**8 May, 1999**

# Internet Peering: A Case Study of the ACCC's Use of its Powers Under Part XIB of the Trade Practices Act, 1974.

Henry ERGAS

Iain Little and Tim Makinson provided assistance in the preparation of this paper, as did Julian Wright of the University of Auckland, who developed the model used in the paper. Additionally, Deena Shiff, Mitchell Landrigan, David Field and Geoff Huston of Telstra provided valuable information and advise. Ramin Marzbani was extremely generous with advice and data. The author has provided advice to Telstra on a range of issues, but not on those discussed here. The views expressed here are those of the author alone, as is the responsibility for any remaining errors.

# TABLE OF CONTENTS

## A Introduction

On May 27, 1998, the Australian Competition and Consumer Commission ("ACCC") determined that certain conduct by Telstra in the supply of Internet services constituted a contravention of the Competition Rule pursuant to s151AJ(2) of the *Trade Practices Act, 1974* ("the Act"). A Competition Notice was consequently issued, and was to come into force on June 5, 1998.[1] This was the first such Notice issued.

The Notice specified the conduct allegedly in contravention of the Competition Rule as follows:

> §20. Save as to an agreement with OzEmail Limited made on or about 22 May 1998, Telstra charges, and has charged at all material times, for Access Provider Services provided to other Internet Access Providers while at the same time not paying for or otherwise compensating other Internet Access Providers for Access Provider Services it receives from other Internet Access Providers ('Conduct').

This paper examines the background to and effects of the Competition Notice. The structure of the paper is as follows. First, the legislative background is set out (section **B**). This is followed by a brief introduction to the Internet (section **C**) and to interconnection arrangements in the Internet (section **D**). Section **E** examines the relevant markets and whether Telstra had power in those markets. Section **F** considers whether the conduct involved taking advantage of market power, section **G** whether it had an anti-competitive effect and section **H** whether the remedies promoted or lessened competition. Section **I** concludes.

## B The legislative background

The Competition Rule is set out in s151AK(1) of Part XIB of the Act. This section provides:

> A carrier or carriage service provider must not engage in anti-competitive conduct.

Section 151AJ(2) provides:

> A carrier or carriage service provider engages in anti-competitive conduct if the carrier or carriage service provider:
>
> (a) has a substantial degree of power in a telecommunications market; and

---

[1] Ultimately, the Notice did not come into effect, as Telstra entered into agreements covering the relevant issues with Optus and connect.com.au.

(b) takes advantage of that power with the effect, or likely effect, of substantially lessening competition in that or any other telecommunications market.

Section 151AJ(3) provides:

> A carrier or carriage service provider engages in anti-competitive conduct if the carrier or carriage service provider:
>
> (a) engages in conduct in contravention of section 45, 45B, 46, 47 or 48; and
>
> (b) the conduct relates to a telecommunications market.

Proof of the threshold test of substantial market power in a telecommunications market is required for this provision to be triggered by conduct which would be in contravention of s46 of the Act requires, as it is for the triggering of section 151AJ(2). Proof is also required of a "taking advantage" of that power. However, while s46 requires proof that power was used for one of the proscribed purposes set out in s46(1)(a), (b) or (c), this requirement has been omitted from s151AJ(2). Rather, the Competition Rule can be breached by conduct that has the effect or likely effect of substantially lessening competition, independently of that conduct's purpose.

Section 151AF defines the concept of a telecommunications market and says:

> Note: ***market*** has a meaning effected by section 4E.

As a result, it is reasonable to assume that issues of market definition and of market power are to be determined as they would under s46.

With respect to"taking advantage", the High Court has found that this requires that the conduct is only possible or profitable because of the absence of competitive constraints.[2] In other words, conduct that the firm would have entered into even if it lacked market power cannot be regarded as being conduct that "takes advantage" of market power. This interpretation is consistent with that adopted under s36 of the *Commerce Act*, 1986 in New Zealand[3], about which the Privy Council has said:

> "In their Lordships' view it cannot be said that a person in a dominant market position 'uses' that position for the purposes of section 36 unless he acts in a

---

2 *Queensland Wire Industries Pty Ltd v Broken Hill Proprietary Co Ltd* (1986) 167 CLR 177. In an Information Paper titled "Anti-competitive conduct in telecommunications markets" (May 1997), the ACCC appears to concur with the approach determined by the High Court, though it qualifies it without giving further explanation as to the qualification. The ACCC says: ".. Taking advantage of market power is, *in most cases*, indicated where the firm has engaged in conduct which would not be profit maximising in a competitive market" (at 24).

3 See *Commerce Commission v Port Nelson Ltd* (1995) 5 NZBLC 99-352, and (though now dated) Y. van Roy Guidebook to New Zealand Competition Laws (1991) at 150 and follows.

