the access charge. We discuss in the next section firms' incentives to choose access charges, possibly different from Ramsey ones, when they have market power.

exerts a positive externality on the other side, in that the decision of using more bandwidth (e.g., browsing the web for a consumer or posting more content for a website) benefits the other side of the market (Section 4.1). Section 4.2 stresses that, in addition, users exerting a larger externality on the other side should be “subsidized” more (implying that differentiated access charges should apply). In the same spirit, Section 4.3 shows that the standard Ramsey recommendation may be reversed when the elastic side exerts a negative externality on the other segment when deciding to use more bandwidth (e.g., by failing to use compression techniques). Finally, Section 4.4 addresses the more complex situation where both consumers and websites have elastic demands; the Ramsey charges must then account both for the demand elasticities of the two segments and for the externality that each side exerts on the other. In addition, and as one would expect, a single instrument is then in general insufficient to achieve first-best welfare maximization.¹⁶

4.1 Taxing the inelastic side

When only one side has an elastic demand, standard Ramsey principles would recommend “taxing” the inelastic side. Here, however, one must take into consideration that this elastic side exerts an externality on the other side. We first show that the standard recommendation applies when the elastic side exerts a positive externality on the other side.

To fix ideas, suppose that consumers’ demand is elastic. We will denote by $U(q)$ the gross surplus that a representative consumer derives from traffic q with a website, with $U' > 0$ and $U'' < 0$. As above in the analysis of variable demand, a consumer’s demand for traffic with a representative connected website is thus a decreasing function $D(p)$ of the usage price p for receiving traffic, and is given by $U'(D(p)) = p$.

Proposition 4 *Suppose consumers’ demand is elastic while websites’ demand is not. The socially optimal access charge taxes the inelastic segment, here the*

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

websites:

$$a^{FB} = \tilde{v} - c_o.$$

Proof. Since consumers' demand is elastic, the access charge now has an impact on traffic and thus on social welfare. Social welfare is equal to

$$\begin{aligned} W &= \sum_i \alpha_i \tilde{\alpha} [U(D(p_i)) - p_i D(p_i)] + \sum_i \tilde{\alpha}_i [\sum_j \alpha_j D(p_j)] (\tilde{v} - \tilde{p}_i) + \sum_i \pi_i \\ &= \tilde{\alpha} \sum_i \alpha_i [U(D(p_i)) + (\tilde{v} - c) D(p_i)]. \end{aligned}$$

Its first-best level thus involves $\alpha^{FB} = \tilde{\alpha}^{FB} = 1$ and a volume of traffic q^{FB} solution to:

$$\max_q \{U(q) + (\tilde{v} - c)q\}.$$

And so,

$$U'(q^{FB}) = c - \tilde{v}.$$

Because consumers' traffic is given by $U'(q) = p^*$, there exists a unique termination charge a^{FB} , defined by

$$p^* = c_t - a^{FB} = c - \tilde{v},$$

which ensures that the social welfare reaches its first-best level in equilibrium; this termination charge is equal to:

$$a^{FB} = c_t - c + \tilde{v} = \tilde{v} - c_o.$$

As one would expect, the socially optimal access charge leaves no surplus to the inelastic side of the market; the websites are charged \tilde{v} per unit of volume. ■

We have assumed that traffic is determined by consumers, who download content from websites. In practice, websites have some control over the traffic they generate. Section 4.3 analyzes the impact of compression techniques for a given amount of content. Here, we content ourselves with a mere reinterpretation of Proposition 4 in the mirror case in which the consumers have an inelastic demand and the websites' posting of content is elastic: Websites may fail to exist or, if they are set up, may economize on content in order to save on Internet connection charges. Properly reinterpreted, Proposition 4 says that if consumers

have an inelastic demand and websites an elastic one, as much cost should be allocated to consumers as possible, and so the socially optimal access charge is now equal to

$$a^{FB} = c_t - v.$$

4.2 Third-degree price discrimination

Access charges should not only allocate the cost between the two sides of the market, but also discriminate, if possible, among users on the same side. For example, different types of websites should be charged different prices. Such third-degree price discrimination has two distinct rationales.

The first rationale is the standard one in pricing theory: Consumers or websites with a low willingness to pay (a high elasticity) should as much as possible be kept on board by being charged lower prices for bandwidth and therefore the traffic “originating” on those websites should face a low access charge when crossing network boundaries.¹⁷ This being said, we should recognize that in general identifying such sites in a publicly observable way will prove very difficult, all the more that the traffic origin can be relabelled so as to arbitrage access charge differentials.

A second, and more novel, motive for discrimination is to account for differences in the externalities exerted on the other side of the market. To illustrate this, suppose that there are two types of consumers: Type A and type B , in proportions α^A and α^B ($\alpha^A + \alpha^B = 1$), with elastic demands $D^A(p^A)$ and $D^B(p^B)$ for bandwidth. To eliminate the first rationale for third-degree price discrimination, let us assume that the two types of consumers have the same elasticity of demand ($-(dD^i/dp^i)/(D^i/p^i)$ is the same for $i = A, B$). The only difference between the two groups is that websites, who still have an inelastic demand, care more about being connected to type- A consumers. That is, their gross surplus is

$$\sum_{i \in \{A, B\}} [\alpha^i D^i(p^i)] \tilde{v}^i,$$

¹⁷The example of the credit card industry is useful in this respect: Until the early 1990s, US supermarkets did not accept credit cards because they perceived the interchange fee and thus the merchant discount to be too high given that their customers have easy access to cash, including through ATMs on the premises. Visa and Master Card lowered the interchange fee on supermarket related transactions in order to get them on board.

where

$$\tilde{v}^A > \tilde{v}^B \geq 0.$$

We will assume that the two types of consumers can be told apart.

An illustration of this situation may be provided by the transoceanic customers discussed in Section 3. Type-*A* consumers are the US consumers. The websites (all based in the US for simplicity) derive less surplus from being connected to European, Australian or other remotely located consumers, because they are not targeted by the advertising, do not buy wares because of transportation costs or low income, and so forth.¹⁸

Suppose that the termination charge is a^i when the off-net website traffic is destined to a consumer of type i . The off-net-cost pricing principle holds and so

$$p^i = c_t - a^i,$$

and

$$\tilde{p} = c_o + \left[\frac{\sum_i \alpha^i D^i(p^i) a^i}{\sum_i \alpha^i D^i(p^i)} \right],$$

as websites are charged for the weighted average access charge, where the weights correspond to actual traffic requests.

