# Peering and Settlement in the Internet: An Economic Analysis

Iain Little and Julian Wright*

September 1999

**Abstract**

This paper explores the implications of Internet peering in the context of a model of competing, vertically integrated Internet Access Providers. We show that if regulation forbids settlement payments between firms, there will be under-investment in capacity and under-pricing of usage, both of which lead to excessive congestion. To overcome these problems, firms that are net providers of Internet infrastructure should be allowed to charge firms that are net users. We characterize the efficient level of these access payments, assuming usage can be appropriately measured. We find that refusal to peer and the charging of settlement payments may well be efficiency enhancing.

*Iain Little, Network Economics Consulting Group (NECG). Julian Wright, University of Auckland.*
*We would like to thank two anonymous referees for valuable comments. We are also indebted to Henry Egras, Director of NECG, for insightful comments as well as financial assistance. Additionally, the staff of Telstra Corp. of Australia have provided valuable institutional information. However, the views expressed here are those of the authors alone, as is the responsibility for any remaining errors.*

## 1. Introduction

Despite the explosive growth of the Internet over the last decade, policy makers and economists have had relatively little to say about how, if at all, Internet interconnection should be regulated or controlled.[1] This lack of literature is all the more surprising given that in many countries a few Internet Access Providers (IAPs) usually dominate the provision and control of the backbone of the Internet, and given the rapid rate at which IAPs' settlement revenue is currently growing.[2] Since IAPs need access to each other's facilities to compete, it would seem the terms of settlement between them could be used to thwart competition. In fact, recently in Australia, the telecommunication regulation authority threatened the largest IAP (Telstra) with fines if it continued to charge other IAPs for the supply of services while not paying for services supplied by other IAPs.[3]

In this paper we argue that interconnection pricing between asymmetric networks is not only legitimate but also likely to be welfare enhancing. Regulations that forbid settlement payments are likely to lead to severe congestion problems. Allowing firms to charge for interconnection, or to refuse to peer with smaller firms, can improve efficiency. To understand these claims, the commercial terms for interconnection of IAPs first needs to be explained.

The original form of interconnection, consistent with the Internet's public sector origins, was that each IAP would bear its own costs, with the costs of the links between these being covered by government subsidies.[4] As the commercial element in the Internet became more dominant, a pattern evolved centered on two forms of interconnection.

The first is generally referred to as *peering*. In these arrangements, one IAP (say IAP 1) grants another IAP (say IAP 2) access to its network in exchange for IAP 1 also gaining access

---

[1] There is, however, a comprehensive literature on other aspects of the Internet, such as end user pricing and quality of service issues.

[2] For the US, the increase in settlement revenues is a direct result of UUnet's 1997 decision (and other IAPs following that decision) to end the era of free Internet backbone access. Worldwide, TARIFICA (1998) reported that just under $4 billion was paid in access fees at the end of 1996. Huston (1999, p.19) describes this evolution of the Internet, as does Baake and Wichmann (1999, p.91).

[3] For a discussion on regulatory involvement in the Australian ISP market, see Ergas (1999).

[4] The Internet, as most experience it, is purely commercial. However, some portions remain subsidized; these include secure US government owned networks, which have come under criticism for selling services to Universities (see Commercial Internet Exchange Association, 1995) and the Internet 2.

to IAP 2's network.[5] In theory, peering arrangements could take a number of forms. In practice, however, peering is generally done without settlement, and hence can be considered a barter transaction.[6] Each IAP bears the cost of the other IAP's use of its network in exchange for the benefit of the use of the other's network.

The second form of interconnection is generally referred to as *transit*. In this arrangement, IAP 2 pays IAP 1 for connectivity to that part of the Internet that is accessible to IAP 1's network (this includes all of its customers and the other networks it peers with). Transit functions are generally provided by access providers who operate backbone networks – that is, the links connecting networks located in different places.

Generally, access providers of similar size peer, while transit services to other networks are purchased from larger and sold to smaller IAPs. The largest group of access providers in the US market, those who own a significant backbone, collectively and exclusively peer.[7] Alternatively, regional, smaller IAPs peer with each other and purchase transit from the larger IAPs to access parts of the Internet that are geographically distant. Further details of the current Internet market structure, as well as its history, are provided by Srinagesh (1997) and Huston (1999).

A number of important policy questions naturally arise from this market structure. Should IAPs be allowed to exclusively peer by restricting peering to only similar networks? Should IAPs be allowed to charge for transit? If so, should the rate they charge be regulated, and in this case at what level?

To answer these questions we provide a model of competing IAPs that determines the efficient terms of peering, as well as the implications of settlement-free peering. In our model, firms each provide some capacity, which under settlement-free peering is available to them both. Firms compete for customers through two-part tariffs, the usage sensitive prices dictating the demand for capacity. We first characterize the welfare maximizing choice of prices and capacity. In this case, we show the social planner would choose usage prices above the marginal cost each firm faces. These prices take into account the extra cost

---

[5] Here, the IAP's network refers to the infrastructure owned by that IAP and the infrastructure owned by its customers. Therefore when two IAPs peer, their customers also benefit from the interconnection agreement.

[6] Peering without settlement is equivalent to bill and keep and sender keep all interconnection.

[7] This group broadly consists of UUNet, MCIWorldcom, Sprint and GTE/BBN. There are exceptions to this; PSINet offer free peering to most access seekers while owning a significant amount of Internet backbone. Probably for this reason, larger backbone providers do not peer with PSINet and therefore it is often considered a smaller IAP.

additional usage would imply for capacity provision. The optimal level of capacity provision is then the amount that is sufficient to satisfy usage demand at these prices; that is, there is no congestion.[8]

In contrast, if regulation forbids settlement payments, competition will lead to congestion: firms, through competition, price at a level in which demand for capacity exceeds supply. Each individual firm under-prices usage to allow its customers to gain a larger share of the rationed capacity. In equilibrium, both firms do this and so neither actually gains. The rationing equilibrium is characterized by each firm's usage price set equal to marginal cost.[9] In addition to under-pricing, we show the equilibrium also suffers from a free-rider problem, in which each firm provides less infrastructure than the other. Because both firms jostle to be the one that ends up providing less capacity, it is likely that they both end up severely rationing customers.

Having shown the negative implications of settlement-free peering, we consider how allowing firms to price for interconnection or to restrict peering to similar firms can overcome these problems. We show that if firms charge for the use of their facilities, the efficient outcome can be achieved. In particular, we consider a settlement regime in which if either firm has a shortfall of infrastructure provision relative to its use of capacity, then it pays the other for the amount of extra capacity it uses at a rate equal to the incremental cost of the supplied infrastructure. This has the interpretation of a payment for transit in the case where one network provides no infrastructure. Under this rule, we show that usage-prices and the level of infrastructure are set at the same levels that the central planner chooses. Firms in competition still set price equal to the marginal cost of providing usage, but since marginal cost now includes the marginal cost of providing extra capacity, this price is the efficient one.