Because websites have an inelastic demand, their price should be set equal to their (average) willingness to pay:

$$\tilde{p} = \frac{\sum_i \alpha^i D^i(p^i) \tilde{v}^i}{\sum_i \alpha^i D^i(p^i)}.$$

Letting $U^i(q^i)$ denote type- i consumers' gross surplus (with $U^{i'}(q^i) = p^i$), the first-best volumes of traffic are solution to

$$\max_{\{q^A, q^B\}} \left\{ \sum_i \alpha^i [U^i(q^i) + (\tilde{v}^i - c) q^i] \right\},$$

leading to

$$U^{i'}(q^i) = c - \tilde{v}^i.$$

¹⁸To clearly separate issues, we purposely assume away the cost heterogeneity that was the focus in Section 3. This cost heterogeneity is easy to reintroduce.

Since $U^i(q^i) = p^i = c_t - a^i$, first-best termination charges are thus given by

$$a^A = \tilde{v}^A - c_o > a^B = \tilde{v}^B - c_o.$$

We therefore have:

Proposition 5 *The access charges should be set so as to attract the customers who are viewed as highly desirable by the other side of the market.*

It is important to note in this respect that a competitive Internet market will not perform this function by itself in the absence of access charge discrimination (while a monopolist would). For, a backbone that would subsidize traffic by type- A consumers to raise the websites' surplus could not recoup the corresponding losses by charging a higher price to these websites since it is constrained by competition on the website market.

4.3 Taxing the elastic side

We have so far assumed that at the margin the elastic side exerts a positive externality on the other side. But when deciding to increase its own bandwidth usage, a user of the Internet may exert negative externalities as well. For example, websites can affect Internet traffic not only through their content but also through the efficiency of their compression techniques: that is, they can forward the same content with less use of the bandwidth. This has two benefits: An economy of bandwidth and a reduction in the consumer's download time. To formalize this, let us return to the case in which consumers and websites have known valuations for exchanging the content. However, this content can be forwarded more or less efficiently. Let θ denote the amount of bandwidth capacity used in the process (earlier, θ was normalized at 1). So the cost is θc . Let $C(\theta)$ denote the website's unit cost of compressing the data, with $C' < 0, C'' > 0$. Last, let \tilde{v} denote the website's gross surplus and $v(\theta)$ denote the consumer's gross surplus, where $v' < 0$ reflects the fact that the consumers dislike download time, and $v'' > 0$.

The first-best level of compression, θ^{FB} , is given by:

$$\max \{v(\theta) - \theta c - C(\theta)\},$$

or

$$C'(\theta^{FB}) = v'(\theta^{FB}) - c.$$

Let us now consider the outcome of a competitive industry with access charge a :

Proposition 6 *Suppose websites and consumers have known valuations, that websites choose the level of compression, and that the industry is competitive.*

a) *The off-net-cost pricing principle holds. Websites are charged $\tilde{p} = c_o + a$, and consumers $p = c_t - a$.*

b) *The socially optimal access charge satisfies:*

$$a = c_t - v'(\theta^{FB}), \text{ or } \tilde{p} = c + [-v'(\theta^{FB})] (> c).$$

Proof. Only part b) requires some elaboration. The websites minimize their total perceived cost (bandwidth plus compression):

$$\min\{\theta\tilde{p} + C(\theta)\},$$

and so

$$C'(\theta) = -\tilde{p}.$$

Because $\tilde{p} = c_o + a$, one obtains the access charge stated in the proposition.

■

In equilibrium, the internalization of the consumer's welfare by the website is not induced by the backbones directly. Because they compete hard for the consumers, they don't reap the benefits of improved quality for the latter. Rather, the website is taxed through the access charge to force it to internalize the externalities.

The websites' demand for bandwidth is elastic in that it reacts to price changes. However, a price increase here induces a socially *beneficial* decrease in bandwidth demand while in Section 4.1, the price increase induced a socially *harmful* decrease in bandwidth demand. This explains the apparently contradictory results.

This observation also points at the inability of the access charge to address multiple targets: Suppose that website demand is "doubly elastic": Social welfare requires that lots of content be created and also that data be compressed. The former goal calls for a low access charge and the latter goal for a high one.

4.4 Elasticity on both sides

The analysis may so far suggest that, at least as long as each side exerts a positive externality on the other, Ramsey access charges should allocate the burden to the least elastic segment. This, however, need not be the case when both segments have elastic demand. This is again because “how much” externality is exerted on the other segment also plays a major role. Suppose for example that a consumer derives surplus v , drawn from a distribution $F(v)$, from being connected with a website; similarly, a website derives a surplus \tilde{v} , drawn from a distribution $\tilde{F}(\tilde{v})$, from being connected with a consumer. Consumers’ and websites’ demands are thus given by $q = D(p) = 1 - F(p)$ and $\tilde{q} = \tilde{D}(\tilde{p}) = 1 - \tilde{F}(\tilde{p})$. 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:

Proposition 7 *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:

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

Proof. See Appendix C. ■

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 $U(q)$, 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 $U(q)/q$:

$$\tilde{p} = c - \frac{U(q)}{q}.$$

Furthermore, the average consumer surplus, U/q , and thus the optimal discount, exceeds the price p charged to consumers: $U/q - p = S/q > 0$. Hence, the total

price charged to the two segments, $p + \tilde{p}$, must be lower than the cost c , and 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 subsidies: the total price must cover the cost c . Prices must therefore be increased so as to fully cover this cost, according to standard Ramsey principles: the departure from first-best prices should be inversely related to the magnitude of demand elasticities:

$$\frac{p - \left(c - \frac{\tilde{U}(\tilde{q})}{\tilde{q}}\right)}{p} = \frac{\lambda}{\eta} \quad , \quad \frac{\tilde{p} - \left(c - \frac{U(q)}{q}\right)}{\tilde{p}} = \frac{\lambda}{\tilde{\eta}},$$

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 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 trade-off 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: $D'(p) \tilde{S}(\tilde{p}) = \tilde{D}'(\tilde{p}) S(p)$.

This optimal trade-off thus relies on demand elasticities as well as on the surpluses generated on both sides. However, and in sharp contrast with the recommendations usually derived from standard Ramsey pricing formulas, the trade-off just described can lead 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 $\tilde{\eta}_S$). For example, in the iso-elastic case ($D(p) = p^{-\eta}$, $\tilde{D}(\tilde{p}) = \tilde{p}^{-\tilde{\eta}}$), the optimality condition (4) can be written as

$$\frac{p^2}{\eta(\eta-1)} = \frac{\tilde{p}^2}{\tilde{\eta}(\tilde{\eta}-1)},$$

so that, everything else being equal, the price charged in one segment increases with the demand elasticity in that segment.¹⁹

5 Market power: Off-net-cost plus markup

Sections 3 and 4 have 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 industry is imperfectly competitive. Intuitively, the relevant marginal cost remains the off-net-cost, but a markup should be added because of market power.²⁰ Thus, while the off-net costs no longer predict average retail prices, they still define the relevant marginal usage prices.

Let us maintain the assumption that the backbones are perfectly substitutable on the consumer segment and posit that they are horizontally differentiated à la Hotelling on the website segment. That is, assume that the two operators are located at the two ends of a segment of unit length (they have addresses $x_1 = 0$ and $x_2 = 1$), that websites are uniformly distributed along this preference segment, and finally that a website located at address $x \in [0, 1]$ derives surplus

$$\alpha\tilde{v} - t(x - x_i)^2$$

from subscribing to network i . The first term is exactly as before (websites derive a value \tilde{v} from sending traffic to a representative connected consumer) while the

¹⁹For example, the Ramsey price for consumers is given by

$$p = \frac{c}{1 + \sqrt{\frac{\tilde{\eta}(\tilde{\eta}-1)}{\eta(\eta-1)}}}.$$

Furthermore, it can be checked that the first-best total price also increases whenever the demand elasticity increases on either side.

²⁰With fixed demands, one cannot distinguish between a markup on fixed fees and a markup on usage prices. We introduce variable demands at the end of this section to further discuss the location of this markup.

second term reflects the disutility from getting a service which does not exactly fit the website's ideal type of service.

In addition, to endogenize the equilibrium number of websites, we further assume that websites have two "outside" options, located at $x_1 = -\mu$ and $x_2 = 1 + \mu$, and providing a value, gross of "transportation costs" but net of any required payment, \tilde{V} . Two possible interpretations are that the content provider may stick to an old technology (e.g., sending content by mail); or that the website design is less ambitious than it could be. Given the number α of connected consumers and the prices \tilde{p}_1 and \tilde{p}_2 charged to websites by the two backbones, a website located at x subscribes to network 1, say, if

$$\alpha(\tilde{v} - \tilde{p}_1) - tx^2 \geq \max \left\{ \alpha(\tilde{v} - \tilde{p}_2) - t(1-x)^2, \tilde{V} - t(\mu+x)^2, \tilde{V} - t(1+\mu-x)^2 \right\}.$$

Appendix D solves for the equilibrium of this generalized Hotelling model, and demonstrates the following proposition:

Proposition 8 *Assume that $v > c_t - a$. Then there exists $\tilde{V}_0 > 0$ and $\tilde{V}_1 > \tilde{V}_0$ such that for $\tilde{V} < \tilde{V}_1$ there exists a unique equilibrium, which is symmetric; furthermore:*

a) *If $\tilde{V} \leq \tilde{V}_0$, the outside options are irrelevant and the equilibrium satisfies (with $\sigma = 1/2t$)*

$$\begin{aligned} p_1 &= p_2 = p^* = c_t - a, \\ \tilde{p}_1 &= \tilde{p}_2 = \tilde{p}^H = c_o + a + \frac{1}{2\sigma}, \\ \pi_1 &= \pi_2 = \pi^H = \frac{1}{4\sigma}. \end{aligned}$$

In the relevant range of the access charge, an increase in the access charge thus increases the equilibrium website price one-for-one, but does not affect the equilibrium profits.

b) *If $\tilde{V}_0 < \tilde{V} \leq \tilde{V}_1$, some websites opt for the outside options and the equilibrium satisfies*

$$\begin{aligned} p_1 &= p_2 = p^* = c_t - a, \\ \tilde{p}_1 &= \tilde{p}_2 = \tilde{p}^e(a) \quad \text{with } 0 < \frac{d\tilde{p}^e}{da} < 1, \\ \pi_1 &= \pi_2 = \pi^e(a) \quad \text{with } \frac{d\pi^e}{da} < 0. \end{aligned}$$

In the relevant range of the access charge, the equilibrium website price increases and the equilibrium profits decrease when the access charge increases.

Differentiation weakens the intensity of price competition for websites and allows backbones to earn a positive profit. However, when the websites' demand for subscription is inelastic (that is, when outside options are irrelevant, so that all websites subscribe to one of the two networks), equilibrium profits are again independent of the access charge; as before, the access charge determines the sharing of the cost between consumers and websites, but has no impact on total welfare, which is again equal to the first-best level.²¹

In contrast, when websites' subscription demand is elastic, the equilibrium profits become dependent on the access charge. A lower access charge increases the opportunity cost of servicing consumers, leading the networks to increase the price they charge consumers for receiving traffic. On the other side, the reduction in the access charge lowers the cost of servicing the websites, leading the two networks to price more aggressively and gain market share at the expense of the alternative options available to the websites; in short, decreasing the access charge allows the networks to extract more rents from the consumers; part of those rents are passed on to websites, who benefit from a reduction in the price they are charged for sending traffic to consumers, and part of it serves to increase networks' profits.

- *Privately and socially optimal access charges.*

²¹Social welfare now accounts for the quality of the match between each website's ideal service and the service actually offered by the selected network. However, the equilibrium achieves the efficient matching. In addition, Proposition 4 can be generalized to differentiated services for websites. If consumers' demand for traffic is elastic, competition between the two operators now leads to

$$\begin{aligned} p_1 &= p_2 = p^* = c_t - a, \\ \tilde{p}_1 &= \tilde{p}_2 = \tilde{p}^H = c_o + a + \frac{1}{2\sigma} \frac{1}{D(c_t - a)}, \\ \pi_1 &= \pi_2 = \pi^H = \frac{1}{4\sigma}. \end{aligned}$$

That is, both backbones earn positive Hotelling profits, whatever the termination charge; the termination charge again affects the usage price for consumers and thus the volume of traffic, and $a^{FB} = \tilde{v} - c_o$ is the unique access charge for which the equilibrium yields the first-best social optimum.