In our model, if the settlement charge is not regulated, firms will agree to a settlement charge that is above the incremental cost of infrastructure provided. Relative to the central planner's solution, this allows firms to increase profits at the expense of consumers. However, with fixed costs to cover and intensive retail competition, settlement charges above incremental cost are needed to allow firms to break-even. More fundamentally, with free-entry into the industry and with multiple firms competing to capture any settlement profit, we argue that the regulation of settlement charges is not appropriate.

---

[8] No congestion is optimal in this model since we ignore, for simplicity, issues of time-varying usage.

[9] Taken literately, the marginal costs in our model are essentially zero. Thus this result is consistent with the fact many Internet providers charge a flat monthly rental to retail customers, with usage prices set to zero.

In practice there is little use of such precise settlement payment schemes, given that they require the burdensome measurement of bi-directional traffic flows.[10] Nevertheless, one can interpret existing schemes as reflecting traffic usage in a less direct way. US transit providers typically offer monthly fees that are negotiated to reflect recent and expected future traffic flows.[11] Where the flow of traffic is roughly balanced, the cost of measuring flows would seem to be particularly inefficient. This suggests an alternative arrangement, which is in fact common place for interconnecting IAPs.

If firms only peer with other firms who provide equally to infrastructure and have similar demand and market shares, we show that ex-post there is no need for settlement contribution. Both firms take into account capacity constraints even when pricing in competition, since they are aware that if they become net users of the common infrastructure, they will be excluded from peering arrangements with the net providers. Prices and capacity provision will be efficient.

Our conclusions are not surprising where firms are asymmetric; if one firm is a net user of the other's network in a peering relationship, it is not equitably contributing to the costs of that shared infrastructure. However, excluding another smaller network from settlement-free peering has been claimed to be anti-competitive. As shown in this paper, both from a static efficiency point of view and a dynamic efficiency point of view, such behavior is likely to be welfare enhancing. Rules that encourage more investment and higher usage prices are indeed desirable. In contrast, regulator enforced settlement-free peering not only leads to expropriation, that arises when one firm uses its competitor's facilities without remuneration, but according to our analysis discourages investment and leads to even greater congestion on the Internet than that which presently exists.

The rest of the paper proceeds as follows. In Section 2 we set-up our model of network competition. Section 3 derives the central planner's desired solution for pricing and capacity provision in this model. We contrast this efficient solution with the equilibrium under settlement-free peering, which we derive in Section 4. Section 5 then develops efficient settlement payment schemes that lead to the central planner's solution. Finally, Section 6 concludes.

---

[10] Although sampling of traffic data offers a partial solution to this problem, it still adds to delay and packet loss.

[11] There are cases outside the US where providers have relied on direct traffic measurement to determine the price of interconnection. The largest Australian Internet backbone provider, Telstra, measures outgoing traffic at each point of interconnection providing a basis for its traffic variable transit prices.

## 2. Model set-up

We suppose there are two IAPs, denoted 1 and 2. Each IAP (or firm) provides customers access to a common infrastructure (the Internet) which it charges customers for usage of.[12] For simplicity we suppose there are a fixed number of potential customers in the population, but the usage of the infrastructure can vary with the price charged. We suppose that firm 1 provides $k_1$, and firm 2 $k_2$, of infrastructure, so that $k_1 + k_2$ is the total capacity available to the average customer. The cost to firm i of providing $k_i$ is denoted $f(k_i)$. Each unit of usage by a customer belonging to firm i, incurs a cost to firm i of $c$ (that is, the marginal cost is $c$). In addition, each customer that subscribes to firm i, imposes a fixed cost on firm i of $f$.

The cost of providing capacity, the fixed cost of serving additional customers, and the marginal cost of usage are easily interpreted. Firstly, the common cost of providing capacity $f(k_i)$ corresponds to the cost of laying cable or purchasing satellite bandwidth to service the average bandwidth requirements of customers, as well as the cost of routers. The fixed cost for serving each customer $f$ corresponds to the costs of adding to the modem pool at a point of presence, upgrading server capacity for e-mail and caching, as well as billing and providing customer service. Finally, the marginal cost of usage $c$ represents the additional expenses that an additional unit of usage imposes on the network. Given that we account for congestion explicitly in the paper, these costs are likely to be trivial, but are included for the sake of generality.

Each customer of firm i consumes $q_i$ units, where each unit can be thought of as a megabyte of data transfer.[13] This level of usage generates utility $u(q_i) + v_0 + \theta_i$, where the function $u$ is the same for all households and does not depend on which firm is used. The parameter $v_0$ represents a fixed surplus from being connected to either network, and $\theta_i$ measures the additional costs and benefits of belonging to a particular network (that relate to factors other than access to the common infrastructure), the value of which depends on the customer's particular tastes. Specifically, customers are endowed with a value of $x$ which is

---

[12] The issues described below are likely to be even more acute with more than two firms. Moreover, it should be noted that despite restricting ourselves to two firms, the model can still capture scenarios with intensive competition or little competition, by varying the parameter, $s$, discussed below.

[13] Customers get value from the transfer of megabytes, regardless of the direction of transfer (for example, regardless of whether the customer downloads information from another user or sends information to another user).

drawn from the uniform distribution on the interval [0,1]. If they subscribe to firm 1, they receive an extra benefit of

$$q_1 = \frac{b}{2s} + \frac{(1-x)}{s},$$

while if they subscribe to firm 2, they receive an extra benefit of

$$q_2 = \frac{x}{2s}.$$

This product differentiation set-up has been used for modelling competition in telecommunications (see Armstrong 1998 and Laffont et al. 1998). The introduction of $b$ follows Carter and Wright (1999), who use it in modelling network competition, to allow for the possibility that when facing equal prices, more customers might prefer firm 1 ($b > 0$) or firm 2 ($b < 0$). This could be because of the value added services provided by one of the IAPs, such as a smaller proportion of customers to modems, or because of the greater reputation that one firm has developed, possibly through brand development.

Given that households' marginal willingness to pay is known and the same for all households, firms can not do better than offer two-part tariffs.[14] Each firm charges a per-unit price $p_i$ and a lump-sum fee (or rental) $r_i$. The share of customers that choose to belong to firm 1 is then

$$s = \frac{1}{2} + \frac{b}{2} + s\,(w_1 - w_2),$$

where $w_i = v(p_i) - r_i$ is the net surplus offered to firm i's consumers and $v(p_i) = \max_q \{u(q_i) - p_i q_i\}$ is the level of indirect utility associated with usage. The firms' profit functions can be written

$$p_1 = s(p_1 - c)q(p_1) + s(r_1 - f) - f(k_1)$$

$$p_2 = (1-s)(p_2 - c)q(p_2) + (1-s)(r_2 - f) - f(k_2).$$

Since $k_1 + k_2$ is the total capacity available to the average customer, we also have the joint capacity constraint that $sq(p_1) + (1-s)q(p_2) \le k_1 + k_2$.