Networks favor the access charge a that maximizes their joint share of websites against outside options, provided that it is compatible with consumers' participation constraint ($v \geq c_t - a$).²² The socially efficient level of the access charge instead optimizes the match between websites' preferred services and the four options available to them (the services offered by the two networks plus the two alternative options). The networks would thus tend to negotiate an excessively low access charge, as compared with what would be socially desirable. However, a is bounded below by $c_t - v$. For a range of parameters, this lower bound is second-best socially optimal as the industry is still plagued by a suboptimal connection of websites due to the backbones' supposed market power on that segment.

- *Elastic demand for websites' services.*

While allowing for elastic subscription, the above framework involves fixed-sized individual demands, thereby eliminating any distinction between usage prices and subscription charges (fixed fees). It is however easy to extend this framework and allow for variable demand for traffic; the analysis shows that market power then affects subscription charges, while usage prices still reflect the (off-net) marginal cost of traffic. Suppose for example that websites adjust their content to the price they are charged, and denote by $\tilde{S}(\tilde{p})$ the surplus they derive (gross of subscription charges) and by $\tilde{D}(\tilde{p}) = -\tilde{S}'(\tilde{p})$ the resulting volume of content. To fix ideas, suppose that backbones compete in two-part tariffs, with backbone i charging a usage price \tilde{p}_i and a fixed subscription fee \tilde{F}_i , and, to keep things simple, neglect websites' outside options ($\tilde{V} = 0$).²³ Then, backbone i 's market share of websites is given by

$$\tilde{\alpha}_i = \frac{1}{2} + \sigma(\tilde{w}_i - \tilde{w}_j),$$

where

$$\tilde{w}_i = \tilde{S}(\tilde{p}_i) - \tilde{F}_i$$

denotes the net surplus offered to websites by backbone i . Furthermore, $p_1 = p_2 = p^* = c_t - a$ and $\alpha = 1$ (all consumers are connected) as before, so that

²²More generally, if consumer subscription were elastic as well, the backbones' ideal level for the access charge would arbitrate between the subscription elasticities of websites and consumers.

²³A similar extension is valid in case b) of Proposition 8 ($\tilde{V} > 0$).

backbone i 's profit can be written as:

$$\begin{aligned}\pi_i &= \tilde{\alpha}_i \left\{ [\tilde{p}_i - (c_o + a)] \tilde{D}(\tilde{p}_i) + \tilde{F}_i \right\} \\ &= \left[\frac{1}{2} + \sigma (\tilde{w}_i - \tilde{w}_j) \right] \left\{ [\tilde{p}_i - (c_o + a)] \tilde{D}(\tilde{p}_i) + \tilde{S}(\tilde{p}_i) - \tilde{w}_i \right\}.\end{aligned}$$

For a given net surplus \tilde{w}_i offered to the websites, backbone i 's best strategy is then to price usage at (perceived, i.e. off-net) marginal cost:

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

Imperfect competition between the two backbones however allows them to charge a positive subscription charge, equal in equilibrium to

$$\tilde{F}_i = \frac{1}{2\sigma},$$

and thus to obtain again supra-competitive profits, $\pi_i = 1/4\sigma$. Usage prices thus reflect off-net marginal costs, while backbones' market power adds a markup to subscription fees.

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 systems. Such systems further require substantial technical cooperation among Internet operators. This hindrance however is likely to be alleviated in the future. Another possible reason for the absence of micropayments relates to the transaction costs involved in charging the opposite side (advertising of micropayments), given that such transaction costs apply to a large amount of small cost transactions.

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 particular way in which the total cost is a priori distributed between senders and receivers does not

matter much. As a consequence, the access charge, which mainly affects how the cost of traffic is divided between senders and receivers, may have little 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 three illustrations. In the first situation, there is perfect competition at both the backbones' and the 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 illustration shows 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 transaction 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 described as 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 websites. 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:

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

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.²⁴ As a result, off-net-cost pricing still prevails:

$$p_i = c_t - a,$$

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. And so websites located on backbone i charge²⁵

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

This implies that consumers bear the full cost of web traffic, and that the access charge is neutral.²⁶ 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 lower price for that category,

²⁴In the absence of network-based price discrimination, the precise timing of consumers' and websites' subscription decisions is not crucial: in any event, the subscription decision of one category of end-user (consumer or website) only depends on the prices offered to this category. 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.

²⁵Since websites pass their traffic cost through to consumers, we need not make any assumption of the $c_o + a$ and \tilde{v} . As a result, s_i can be either positive or negative, depending on the value of the access charge a .

²⁶A similar result can be found in Rochet-Tirole (2000) 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 for cash and card payments) leads to a neutrality of the interchange fee.

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:

$$\pi_i = \alpha_i \hat{D}[p_i - (c_t - a)] + \tilde{\alpha}_i \hat{D}[\tilde{p}_i - (c_o + a)], \quad (5)$$

and the standard Bertrand argument still applies to both categories of end users, so that again, $p = c_t - a$ 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 $p = \min\{\tilde{p}_1, \tilde{p}_2\}$ 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

$$\hat{D} = D(s^M(p + \tilde{p})). \quad (6)$$

Therefore, backbones' profits are still given by (5), with \hat{D} now given by (6). As a result, Bertrand competition between the two backbones leads again to off-net-cost pricing, $p_i = c_t - a$ 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 set-up 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 modeled in Section 5, this total price is higher than c but remains independent of the access charge, and so is the equilibrium traffic.

For example, suppose that the backbones are perfectly substitutable on the consumer segment but horizontally differentiated à la Hotelling on the website segment, in the sense that a website located at address $x \in [0, 1]$ incurs a fixed cost $t(x - x_i)^2$ when operating through network i . Consumer prices are as before

equal to $p^* = c_t - a$. If it subscribes to network i , a website will thus obtain a profit equal to

$$\Pi(\tilde{p}_i + p^*) - t(x - x_i)^2 = \max_{\hat{P}} \left[\hat{P} - C - (p + \tilde{p})q \right] D(\hat{P}) - t(x - x_i)^2.$$

In addition, websites can opt for two “outside” options which give them $\tilde{\Pi} - t(x - \tilde{x}_i)^2$. The two backbones’ market shares thus depend on the two total prices:

$$\tilde{\alpha}_i = \tilde{\alpha}_i(\tilde{p}_1 + p^*, \tilde{p}_2 + p^*),$$

and backbone i ’s profit writes as (using $p^* = c_t - a$):

$$\begin{aligned} \pi_i &= \alpha_i \tilde{\alpha}_i [p_i - (c_t - a)] + \tilde{\alpha}_i [\tilde{p}_i - (c_o + a)] \\ &= \tilde{\alpha}_i(\tilde{p}_1 + p^*, \tilde{p}_2 + p^*)(\tilde{p}_i + p^* - c), \end{aligned}$$

which is thus unaffected by the access charge a . Consequently, the price \tilde{p} (and thus the commodity price \hat{P} , the volume of traffic, etc.) is also unaffected by the access charge.

7 Peer vs customer relationships

Current interconnection agreements fall into two categories, peering and customer relationships. There are roughly two features distinguishing these contracts:

- *address advertising*: Under a peering contract, each party advertises its customers and only its customers to the other party. That is, parties do not offer transit to third parties and accept only traffic destined to their own customers. In contrast, a customer relationship is a hierarchical one in which the hierarchical superior advertises the rest of the Internet to the customer, and conversely advertises the customer to all other parties. That is, the superior offers to route the customer’s traffic to and from the rest of the Internet.
- *payments*: Many peers still use “bill-and-keep” and so terminate each other’s traffic but do not charge each other for doing so ($a = 0$). In contrast, customers pay for the traffic they receive and originate.

In our view, the key distinction between peer and customer relationships is the former feature, since some backbones already charge for peering under some

circumstances. The analysis in this section has an exploratory nature. It shows that game theory can be used to formalize and help think about contractual differences in interconnection agreements. In particular, it brings some insights as to why the peer status may be perceived as desirable in some environments – the peer status is indeed often viewed as desirable in the Internet industry: some of the top backbones (Genuity, AT&T, Level 3, Qwest) have had a deliberate policy of moving from the customer to the peer status; in contrast, peers are usually reluctant to return to customer status.

Our analysis so far has looked at the competition for customers among backbones linked under paid peering contracts. This section asks: Could a large customer (an ISP or content provider, say) gain from entering peering contracts with the backbones? Our key insight here is that a customer may be able to obtain a better deal, indeed possibly a price below the marginal cost of serving it, if a) it enters peering relationships with all backbones, and b) the backbones have market power with respect to websites. In a nutshell, peering with backbone 1 is a way for the ISP to threaten backbone 2 with a lack of connectivity with the ISP's subscribers if the negotiation with backbone 2 breaks down; and this threat is effective if backbone 2 has something to lose, i.e. is profitable on the other side.

To make this point, we focus on a consumer (an ISP such as AOL,²⁷ say) that is large enough to be able to negotiate with the two backbones. We assume that all other end-users are customers of the backbones (so we do not look at situations in which the majority of end-users attempt to use the tactics studied here). It will be clear that the ideas apply much more generally. Consider the following timing:

Stage 1: The two backbones enter a paid peering agreement with termination fee *a*.

Stage 2: The ISP makes take-it-or-leave-it offers to the backbones.

Stage 3: The backbones make take-it-or-leave-it offers to the websites and to the other consumers.

This timing deserves some comments. The interaction between backbones,

²⁷It could alternatively be a large content provider (e.g., a web-hosting ISP), or even a small backbone currently under a customer contract and contemplating a move from customer to peer status.

and consumers other than the ISP and websites (stage 3) is as earlier: The backbones set prices for the websites. There is no loss of generality here in assuming that the backbones offer customer contracts. The new feature is in stage 2.

We assume that backbones are perfectly substitutable on the consumer side and have market power on the website side. For illustrative purposes only, we use the (no outside good) Hotelling model of Section 5 in which the backbones are located at the two extremes of the Hotelling segment.

We ask: Is this an equilibrium for the ISP to stick to the customer status? Suppose it were. Then, from perfect backbone substitutability and the off-net-cost pricing principle, the ISP, like other consumers at stage 3, obtains price

$$p^C = p^* = c_t - a,$$

where “ C ” stands for “customer” status. Note that even though the backbones want to connect to the ISP so as not to be handicapped in their stage 3 rivalry when offering contracts to websites, they are unwilling to lose money on the ISP: *If a backbone is not connected directly to the ISP, it will be connected indirectly through the conjunction of this ISP’s customer contract with the other backbone and the peering agreement between backbones.* In other words, a backbone can never lose connectivity with a customer in a world of backbone peering and customer contracts with the rest of the Internet.

Suppose that the ISP represents a fraction x of consumer traffic. Let the ISP deviate and offer *peer* contracts to the two backbones. Assuming that backbone 1 accepts to peer, backbone 2, if it rejects the peering proposal, loses connectivity with the ISP,²⁸ and therefore offers quality of service $(1 - x)\tilde{v}$ to the websites, while backbone 1, provided it accepts the ISP’s peering proposal, offers quality \tilde{v} . The key point then is that backbone 2’s profit in the stage 3 website market decreases. Although the exact amount is irrelevant, this decrease in the Hotelling

²⁸This loss of connectivity with the ISP is costly to the backbone only if the ISP’s customers are single homed and sufficiently locked-in. For a further discussion of installed bases and their strategic use in the Internet industry, see Crémer et al. (2000).

model for x small²⁹ is

$$x\tilde{v}/3$$

Since by assumption the ISP can offer a per unit price which leaves no profit to the backbone 1, it will choose a price p^P (where “ P ” stands for “peer” status) such that

$$x \left[p^P - (c_t - a) + \frac{\tilde{v}}{3} \right] = 0,$$

or

$$p^P = (c_t - a) - \frac{\tilde{v}}{3} < p^C.$$

Interestingly, the off-net-cost pricing principle no longer holds (even though the price remains guided by the off-net cost).

The “subsidy” relative to the marginal cost ($c_t - a$) reduces each backbone’s profit by $x\tilde{v}/3$. The ISP is therefore strictly better off under peering contracts.