This says the weighted-average usage has to be less than the capacity provided. Of course, customers can attempt usage above capacity. Though in practice this causes some

[14] See for instance, Laffont et al. (1998, p.20).

usage to be delayed, thus lowering customers' utility, we model any usage above capacity as generating no utility.[15] We will later make explicit how usage is rationed in this situation.

## 3. Central planner's solution

The central planner chooses the variables $k_1, k_2, p_1, p_2, r_1, r_2$ to maximize the total of consumer and producer welfare, which is

$$W = sw_1 + (1-s)w_2 + \frac{s(1-s)}{2s} + \frac{bs}{2s} + \frac{1}{4s} + p_1 + p_2,$$

subject to the capacity constraint $sq_1 + (1-s)q_2 \leq k_1 + k_2$.[16] After some manipulation the following first order conditions can be derived (the derivations of this result is contained in Appendix 7.1).

$$\begin{aligned} p_1 &= c + f'(k_1) \\ p_2 &= c + f'(k_2) \\ f'(k_1) &= f'(k_2) \\ k_1 + k_2 &= q(c + f'(k_1)) \\ r_1 &= r_2 \end{aligned}$$

The most important feature to note of this solution, is that the marginal prices charged to customers take into account the additional cost of providing the capacity for the usage that is demanded. In this static model, it is not efficient to build more capacity than is needed by customers. Likewise, it is not efficient to under provide capacity, since then customers' demand would be quantity rationed. The most efficient solution is to use prices to signal to customers the true cost of additional capacity and so let customers indirectly choose the appropriate amount of capacity.

Another implication of this solution is that firm 1's share of the market will be $s = \frac{1}{2} + \frac{b}{2}$. This is driven by the fact the only asymmetry in the set-up between the two firms is the possible customer preference towards one of the two networks (that is, $b \neq 0$). The level of the lump-sum fees is not determined by the solution, since it is simply a transfer between consumers and producers. For reasonable parameter values, there is a range of values for rentals, $r_i$, which allow firms to break-even and leave customers willing to join. Without

---

[15] One justification for this simplification is that some customers will experience lower but positive utility from this delayed usage, while others may actually experience disutility due to their wasted time.

[16] We assume for the relevant parameter values, the consumers will want to participate.

specifying $r_i$, we cannot be sure of the exact division between consumer and producer surplus. However, if there are fixed costs to recover, then rentals must lie above the cost of connecting customers, $f$.

The solution above determines the level of total capacity provided, but how is the provision of this capacity divided between the two firms? If the costs of building infrastructure are simply proportional to the capacity provided, then (ignoring break-even constraints) it would be just as efficient to have one firm providing all capacity as to have both firms share the provision of capacity; so the division between $k_1$ and $k_2$ is not determined. If there are increasing marginal costs to each firm's production of the common infrastructure, then it would make sense to share the production. Alternatively, if there are significant economies of scale in the production of capacity, it would make sense to have one firm provide all infrastructure (this later case is a corner solution, so does not solve the first order conditions for $k_1$ and $k_2$ above, which are for interior solutions).

## 4. The decentralized solution under peering

Under a peering approach neither network owner pays for the use of the other's network. In this case, firms first choose their investment in capacity $k_i$, and given this they choose their prices, both decisions being made non-cooperatively.

### 4.1 Is the capacity constraint binding?

Taking capacity as given at some level $k_1+k_2$, we first see whether there is any equilibrium where we can ignore the capacity constraint. As shown in Appendix 7.2, the equilibrium in prices (where each firm sets its two-part tariff to maximize its profit, taking as given what the other firm is doing), implies

$$p_1 = c$$
$$p_2 = c$$
$$r_1 = f + \frac{s}{s}$$
$$r_2 = f + \frac{1-s}{s}.$$

With these prices, equilibrium profits are

$$p_1 = \frac{s^2}{s} - f(k_1)$$

$$p_2 = \frac{(1-s)^2}{s} - f(k_2),$$

where $s = \frac{1}{2} + \frac{b}{6}$. Clearly, each firm will want to choose the minimal amount of $k$ possible. However, total capacity $k_1 + k_2$ has to be enough to cover $sq(c) + (1-s)q(p)$, otherwise some customers are rationed. This suggests only two outcomes are possible in equilibrium with settlement free peering: either the capacity constraint is just binding or customers are rationed. We consider each in turn.

### 4.2 Capacity constraint just binding?

In this subsection we see whether there is any equilibrium in which the capacity constraint is just binding; that is, there is no rationing. Each firm sets its price taking into account the binding capacity constraint.

The Lagrangean for firm 1 is written

$$L_1 = s(p_1 - c)q_1 + s(r_1 - f) - f(k_1) + l_1(k_1 + k_2 - sq_1 - (1-s)q_2),$$

while the Lagrangean for firm 2 is

$$L_2 = (1-s)(p_2 - c)q_2 + (1-s)(r_2 - f) - f(k_2) + l_2(k_1 + k_2 - sq_1 - (1-s)q_2).$$

Maximizing each of these, taking $k_1$ and $k_2$ as given, we get the first order conditions, which are derived in Appendix 7.3.

$$p_1 = c + l_1$$
$$p_2 = c + l_2$$
$$r_1 = f + \frac{s}{s} - l_1 q_2$$
$$r_2 = f + \frac{1-s}{s} - l_2 q_1$$
$$k_1 + k_2 = sq_1 + (1-s)q_2.$$

Substituting these results back into the profit functions yields

$$p_1 = \frac{s^2}{s} + s\,l_1(q_1 - q_2) - f(k_1)$$

$$p_2 = \frac{(1-s)^2}{s} + (1-s)\,l_2(q_2 - q_1) - f(k_2).$$

Given that the capacity constraint is binding, $l_1 = l_2 > 0$. In this case, prices are above the marginal cost each firm incurs for usage. However, this is not an equilibrium. Consider what happens if one firm decreases its usage price by a small amount, with a corresponding adjustment to its rental to keep customers at the same level of utility. Note that the level of capacity has been set, and we are taking as given the other firm's price. The lower usage price will generate more demand, so that total usage demand will exceed capacity. Starting from the symmetric equilibrium above, firm 1's customers usage will slightly exceed half the capacity, while firm 2's customers demand exactly half the available capacity. Some of this additional usage is rationed (that is the customers receive no benefit from this additional usage). Since in any realistic rationing process all customers share in the rationing of usage (we give an example of such a process in the next section), users of firm 1 gain some additional non-rationed usage at the lower price. This extra benefit can be captured by firm 1 by raising the lump sum fee customers are charged, while still leaving the marginal consumer indifferent between the two networks. Thus a non-rationing equilibrium does not exist.[17] Firms in competition, under settlement free peering, under-price usage.

In fact, it is shown in the appendix, to a first approximation, the extra profit firm 1 gets from undercutting is proportional to $l_1$. That is, firm 1's incentive to under-price is proportional to the extra benefit it gets from relaxing the capacity constraint. Only when prices have been driven down to marginal cost is there no incentive to under-price. As previously argued, this involves rationing.

### 4.3 Rationing equilibrium

In this subsection we characterize the equilibrium with rationing. We suppose that usage above capacity is rationed in a random fashion that neither firm can control[18] (other than through pricing[19]). Being rationed in this context means that the attempted usage is denied and no utility is received. The proportion of a customer's usage that gives them no utility (due to congestion) is assumed to be equal across all customers, and is such that the total utility bearing usage is equal to total capacity. This set-up is a simple way of capturing the

---

[17] The proof is contained in Appendix 7.3.

[18] It is unlikely, in a practical sense, that any firm can control rationing by transporting its own packets and dropping those of its competitors.

[19] MacKie-Mason and Varian (1994) provide a model of how retail prices can be used to ease congestion.

idea that with excess demand for capacity, there will be delay for all consumers, with lower utility due to delay. In this case, a customer's attempts at using the Internet are proportionately split between successful and unsuccessful usage.

Let $q^r(p_1)$ denote the amount rationed to customers of firm 1 and $q^r(p_2)$ denote the amount rationed to customers of firm 2. According to this proportional rationing, these can be described as $q_1^r = gq_1$ and $q_2^r = gq_2$, where the proportion of usage that is rationed is $1-g$ and

$$g = \frac{(k_1 + k_2)}{sq_1 + (1-s)q_2}.$$

Note that under this set-up, total rationed usage $(sq_1^r + (1-s)q_2^r)$ adds up to total capacity $(k_1 + k_2)$. Since customers' utility depends only on successful usage,[20] the share of customers that belong to firm 1 is

$$s = \frac{1}{2} + \frac{b}{2} + s\,(w_1 - w_2),$$

where $w_i = z(p_i, g) - r_i$ is the net surplus offered to firm i's consumers and $z(p_i, g) = \max_{q_i}\{u(gq_i) - gp_iq_i\}$.

The firms' profit functions are now

$$p_1 = s(p_1 - c)q_1^r + s(r_1 - f) - f(k_1)$$
$$p_2 = (1-s)(p_2 - c)q_2^r + (1-s)(r_2 - f) - f(k_2).$$

In Appendix 7.4, the first order conditions are shown to be

$$p_1 = p_2 = c$$
$$r_1 = \frac{s}{s} + f$$
$$r_2 = \frac{1-s}{s} + f.$$

Substituting the first order conditions back into the profit functions implies

---

[20] Note that when an individual customer chooses $q_i$ it treats $g$ as a constant since it is too small to take into account its own affect on the amount of rationing. However, when a firm changes its usage prices, $g$ will vary.

$$p_1 = \frac{s^2}{s} - f(k_1)$$

$$p_2 = \frac{(1-s)^2}{s} - f(k_2)$$

As in the case where we ignored the capacity constraint, the rationing equilibrium leads to firms ignoring the cost of capacity in their usage price. Moreover, firms again want to contribute the minimal amount of capital to the infrastructure as is necessary to obtain customer participation (recall customers' utility depends on the rationed level of quantity). Thus total capacity is lowered until the customers participation constraint is binding.

Any combinations of $k_1$ and $k_2$ that satisfy this minimal level of total capacity ($k_1 + k_2$) will do. However, if the firms act in a decentralized and simultaneous way, it is unlikely they end up providing even this low level of infrastructure; each will hope the other provides more, while itself providing less.

A classic free-rider problem emerges.[21] Each firm uses the other firm's capacity for free and thus receives maximal profits when the other firm provides all of the capital. This free-rider problem could surface in two ways in practice. In one case, investment is delayed, since each firm waits for the other to build the capacity. In any realistic setting, investment decisions are made through time, and so there will be complicated dynamics and game playing between firms to try to avoid being the one to provide the majority of the capacity needed. In such a world, coordination failures and delay are likely, given that both have an individual incentive to reach an equilibrium where the other firm does most of the investment. In the other case, one firm takes the lead, makes the investment, and under settlement free peering allows the other firm to use its facilities. This later case is clearly inequitable. If the competitor is allowed to use the incumbent's capacity for free, this is tantamount to a form of expropriation. It suggests some payment should be required when there is an imbalance in how much a firm uses capacity versus how much capacity it provides.

---

[21] MacKie-Mason and Varian (1996) provide two examples of such free-riding. In one, a new Internet Provider hosts a number of World Wide Web servers near an existing IAP and purchases a very short connection to the IAP. Such an Internet Provider could receive considerable revenue from customers whose World Wide Web servers are being hosted, while imposing a considerable load on the existing networks, without every delivering much incoming traffic. In the other example, networks provide substantial backbone infrastructure for transit between two networks, without receiving a direct share of the end-user payment for either end.

## 5. Settlement payments

In this section, we continue to assume firms first choose their investment in capacity, and given this they choose their prices, both decisions being made non-cooperatively. However, we also suppose firms put in place settlement rules before they decide how much to invest in capacity. As long as firms can measure the level of usage from their customers, we show there is a simple rule that leads to the efficient outcome. The rule states that if one firm uses more capacity than it provides, it pays the other firm a rate $t$ on this difference. We show that when $t$ is set at the incremental cost of the capital provided, this rule leads to the social planner's efficient outcome discussed in Section 3. We also show that, starting from this efficient solution, both firms will prefer to set a higher rate for these payments. This raises their profit at the expense of consumers. Despite this later result, we give reasons why allowing IAPs to set their own charges for interconnection is likely to be desirable in the Internet industry.

Under the settlement rule there are two possibilities to consider. Either one of the firms under-provides relative to its usage (case 1) or both firms each provide exactly enough to cover their own usage (case 2). We consider each case in turn.