Proposition 9 *A backbone can never lose connectivity with a party in a world of backbone peering and customer contracts with the rest of the Internet. In contrast, peering offers may allow a large end-user (ISP, content provider,...) to threaten unilateral reductions in the backbones’ connectivity, and thereby to obtain prices below those predicted by the off-net-cost pricing principle.*

This section has shown that, for an organization that has a full set of peering relationships, (paid) peering may be cheaper than transit. It bears emphasizing that our analysis is only a first step toward understanding the institutions of interconnection contracts. It should be extended in several directions. First, reaching the status of peer may confront difficulties that are understated by the nature of the game we postulated. In particular, we assumed that the ISP has full information about the backbone’s cost functions; conversely the backbones are fully aware of the cost that the ISP’s traffic will impose on them. To be sure, such uncertainties would affect the negotiation of both customer and peering

²⁹More generally, one can show that, as long as x is not so large (relative to product differentiation) that the lower connectivity backbone no longer attracts websites, the ISP can offer

$$p' = (c_1 - a) - \left[\frac{\tilde{v}}{3} - \frac{x(\tilde{v})^2}{18t} \right].$$

contracts. In particular, the ISP may think twice about seeking peer status with all backbones, because the loss of connectivity with one of them (due to the ISP's unwittingly being too greedy) could hurt the ISP greatly, especially if its customers were to switch quickly in case of degradation of service. Along similar lines, it would be useful to look at different bargaining structures and the credibility of the ISP's willingness to play its card. Second, we have assumed that the ISP makes simultaneous offers. Because of the stickiness and staggered nature of interconnection agreements, the ISP's move from customer to peer status would be gradual. A reasonable conjecture in such environments is that the ISP would be able to extract better terms once it already has a large set of peering contracts than from the first backbones it starts peering with; that is, the bargaining position would seem to get stronger as the ISP nears the ubiquitous peer status. Third, we leave for future research the analysis of situations in which a sizeable fraction (in volume) of end-users try to play the peering game. If, here, enough ISPs insist on a peering contract, then there is no backbone profit left to be defended and so the backbones have no reason to accede to all peering requests, which would force them to lose money.

References

- Armstrong, M. (1998), "Network Interconnection," *Economic Journal*, 108, 545-564.
- Caillaud, B. and B. Jullien (2000), "Competing Cybermediaries," forthcoming, *European Economic Review Papers and Proceedings*.
- Carter, M. and J. Wright (1999a), "Local and Long-Distance Network Competition," mimeo, Universities of Canterbury and Auckland.
- Carter, M. and J. Wright (1999b), "Interconnection in Network Industries," *Review of Industrial Organization*, 14:1-25.
- Cherdron, M. (2000), "Interconnection, Termination-Based Price Discrimination, and Network Competition in a Mature Telecommunications Market," mimeo, Mannheim University.
- Crémer, J., P. Rey and J. Tirole (2000), "Connectivity in the Commercial Internet," *Journal of Industrial Economics*, 48:433-472.
- Dessein, W. (1999a), "Network Competition with Heterogeneous Calling Pattern," mimeo.
- Dessein, W. (1999b), "Network Competition in Nonlinear Pricing," mimeo.
- Gans, J. and S. P. King (1999), "Using 'Bill-and-Keep' Interconnect Arrangements to Soften Network Competition," mimeo, <http://www.mbs.unimelb.edu.au/jgans>.
- Gao, L. (2000), "On Inferring Autonomous System Relationships in the Internet", mimeo.
- Hahn, J.H. (2000), "Network Competition and Interconnection with Heterogenous Subscribers," mimeo, Oxford University.
- Huston, G. (1999) "Interconnection, Peering and Settlements," http://www.isoc.org/inet99/proceedings/1e/1e_1.htm.
- Jehiel, P., B. Moldovanu, and E. Stacchetti (1996) "How (Not) to Sell Nuclear Weapons," *American Economic Review*, 86: 814-829.

- Jeon, D.S., Laffont, J.-J. and J. Tirole (2000), "On the Receiver Pays Principle," mimeo, University Pompeu Fabra, Barcelona, and IDEI, Toulouse.
- Laffont, J.-J. and J. Tirole (1999), *Competition in Telecommunications*, Cambridge: MIT Press.
- Laffont, J.-J., P. Rey and J. Tirole (1998a), "Network Competition: I. Overview and Nondiscriminatory Pricing," *The Rand Journal of Economics*, 29, 1-37.
- Laffont, J.J., P. Rey and J. Tirole (1998b), "Network Competition: II. Price Discrimination," *The Rand Journal of Economics*, 29(1), 38-56.
- Marcus, S. (1999), *Designing Wide Area Networks and Internetworks: A Practical Guide*, Reading, MASS: Addison-Wesley.
- Rochet, J.-C. and J. Tirole (2000), "Cooperation among Competitors: The Economics of Payment Card Associations," mimeo, IDEI.
- Schmalensee, R.(2000) "Payment Systems and Interchange Fees," mimeo, MIT.
- Schwartz, M. and D. R. Vincent (2000), "The No Surcharge Rule in Electronic Payments Markets: A Mitigation of Pricing Distortions?" mimeo.
- Wright, J. (2000), "An Economic Analysis of a Card Payment Network," mimeo, University of Auckland.

Appendix

A Proof of Proposition 2

For consumers, the relative “attractiveness” of backbone 2, say, can be measured by the surplus differential derived from joining backbone 2 rather than backbone 1, or

$$A_1 = (\tilde{\alpha}_1 p_1 + \tilde{\alpha}_2 \hat{p}_1) - (\tilde{\alpha}_2 p_2 + \tilde{\alpha}_1 \hat{p}_2) = \hat{p}_1 - p_2 + \tilde{\alpha}_1 [(p_1 - \hat{p}_1) + (p_2 - \hat{p}_2)].$$

In any equilibrium where the market for consumers is shared, consumers must be indifferent between subscribing to one or the other network, that is,

$$A_1 = 0. \tag{7}$$

Furthermore, requiring the end user equilibrium to be stable amounts to imposing that A_1 must remain equal to 0 when a small perturbation affects $\tilde{\alpha}_1$. This, in turn, imposes

$$\frac{\partial A_1}{\partial \tilde{\alpha}_1} = 0. \tag{8}$$

This second condition implies

$$p_1 - \hat{p}_1 = \hat{p}_2 - p_2,$$

which, combining with (7), yields $\hat{p}_1 = p_2$. Combined with (8), this leads to

$$\begin{aligned} \hat{p}_1 &= p_2, \\ \hat{p}_2 &= p_1. \end{aligned} \tag{9}$$

In other words, in any stable shared market for consumers, the off-net price charged to consumers by one backbone must be equal to the on-net price of the rival backbone. The same argument applies to any stable shared market for websites, leading to:

$$\begin{aligned} \hat{\tilde{p}}_1 &= \tilde{p}_2, \\ \hat{\tilde{p}}_2 &= \tilde{p}_1. \end{aligned} \tag{10}$$