### Case 1: The asymmetric solution

Suppose, without loss of generality, firm 1 provides more capacity than it uses, while firm 2 provides less than it uses. According to the rule above, firm 1's profit function is now

$$\pi_1 = s(p_1 - c)q_1 + s(r_1 - f) + t[(1-s)q_2 - k_2] - f(k_1),$$

while firm 2's profit is

$$\pi_2 = (1-s)(p_2 - c)q_2 + (1-s)(r_2 - f) - t[(1-s)q_2 - k_2] - f(k_2).$$

Firm 1 chooses $p_1$ and $r_1$ to maximize $\pi_1$ subject to the constraint that firm 2's demand for capacity can be meet. Since firm 1 has $k_1 - sq_1$ capacity remaining after its own use, it faces the constraint that $(1-s)q_2 - k_2 \leq k_1 - sq_1$. Firm 2 simply chooses $p_2$ and $r_2$ to maximize $\pi_2$. The first order conditions are

$$p_1 = c + \lambda$$
$$p_2 = c + t$$
$$r_1 = f + \frac{\sigma}{s} + (t - \lambda)q_2$$
$$r_2 = f + \frac{1-s}{\sigma}$$
$$k_1 + k_2 = sq_1 + (1-s)q_2 .$$

Under the rule that $t = l$ , we get

$$p_1 = \frac{s^2}{s} + lk_1 - f(k_1)$$

$$p_2 = \frac{(1-s)^2}{s} + lk_2 - f(k_2).$$

If each firm set $k_1$ and $k_2$ independently, it will choose $l = f'(k_1)$ so $f'(k_1) = f'(k_2)$. Thus, ignoring the level of the lump-sum payments, the equilibrium satisfies all the conditions of the central planner's solution

$$p_1 = c + f'(k_1)$$

$$p_2 = c + f'(k_2)$$

$$f'(k_1) = f'(k_2)$$

$$k_1 + k_2 = q(c + f'(k_1)).$$

The key result is that firm 1 (the capacity provider) should charge firm 2 (the capacity user), based on the capacity used, at a rate equal to the incremental cost of providing infrastructure $t = f'(k_1)$. Clearly any charge in the other direction would not be appropriate (and in fact would lead to an inefficient outcome). This solution continues to hold even in the corner solution, in which it is efficient for firm 1 to provide the entire infrastructure. Then $p_1 = c + f'(k_1)$, $p_2 = c + f'(k_1)$ and $k_1 = q(c + f'(k_1))$ and the payment from firm 2 to firm 1 is $(1-s)q(c + f'(k_1))f'(k_1)$.

Unlike the solution studied in Section 4.2, the solution here is indeed an equilibrium. Neither firm has an incentive to under-price, given the price the other firm sets. If firm 2 lowers its price, the additional usage cost is $c + t$ per-unit. By lowering its price it receives less than $c + f'(k_2)$ per-unit. Since $t = l = f'(k_1) = f'(k_2)$, it will face a loss on each additional increment it sells. If firm 1 lowers its price, it can sell more, but only by excluding access to firm 2. Since the per-unit amount $t$ it receives from firm 2, more than covers its retail margin (which must be less than $f'(k_1)$ after dropping its price), it does not have an incentive to lower its price.

An important question is whether firms will want to adopt such a settlement regime voluntarily or whether regulation is needed to achieve the efficient level of $t$ above; that is, what level of $t$ will the firms want to agree on? Although we do not characterize the optimal level of $t$ chosen by the two firms (since this will involve solving the bargaining problem that arises between the two firms), we can show that both firms will prefer a level of $t$ that is above the incremental cost of infrastructure provided by firm 1. The effect is to raise overall prices for each firm. Firm 2's usage price increases, but its fixed price remains unchanged. Firm 1 actually lowers its usage price, but raises its fixed price by more than the offsetting amount. The net result is to reallocate surplus from consumers to the firms. In fact, in Appendix 7.5, we show that starting from the efficient level of $t$, each firm's incentive to raise $t$ is proportional to the level of infrastructure that it provides.[22] Thus if one of the firms provides no infrastructure (the pure transit case), then its preferred level of $t$ is precisely the efficient one. The infrastructure provider will, however, strongly prefer a higher $t$, and so the result of bargaining is likely to be a $t$ above the incremental cost of the capital provided.

Despite this, there are strong reasons to believe that the incentive firms have to charge above cost for settlement is not a long-term problem. With fixed costs to cover and intensive retail competition, above incremental cost settlement charges would be needed to allow firms to break-even. This inevitably moves us away from the social planner's first-best optimum in section 3, but not necessarily the Ramsey optimum that is more relevant in this case. The Ramsey optimum will involve settlement charges such that infrastructure providers can just break-even. There are at least two reasons why regulation may not be needed to achieve such charges. In reality there are likely to be multiple IAP's, not just two, and so there will be limits on the ability of any individual firm from raising settlement charges above cost (this would lead to competition to gain settlement revenue). More generally, with free entry into the industry, if settlement charges were sustained above the cost of providing the associated capital, there would be more entry and infrastructure building until this profit opportunity is eliminated. Given that regulation is particularly difficult to implement in a rapidly evolving industry such as the Internet, these results suggest policy makers should be very cautious about attempting to regulate the level of settlement charges.

**Case 2: The symmetric solution**

An alternative outcome of using settlement payments might be that both firms exactly cover their need for capacity, and so ex-post there is no settlement payment between them. However, according to the settlement rule above, they must take into account if they do

[22] The intuition for this result for firm 2 is straightforward. The increase in profits from its higher price is offset by the higher settlement payments, except to the extent it provides its own capacity and so does not pay settlement payments. For firm 1, the benefits arise from the demand it serves itself plus the share of firm 2's demand that it services through its excess capacity. In total the benefits are thus proportional to its total amount of capacity.

become a net user of capacity, they will have to make the appropriate settlement payment described previously. In this case, firm 1 maximizes

$$p_1 = s(p_1 - c)q_1 + s(r_1 - f) - f(k_1) \text{ subject to } sq_1 \leq k_1,$$

while firm 2 maximizes

$$p_2 = (1-s)(p_2 - c)q_2 + (1-s)(r_2 - f) - f(k_2) \text{ subject to } (1-s)q_2 \leq k_2.$$

This set-up can also be interpreted as the outcome if networks reach an agreement that they will only interconnect if they provide equal contributions to the infrastructure and they are of roughly equal size. The first order conditions that result are

$$p_1 = c + l_1$$
$$p_2 = c + l_2$$
$$r_1 = f + \frac{s}{s}$$
$$r_2 = f + \frac{(1-s)}{s}$$
$$sq_1 = k_1$$
$$(1-s)q_2 = k_2.$$

Substituting these back into the profit functions yields

$$p_1 = \frac{s^2}{s} + l_1 k_1 - l_1 f(k_1)$$
$$p_2 = \frac{(1-s)^2}{s} + l_2 k_2 - f(k_2).$$

Maximizing profit with respect to each firm's choice of capital provision, implies $l_1 = f'(k_1)$ and $l_2 = f'(k_2)$. Since it is only efficient for both firms to fully provide for their own usage when both firms have equal costs of providing for additional capacity, it must also be that $f'(k_1) = f'(k_2)$. In this case, the solution is precisely the solution to the central planner's problem discussed in Section 3 (apart from the fact the lump-sum charges do not have to be equal in the competitive equilibrium). Furthermore, it is clear this is an equilibrium. If either firm lowers its price, it will become a net user, in which case it will make a payment for the extra usage, at a rate greater than the margin generated by the lower price.

## 6. Conclusion

We have found that there is little incentive for asymmetric firms to invest in Internet infrastructure under a regulated regime of settlement-free peering. While enforced peering may appear to have immediate consumer benefits through lower usage prices, this is shown to be not the case; the resulting prices just cause excessive congestion. More importantly, a regulator can expect this congestion problem will worsen in the medium- to long-term as an inefficiently low level of investment in infrastructure is carried out. Allowing IAPs to charge settlement fees, or allowing them to peer and exclude firms that are net users of their infrastructure solves these problems. The first-best outcome is achieved if settlement fees are set at the incremental cost of capital provided. While firms will have an incentive to agree to higher fees, we argued that regulating settlement fees is not likely to be beneficial.

## 7. Appendix

This appendix contains derivations of the key results stated in the sections above.

### 7.1 The central planner's solution

The central planner chooses $k_1, k_2, p_1, p_2, r_1, r_2$ to maximize the total of consumer and producer welfare, which is

$$W = sw_1 + (1-s)w_2 + \frac{s(1-s)}{2s} + \frac{bs}{2s} + \frac{1}{4s} + p_1 + p_2,$$

subject to the constraint that $sq_1 + (1-s)q_2 \leq k_1 + k_2$. The Lagrangean is

$$L = sw_1 + (1-s)w_2 + \frac{s(1-s)}{2s} + \frac{bs}{2s} + \frac{1}{4s} + s(p_1 - c)q(p_1) + s(v_1 - w_1 - f) - f(k_1)$$
$$+ (1-s)(p_2 - c)q(p_2) + (1-s)(v_2 - w_2 - f) - f(k_2) + l(k_1 + k_2 - sq_1 - (1-s)q_2),$$

where profits have been written in terms of $w_1$ and $w_2$ rather than $r_1$ and $r_2$; that is, these now become the choice variables. Differentiating with respect to $p_1$ and $p_2$ yields the first order conditions

$$\frac{\partial L}{\partial p_1} = s(p_1 - c)q'(p_1) - lsq'(p_1) = 0$$
$$\frac{\partial L}{\partial p_2} = (1-s)(p_2 - c)q'(p_2) - l(1-s)q'(p_2) = 0,$$

where we have used that $v'(p_i) = -q(p_i)$. Solving gives the solution for prices

$$p_1 = c + l$$
$$p_2 = c + l.$$

Differentiating the Lagrangean with respect to $w_1$ and $w_2$ yields

$$\frac{\partial L}{\partial w_1} = s - s + s\ \frac{\partial L}{\partial s} = 0$$

$$\frac{\partial L}{\partial w_2} = (1-s) - (1-s) - s\ \frac{\partial L}{\partial s} = 0,$$

which implies

$$\frac{\partial L}{\partial s} = w_1 - w_2 + \frac{(1-2s)}{2s} + \frac{b}{2s} + (p_1 - c)q(p_1) + (v_1 - w_1 - f)$$
$$- (p_2 - c)q(p_2) - (v_2 - w_2 - f) - l\,(q_1 - q_2)$$
$$= 0$$

Substituting in that $p_i = c + l$ from above and cancelling common terms (note $q_1 = q_2$ and $v_1 = v_2$) gives

$$\frac{(1-2s)}{2s} + \frac{b}{2s} + f'(k_1)q(p_1) - f'(k_2)q(p_2) = 0\,.$$

Differentiating the Lagrangean with respect to $k_1$ and $k_2$ yields

$$\frac{\partial L}{\partial k_1} = l - f'(k_1) = 0$$

$$\frac{\partial L}{\partial k_2} = l - f'(k_2) = 0.$$

This implies that in the efficient solution the capacity constraint is just binding, since capacity is costly to build $l = f'(k_i) > 0$. It also implies that $f'(k_1) = f'(k_2)$, which says the additional cost of firm 1 building additional capacity is equal to the additional cost of firm 2 building additional capacity. If firm 1 is subject to greater economies of scale up to some point, than firm 2, then it may be efficient for firm 1 to build more than half the capacity. Finally, the last three equations above combined yield the result that $s = \frac{1}{2} + \frac{b}{2}$. This then implies $r_1 = r_2$. The remaining results of the section follow in a straightforward manner.

### 7.2 Is the capacity constraint binding?

In this case each firm maximizes its profit; these are (written in terms of $w_1$ and $w_2$)

$$p_1 = s(p_1 - c)q(p_1) + s(v_1 - w_1 - f) - f(k_1)$$

$$p_2 = (1-s)(p_2 - c)q(p_2) + (1-s)(v_2 - w_2 - f) - f(k_2)\,.$$

The first order conditions are then

$$\frac{\partial p_1}{\partial p_1} = s(p_1 - c)q'(p_1) = 0$$
$$\frac{\partial p_2}{\partial p_2} = (1-s)(p_2 - c)q'(p_2) = 0$$
$$\frac{\partial p_1}{\partial w_1} = -s + s\left((p_1 - c)q(p_1) + (v_1 - w_1 - f)\right) = 0$$
$$\frac{\partial p_2}{\partial w_2} = s - s\left(-(p_2 - c)q(p_2) - (v_2 - w_2 - f)\right) = 0,$$

where again we have used that $v'(p_i) = -q(p_i)$. Combining these first order conditions, and noting that $v_i - w_i = r_i$, generates the results in the text. The rest of the results for this subsection follow immediately.

### 7.3 Capacity constraint just binding?

The Lagrangean for firm 1 is written

$$L_1 = s(p_1 - c)q_1 + s(v_1 - w_1 - f) - f(k_1) + l_1(k_1 + k_2 - sq_1 - (1-s)q_2),$$

while the Lagrangean for firm 2 is

$$L_2 = (1-s)(p_2 - c)q_2 + (1-s)(v_2 - w_2 - f) - f(k_2) + l_2(k_1 + k_2 - sq_1 - (1-s)q_2).$$

Differentiating, we get the following conditions

$$\frac{\partial p_1}{\partial p_1} = s(p_1 - c)q'(p_1) - s\,l_1 q'(p_1) = 0$$
$$\frac{\partial p_2}{\partial p_2} = (1-s)(p_2 - c)q'(p_2) - (1-s)\,l_2 q'(p_2) = 0$$
$$\frac{\partial p_1}{\partial w_1} = -s + s\left((p_1 - c)q(p_1) + (v_1 - w_1 - f)\right) - l_1(q_1 - q_2) = 0$$
$$\frac{\partial p_2}{\partial w_2} = s - s\left(-(p_2 - c)q(p_2) - (v_2 - w_2 - f)\right) - l_2(q_1 - q_2) = 0.$$

Combining these equations gives the first order conditions

$$p_1 = c + l_1$$
$$p_2 = c + l_2$$
$$r_1 = f + \frac{s}{s} - l_1 q_2$$
$$r_2 = f + \frac{1-s}{s} - l_2 q_1$$
$$k_1 + k_2 = sq_1 + (1-s)q_2.$$