Now, backbone i 's profit is then given by:

$$\begin{aligned}
\pi_i &= \alpha_i \tilde{\alpha}_i (p_i + \tilde{p}_i - c) + \alpha_i \tilde{\alpha}_j (\hat{p}_i - (c_t - a)) + \alpha_j \tilde{\alpha}_i (\widehat{\tilde{p}}_i - (c_o + a)) \\
&= \alpha_i \tilde{\alpha}_i [p_i - (c_t - a)] + \alpha_i \tilde{\alpha}_i [\tilde{p}_i - (c_o + a)] + \alpha_i \tilde{\alpha}_j [\hat{p}_i - (c_t - a)] + \alpha_j \tilde{\alpha}_i [\widehat{\tilde{p}}_i - (c_o + a)] \\
&= \alpha_i [\tilde{\alpha}_i p_i + \tilde{\alpha}_j \hat{p}_i - (c_t - a)] + \tilde{\alpha}_i [\alpha_i \tilde{p}_i + \alpha_j \widehat{\tilde{p}}_i - (c_o + a)] \\
&= \alpha_i [\hat{p}_i + \tilde{\alpha}_i (p_i - \hat{p}_i) - (c_t - a)] + \tilde{\alpha}_i [\tilde{p}_i + \alpha_i (\tilde{p}_i - \widehat{\tilde{p}}_i) - (c_o + a)].
\end{aligned}$$

If the end-user equilibrium is stable, a change in consumers' behavior does not affect the behavior of the websites. Furthermore, since α_i can be drastically increased (to 1) or reduced (to 0) with an infinitesimal change in either p_i or \hat{p}_i , in a shared market equilibrium it must be the case that this expression of the profit is independent of α_i , that is:

$$\frac{\partial \pi_i}{\partial \alpha_i} = [\hat{p}_i + \tilde{\alpha}_i (p_i - \hat{p}_i) - (c_t - a)] + \tilde{\alpha}_i (\tilde{p}_i - \widehat{\tilde{p}}_i) = 0, \quad (11)$$

which yields:

$$\begin{aligned}
\hat{p}_1 + \tilde{\alpha}_1 (p_1 - \hat{p}_1 + \tilde{p}_1 - \widehat{\tilde{p}}_1) &= c_t - a, \\
\hat{p}_2 + \tilde{\alpha}_2 (p_2 - \hat{p}_2 + \tilde{p}_2 - \widehat{\tilde{p}}_2) &= c_t - a.
\end{aligned}$$

Using ($\hat{p}_1 = p_2, \hat{p}_2 = p_1$), these conditions can be rewritten as

$$\begin{aligned}
p_2 + \tilde{\alpha}_1 (p_1 - p_2 + \tilde{p}_1 - \tilde{p}_2) &= c_t - a, \\
p_1 + \tilde{\alpha}_2 (p_2 - p_1 + \tilde{p}_2 - \tilde{p}_1) &= c_t - a.
\end{aligned}$$

Subtracting one equality to the other then yields (using $\tilde{\alpha}_1 + \tilde{\alpha}_2 = 1$)

$$(\tilde{\alpha}_1 + \tilde{\alpha}_2) (\tilde{p}_1 - \tilde{p}_2) = 0,$$

that is,

$$\tilde{p}_2 = \tilde{p}_1.$$

Therefore, the equilibrium prices charged to websites are symmetric: $\tilde{p}_2 = \tilde{p}_1$ and, using (9),

$$\widehat{\tilde{p}}_2 = \widehat{\tilde{p}}_1 = \tilde{p}_2 = \tilde{p}_1.$$

Similarly, if the end-user equilibrium is stable, then in a shared market equilibrium the above expression of the profit must be independent of $\tilde{\alpha}_i$, that is:

$$\frac{\partial \pi_i}{\partial \tilde{\alpha}_i} = 0, \quad (12)$$

which implies that the prices charged to consumers are also symmetric:

$$\hat{p}_2 = \hat{p}_1 = p_2 = p_1.$$

The no-deviation conditions (11) and (12) then imply that the off-net-cost pricing principle applies:

$$\begin{aligned} \hat{p}_2 &= \hat{p}_1 = p_2 = p_1 = c_t - a, \\ \widehat{\tilde{p}}_2 &= \widehat{\tilde{p}}_1 = \tilde{p}_2 = \tilde{p}_1 = c_o + a. \end{aligned}$$

Consider now a corner equilibrium, such as $\alpha_1 = \tilde{\alpha}_2 = 1, \alpha_2 = \tilde{\alpha}_1 = 0$, for which $\pi_1 = \hat{p}_1 - (c_t - a)$ and $\pi_2 = \widehat{\tilde{p}}_2 - (c_o + a)$. Since equilibrium profits cannot be negative, we have:

$$\begin{aligned} \hat{p}_1 &\geq c_t - a, \\ \widehat{\tilde{p}}_2 &\geq c_o + a. \end{aligned}$$

In addition, $\alpha_1 = 1$ implies $p_2 \geq \hat{p}_1$; furthermore, the constraint must be binding, since otherwise, backbone 1 could enhance its profits by slightly increasing \hat{p}_1 . Hence,

$$p_2 = \hat{p}_1$$

and similarly,

$$\tilde{p}_1 = \widehat{\tilde{p}}_2.$$

The stability of the end-user equilibrium imposes

$$\frac{\partial A_1}{\partial \tilde{\alpha}_1} \leq 0, \quad (13)$$

that is,

$$p_1 \leq \hat{p}_2. \quad (14)$$

Then, if $\hat{p}_1 > c_t - a$, backbone 2 could increase its profits by attracting all consumers through a slight reduction in p_2 (note that under (14), backbone 2 would keep all websites). Hence, necessarily

$$p_2 = \hat{p}_1 = c_t - a,$$

and similarly,

$$\tilde{p}_1 = \hat{p}_2 = c_o + a,$$

so that both backbones make zero profit.

Finally, $p_1 < c_t - a$ cannot be part of a subgame perfect equilibrium. And if $\hat{p}_2 > c_t - a$, backbone 1 could make a positive profit by attracting all websites (through a slight reduction in \tilde{p}_1) and charging p_1 just below \hat{p}_2 . Hence,

$$p_1 = \hat{p}_2 = c_t - a,$$

and, similarly:

$$\tilde{p}_2 = \hat{p}_1 = c_o + a.$$

A similar reasoning applies to the other corner equilibria.

B Proof of Proposition 3

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 can be non-negative 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 non-negative only if $\alpha_2 = \tilde{\alpha}_2 = 0$, a contradiction.