The profit functions arising from substituting these first order conditions into the original profit functions are

$$p_1 = \frac{s^2}{s} + s\,l_1(q_1 - q_2) - f(k_1)$$
$$p_2 = \frac{(1-s)^2}{s} + (1-s)\,l_2(q_2 - q_1) - f(k_2).$$

Now to show that this is not an equilibrium, consider the case firm 1 charges a slightly lower price $p_1' = c + l_1 - dp_1$ and adjusts $r_1$ so that the share of customers joining network 1 remains unchanged. By looking at the marginal consumer (the consumer who is just indifferent between the two networks), this occurs when

$$\left[u(q_1 + dq_1) - (p_1 - dp_1)(q_1 + dq_1) - (r_1 + dr_1) + \frac{1-s}{2s}\right] - \left[u(q_2 - dq_2) - p_2(q_2 - dq_2) - r_2 + \frac{s}{2s}\right]$$
$$= \left[u(q_1) - p_1 q_1 - r_1 + \frac{1-s}{2s}\right] - \left[u(q_2) - p_2 q_2 - r_2 + \frac{s}{2s}\right]$$
$$= 0,$$

where $dr_1$ is the change in the fixed fee charged by firm 1. Re-arranging we get

$$dr_1 = [u(q_1 + dq_1) - u(q_1)] - [u(q_2 - dq_2) - u(q_2)] + q_1 dp_1 - p_1 dq_1 + dp_1 dq_1 - p_2 dq_2 .$$

Thus the change in profits of firm 1 is

$$dp_1 = s(p_1 - dp_1 - c)(q_1 + dq_1) - s(p_1 - c)q_1 + s dr_1$$
$$= -scdq_1 - sp_2 dq_2 + s[u(q_1 + dq_1) - u(q_1) + u(q_2) - u(q_2 - dq_2)].$$

Since $q_1$ is chosen so that $u'(q_1) = p_1$ and $q_2$ is chosen so that $u'(q_2) = p_2$, then to a first approximation

$$dp_1 = s(p_1 - c)dq_1$$
$$= s\,l_1 dq_1.$$

Since the original price $p_1 = c + l_1$ is greater than $c$ (i.e. $l_1 > 0$), this condition says the firm can profit by lowering its price. That is, there is no equilibrium where the capacity is just binding. Firms in competition, under settlement free peering, under-price usage.

### 7.4 Rationing equilibrium

The firms profit functions are

$$p_1 = s(p_1 - c)q_1^r + s(r_1 - f) - f(k_1)$$

$$p_2 = (1-s)(p_2 - c)q_2^r + (1-s)(r_2 - f) - f(k_2).$$

It is easier to find the first order conditions for this case if we first transform the problem so the firms' pick the total tariff charged and the quantity of rationed usage that customers will receive. This is equivalent to letting the firm pick the price and rental above. Thus the profit functions can be re-written

$$p_1 = s(T_1 - cq_1^r - f) - f(k_1)$$

$$p_2 = (1-s)(T_2 - cq_2^r - f) - f(k_2),$$

where $s = \frac{1}{2} + \frac{b}{2} + s\,(w_1 - w_2)$ and $w_i = u(q_i^r) - T_i$, where $T_i = p_i q_i^r + r_i$. Since $q_i$ is chosen to $\max_{q_i}\{u(q_i^r) - T_i\}$ we get that $u'(q_i^r) = p_i$.

Differentiating the profit function for firm 1 with respect to $T_1$ and $q_1^r$, we get the following first order conditions

$$T_1 = \frac{s}{s} + cq_1^r + f$$

$$T_1 = \frac{cs}{s\,u'(q_1^r)} + cq_1^r + f.$$

Equating these two results implies $u'(q_1^r) = c$. Using the result above that $u'(q_i^r) = p_i$, we get that $p_1 = c$, and so $r_1 = \frac{s}{s} + f$. Applying the same steps to firm 2, we end up with the first order conditions as

$$p_1 = p_2 = c$$

$$r_1 = \frac{s}{s} + f$$

$$r_2 = \frac{1-s}{s} + f,$$

and the profits in equilibrium as

$$p_1 = \frac{s^2}{s} - f(k_1)$$

$$p_2 = \frac{(1-s)^2}{s} - f(k_2).$$

### 7.5 Settlement payments: case 1

The profit functions (re-written) are

$$p_1 = s(p_1 - c)q_1 + s(v_1 - w_1 - f) + t[(1-s)q_2 - k_2] - f(k_1)$$

$$p_2 = (1-s)(p_2 - c)q_2 + (1-s)(v_2 - w_2 - f) - t[(1-s)q_2 - k_2] - f(k_2).$$

Firm 1 chooses $p_1$ and $w_1$ to maximize $p_1$ subject to the constraint that

$$(1-s)q_2 - k_2 \le k_1 - sq_1.$$

Firm 2 simply chooses $p_2$ and $w_2$ to maximize $p_2$. Differentiating we get the following conditions

$$\frac{\partial L}{\partial p_1} = sq_1 + s(p_1 - c)q'(p_1) - sq_1 - s\,l\,q'(p_1) = 0$$

$$\frac{\partial L}{\partial w_1} = -s + s\,[(p_1 - c)q_1 + (v_1 - w_1 - f) - tq_2 - l\,q_1 + l\,q_2] = 0$$

$$\frac{\partial p}{\partial p_2} = (1-s)q_2 + (1-s)(p_2 - c)q'(p_2) - (1-s)q_2 - (1-s)tq'(p_2) = 0$$

$$\frac{\partial p}{\partial w_2} = -(1-s) - s\,[-(p_2 - c)q_2 - (v_2 - w_2 - f) + tq_2] = 0,$$

plus the binding constraint $(1-s)q_2 - k_2 = k_1 - sq_1$. Solving these equations and simplifying yields the first order conditions

$$p_1 = c + l$$

$$p_2 = c + t$$

$$r_1 = f + \frac{s}{s} + (t - l)q_2$$

$$r_2 = f + \frac{1-s}{s}.$$

Substituting these conditions back into the profit functions yields

$$p_1 = \frac{s^2}{s} + s\,l\,q_1 + s(t - l)q_2 + t[(1-s)q_2 - k_2] - f(k_1)$$

$$p_2 = \frac{(1-s)^2}{s} + (1-s)tq_2 - t(1-s)q_2 + tk_2 - f(k_2)\,.$$

Substituting in the binding constraint from above into firm 1's profit function, and simplifying both firms' profit functions we get

$$p_1 = \frac{s^2}{s} + s(t - l)(q_2 - q_1) + tk_1 - f(k_1)$$

$$p_2 = \frac{(1-s)^2}{s} + tk_2 - f(k_2).$$

Then if we choose $t = l$ these reduce to

$$p_1 = \frac{s^2}{s} + l k_1 - f(k_1)$$

$$p_2 = \frac{(1-s)^2}{s} + l k_2 - f(k_2).$$

If each firm set $k_1$ and $k_2$ independently, it chooses $l = f'(k_1)$ and $l = f'(k_2)$, which implies $f'(k_1) = f'(k_2)$. Since $p_1 = c + l = c + t = p_2$, this implies $v_1 = v_2$. The share of customers that belong to firm 1 is $s = \frac{1}{2} + \frac{b}{2} + s\,[(v_1 - r_1) - (v_2 - r_2)]$. Solving for s we get $s = \frac{1}{2} + \frac{b}{6}$, and thus the equilibrium satisfies the conditions of the central planner's solution, except that rentals are not the same across firms.