C Proof of Proposition 7

Denoting consumers' gross surplus by $U(q) = S(P(q)) + P(q)q$, where $P(q) = D^{-1}(q)$, and similarly for websites, social welfare is equal to

$$W = \tilde{q}U(q) + q\tilde{U}(\tilde{q}) - cq\tilde{q}.$$

Its first-best level is thus characterized by (using $U'(q) = P(q)$):

$$p = c - \frac{\tilde{U}(\tilde{q})}{\tilde{q}}, \quad (15)$$

$$\tilde{p} = c - \frac{U(q)}{q}, \quad (16)$$

or, equivalently (using $U/q = p + S/D$):

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

Ramsey prices maximize W subject to the budget constraint:

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

Denoting by λ the Lagrange multiplier associated with this budget constraint, the solution is characterized by standard Ramsey equations:

$$\frac{p - \left(c - \frac{\tilde{U}(\tilde{q})}{\tilde{q}}\right)}{p} = \frac{\lambda}{\eta}, \quad (17)$$

$$\frac{\tilde{p} - \left(c - \frac{U(q)}{q}\right)}{\tilde{p}} = \frac{\lambda}{\tilde{\eta}}, \quad (18)$$

where $\eta = -pD'(p)/D(p)$ and $\tilde{\eta} = -\tilde{p}\tilde{D}'(\tilde{p})/\tilde{D}(\tilde{p})$ denote demand elasticities for consumers and websites. Using again $U/q = p + S/D$ and $p + \tilde{p} = c$, these conditions boil down to:

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

that is:

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

D Proof of Proposition 8

When both networks and the outside options have positive market shares,³⁰ backbone 1's market share is of the form $[\underline{x}_1, \bar{x}_1]$, where boundaries are characterized by

$$\begin{aligned} \alpha(\tilde{v} - \tilde{p}_1) - t\underline{x}_1^2 &= \tilde{V} - t(\mu + \underline{x}_1)^2, \\ \alpha(\tilde{v} - \tilde{p}_1) - t\bar{x}_1^2 &= \alpha(\tilde{v} - \tilde{p}_2) - t(1 - \bar{x}_1)^2, \end{aligned}$$

that is,

$$\begin{aligned} \underline{x}_1 &= -\frac{\mu}{2} + \frac{\tilde{V} - \alpha(\tilde{v} - \tilde{p}_1)}{2t\mu}, \\ \bar{x}_1 &= \frac{1}{2} + \frac{\alpha(\tilde{p}_2 - \tilde{p}_1)}{2t}. \end{aligned}$$

³⁰The condition $c_o + a < \tilde{v} - \tilde{V} + (1 - \mu)t\mu$ ensures that some websites subscribe to one or the other network, whereas the condition $c_o + a > \tilde{v} - \tilde{V} + \left(1 - \mu - \frac{2-3\mu}{1-\mu}\right)t\mu$ ensures that some websites choose not to subscribe (i.e., $\tilde{\alpha} < 1$).

Assuming that all consumers are connected ($\alpha = 1$), backbone 1's market share of websites is thus given by³¹

$$\begin{aligned}\tilde{\alpha}_1 &= \bar{x}_1 - \underline{x}_1 = \frac{1 - \mu}{2} + \frac{\tilde{p}_2 - \tilde{p}_1}{2t} - \frac{\tilde{V} - (\tilde{v} - \tilde{p}_1)}{2t\mu} \\ &= \frac{1 - \mu}{2} - \frac{\tilde{V} - \tilde{v}}{2t\mu} + \frac{1}{2t} \left[\tilde{p}_2 - \tilde{p}_1 - \frac{1}{\mu} \tilde{p}_1 \right] \\ &\equiv \frac{d}{2} + \sigma [\tilde{p}_2 - (1 + \rho) \tilde{p}_1],\end{aligned}$$

with

$$\begin{aligned}d &= 1 - \mu - \frac{\tilde{V} - \tilde{v}}{t\mu}, \\ \sigma &= \frac{1}{2t}, \\ \rho &= \frac{1}{\mu},\end{aligned}$$

and similarly

$$\tilde{\alpha}_2 = \frac{d}{2} + \sigma [\tilde{p}_1 - (1 + \rho) \tilde{p}_2].$$

As long as all consumers subscribe, backbone i 's profit is given by:

$$\begin{aligned}\pi_i &= \alpha_i \tilde{\alpha}_i (p_i + \tilde{p}_i - c) + \alpha_i \tilde{\alpha}_j (p_i - (c_t - a)) + \alpha_j \tilde{\alpha}_i (\tilde{p}_i - (c_o + a)) \\ &= \alpha_i \tilde{\alpha} [p_i - (c_t - a)] + \tilde{\alpha}_i [\tilde{p}_i - (c_o + a)].\end{aligned}$$

Some potential websites may not subscribe to a network if the outside options are sufficiently attractive. The total number of connected websites is now given by:

$$\tilde{\alpha} = \tilde{\alpha}_1 + \tilde{\alpha}_2 = d - 2\sigma\rho(\tilde{p}_1 + \tilde{p}_2)$$

if this expression is lower than 1.

The standard Bertrand argument again applies to the inelastic consumer segment. The price paid by consumers does not affect their demand (as long as

³¹If outside options are not attractive, all websites are connected to one or the other network and

$$\tilde{\alpha}_1 = 1 - \tilde{\alpha}_2 = \bar{x}_1 = \frac{1}{2} + \frac{\alpha(\tilde{p}_2 - \tilde{p}_1)}{2t}.$$

it remains below v) and thus has no effect on the website business. Therefore, $p_1 = p_2 = c_t - a$, which by assumption is lower than v .

Next, note that, given network j 's price $p_j = c_t - a$, no price p_i can, for network i , dominate $p_i = c_t - a$: by setting a different price p_i , network i cannot modify the number of connected consumers and thus cannot affect the website business, and it can only reduce its profit (i.e., make losses) on the consumer segment. Therefore, given $p_j = c_t - a$, network i 's maximal profit is given by

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

which is concave in \tilde{p}_i . Simple computations yield:

$$\begin{aligned} \tilde{p}_1 &= \tilde{p}_2 = \tilde{p}^e(a) = \frac{d}{1+2\rho} \cdot \frac{1}{2\sigma} + \frac{1+\rho}{1+2\rho} (c_o + a), \\ \pi_1 &= \pi_2 = \pi^e(a) = \frac{1+\rho}{\sigma(1+2\rho)^2} \left[\frac{d}{2} - \sigma\rho(c_o + a) \right]^2. \end{aligned}$$

The thresholds \tilde{V}_0 and \tilde{V}_1 are respectively such that $\tilde{\alpha}(\tilde{p}^e, \tilde{p}^e) = 1$ and 0.