In case that $t$ and $l$ are not restricted to be equal we get the following results, summarized from above:

$$p_1 = c + l$$

$$p_2 = c + t$$

$$r_1 = f + \frac{s}{s} + (t - l)\,q_2$$

$$r_2 = f + \frac{1-s}{s}$$

$$s = \frac{1}{2} + \frac{b}{6} + \frac{s}{3}\left[v_1(c + l) - v_2(c + t) - (t - l)\,q_2\right]$$

$$k_1 + k_2 = sq_1 + (1 - s)\,q_2$$

and

$$p_1 = \frac{s^2}{s} + s(t - l)\,(q_2 - q_1) + tk_1 - f(k_1)$$

$$p_2 = \frac{(1-s)^2}{s} + tk_2 - f(k_2).$$

We would like to know whether the firms will voluntarily agree to set $t = l$, the efficient solution. To answer this question, we differentiate $p_1$ and $p_2$ with respect to $t$ and evaluate at the efficient solution $t = l$. In doing so we can ignore $\frac{dk_1}{dt}$ and $\frac{dk_2}{dt}$ terms. If the choice of $k_1$ and $k_2$ is made after the choice of $t$, then since infrastructure is chosen optimally (that is, $\frac{dp_1}{dk_1} = 0$ and $\frac{dp_2}{dk_2} = 0$), then these terms will have no effect on profits. If the choice of $k_1$ and $k_2$ is made before the choice of $t$ is known, then changing $t$ does not affect $k_1$ and $k_2$, and so again we can ignore the two terms. Thus our results here apply whether the choice of $t$ is being made ex-ante or ex-post. Differentiating profits with respect to settlement charges we get

$$\frac{dp_1}{dt} = \left[\frac{2s}{s} + (t - l)\,(q_2 - q_1)\right]\frac{ds}{dt} + s(q_2 - q_1)\left(1 - \frac{dl}{dt}\right)$$

$$+ s(t - l)\left(\frac{\partial q_2}{\partial p_2} - \frac{\partial q_1}{\partial p_1}\frac{dl}{dt}\right) + k_1$$

$$\frac{dp_2}{dt} = \frac{-2(1-s)}{s}\frac{ds}{dt} + k_2.$$

To evaluate the derivatives we note

$$\frac{ds}{dt} = \frac{s}{3}\left[(q_2 - q_1)\frac{dl}{dt} - (t - l)\frac{\partial q_2}{\partial p_2}\right],$$

which equals zero when $t = l$. Starting from the point $t = l$, an increase in $t$ causes an equal rise in overall prices of the two firms, so that the market share does not change at the margin. To understand why note an increase in $t$ by $dt$, raises $r_1$ by $q_2(dt - dl)$. This is partially offset by the decrease in $p_1$ by $dl$ (it can be shown that at $t = l$, $dl < 0$). The net effect on consumers who choose firm 1 is $q_1 d\lambda - q_2 d\lambda - q_2 dt$. This compares to the net effect on consumers who choose firm 2, which is $-q_2 dt$. Provided $q_1 = q_2$, which will be the case at the point $t = \lambda$, the net effect on consumers is the same across the two firms $(-qdt)$. Thus, at the margin, the market share will not depend on $t$, and so we get the result that

$$\left(\frac{dp_1}{dt}\right)_{t=l} = k_1 \qquad \text{and} \qquad \left(\frac{dp_2}{dt}\right)_{t=l} = k_2.$$

### 7.6 Settlement payments: case 2

In this case, each firm maximizes a Lagrangean function. Firm 1 maximizes

$$L_1 = s(p_1 - c)q_1 + s(v_1 - w_1 - f) + l_1(k_1 - sq_1) - f(k_1),$$

while firm 2 maximizes

$$L_2 = (1-s)(p_2 - c)q_2 + (1-s)(v_2 - w_2 - f) + l_2(k_2 - (1-s)q_2) - f(k_2).$$

First order conditions are easily derived as

$$p_1 = c + l_1$$
$$p_2 = c + l_2$$
$$r_1 = f + \frac{s}{s}$$
$$r_2 = f + \frac{1-s}{s},$$

plus the binding constraints $sq_1 = k_1$ and $(1-s)q_2 = k_2$. The profit functions are then

$$p_1 = \frac{s^2}{s} + s\,l_1 q_1 - f(k_1)$$

$$p_2 = \frac{(1-s)^2}{s} + (1-s)\, l_2 q_2 - f(k_2)\,.$$

Substituting in the constraints we get

$$p_1 = \frac{s^2}{s} + l_1 k_1 - f(k_1)$$

$$p_2 = \frac{(1-s)^2}{s} + l_2 k_2 - f(k_2)\,.$$

When each firm chooses the level of $k_i$ to maximize its profit we get the result in the main text.

## 8. References


Armstrong, M. 1998. "Network Interconnection," *The Economic Journal*, 108, pp. 545-564.

Baake, P. and Wichmann, T. 1999. "On the Economics of Internet Peering," *Netnomics*, 1, pp. 89-105.

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

Commercial Internet Exchange Association (CIX). 1995. "*White Paper: A Telecommunications Policy Framework for Internet Service Providers*," March (www.cix.org).

Ergas, H. 1999. "*Internet Peering: A Case Study of the ACCC's Use of its Powers Under Part XIB of the Trade Practices Act, 1974*," Network Economics Consulting Group.

Huston, G. 1999. "*ISP Survival Guide*," John Wiley & Sons Inc.: New York.

Laffont, J.-J, Rey, P., and J. Tirole. 1998. "Network Competition: I. Overview and Nondiscriminatory Pricing," *Rand Journal of Economics*, 29. pp. 1-37.

Mackie-Mason, J.K. and H.R. Varian. 1994. "Pricing Congestible Network Resources," *IEEE Journal of Selected Areas in Communications*, 13(7), pp. 1141-49.

Mackie-Mason, J.K. and H.R. Varian. 1996. "Economic FAQs about the Internet," *The Journal of Economic Perspectives*, 8(3), pp. 75-96

Srinagesh, P. 1997. "Internet Cost Structures and Interconnection Agreements," in *Internet Economics*, ed. McKnight, L. and Bailey, J. MIT Press: USA.

TARIFICA. 1998. "The Network Economics of the Internet," *Telecoms Strategies & Trends*, January (1).