> way which a person not in a dominant position but otherwise in the same circumstances would have acted.[4]"

This interpretation is also consistent with that given by the United States Supreme Court in relation to §2 of the Sherman Act. Areeda and Hovenkamp summarise this interpretation in the following terms:

> Our concern about monopoly and the opportunities of rivals must not be allowed to obscure the objective of antitrust law, which seeks to protect the process of competition on the merits and the economic results associated with workable competition. Accordingly, aggressive but not predatory pricing, higher output, improved product quality, energetic market penetration, successful research and development, cost reducing innovations and the like .. are therefore not to be considered 'exclusionary' for §2 purposes even though they tend to exclude rivals and may even create a monopoly.[5,6]

Finally, it is clear that the concept of harming competition cannot be equated with that of harming competitors.[7] In *Queensland Wire Industries*, Mason CJ and Wilson J said that:

> The object of s46 is to protect consumers, the operation of the section being predicated on the assumption that competition is a means to that end. Competition by its very nature is deliberate and ruthless. Competitors jockey for sales, the more effective competitors injuring the less effective by taking sales away. Competitors almost always try to 'injure' each other in this way. This competition has never been a tort .. and these injuries are the inevitable consequence of the competition s46 is designed to foster.[8]

Equally, in *Eastern Express Pty Ltd v General Newspapers Pty Ltd* (1991) 30 FCR 385 at 406 Wilcox J says that:

> The more competitive the market, the more the principles underlying Pt IV are applied, the greater the damage likely to be sustained by less efficient participants. (..) Something more than competition .. is required before s46 is offended (..) the relevant conduct must be conduct which undermines competition.

---

4 Telecom Corp of NZ Ltd v Clear Communications Ltd (1995) 1 NZLR 385; (1994) 6 TCLR 138; 5 NZBLC 103,552.

5 P. Areeda and H. Hovenkamp <u>Antitrust Law</u> (1996) at §651b, citations omitted.

6 A similar concept can be found in the competition law of the European Union, though less clearly formulated and less consistently applied. See V. Korah <u>An Introductory Guide to EC Competition Law and Practice</u> (5th ed. 1994) at 84 and follows.

7 See generally P. Clarke and S. Corones <u>Competition Law and Policy: Cases and Materials</u> (1999) at 347 and follows.

8 (1989) 167 CLR 177 at 191.

In short, for the Competition Rule to have been breached by the conduct referred to above, proof is required that:

(1) Telstra had substantial power in relevant markets for telecommunications markets;

(2) The conduct at issue would not have been possible or profitable in the absence of that power. This requires construction of a 'counter-factual' in which Telstra lacks market power but which otherwise has the same features as the situation at issue; and

(3) The conduct was for a purpose proscribed under s46 or which would have had the effect, or likely effect, of substantially harming the competitive process.

## C What is the Internet

The Internet is the term "commonly used as a reference for the loosely administered collection of interconnected networks around the globe that share a common addressing structure for the interchange of traffic".[9] Currently, the Internet is composed of more than 60,000 constituent networks[10] connecting some 30,000,000 hosts.[11] The public links connecting these hosts are estimated to have a capacity equivalent to 75 Gbps[12], which is somewhere between 5 and 10 per cent of the world-wide capacity of the Public Switched Telephone Network.[13]

Technically, the Internet is distinguished by its reliance on the Internet Protocol ("IP"), a packet switched protocol[14] operated in connectionless mode. The IP is implemented through *routers*, which form the "nervous system" of the Internet. These routers receive incoming packets and perform both *switching* functions (that is, deciding how the packet should travel towards its destination) and *forwarding*

---

9 G. Huston ISP Survival Guide (1999) at 632.

10 Ibid at 91.

11 K.G. Coffman and A. M. Odlyzko "The size and growth rate of the Internet", mimeo (October 1998).

12 Ibid, at 11. Note that this estimate excludes leased lines connecting end-users to ISPs.

13 However, it is likely that Internet links are somewhat more heavily utilised than PSTN links.

14 *Packet switching* is a data communications technology in which data are broken down into blocks of bytes ("packets") which are then transmitted separately. It is to be distinguished from the *circuit switching* used in the PSTN. In circuit switched networks, the network establishes a connection ("circuit") between the called and calling party for the duration of a session, and the information exchanged during that session is transmitted as a continuous stream over that circuit. Packet switching techniques can be further classified into *connection-oriented* and *connectionless*. In connection-oriented techniques, a path ("state") is established between the sender and recipient at the outset of a session, and then closed ("torn down") at the end of the session. The data being transmitted then use that path for the duration of the session. In connectionless techniques, no such state is established, and each packet is routed independently. In general, packet switching is more efficient than circuit switching when the information being transmitted is both bursty (as against being continuous) and asymmetric (that is, more information is being sent in one direction than in the other). See generally L. L. Davidson and B. S. Davie Computer Networks: A Systems Approach (1996) at 151 and follows.

functions (that is, the actual process of sending the packet from the input port onto its next hop).

The means by which routers carry out these functions are complicated, and a description would require considerable technical detail. Nonetheless, the essential features of Internet routing can be conveyed through a simple analogy.

Imagine a world in which everybody might need to go to any specific house, anywhere in the world, and where streets could (and frequently did) change what other streets they connected to and could (and frequently did) appear and disappear. Printed maps would be useless – too many maps would be needed, and they would become out of date too quickly.

Instead this world has a police officer (the router) on each street corner to give directions to travellers (the data packets). These police officers would be able to see the next corner in each direction – and so could direct travellers there. They would periodically shout to the police officers at the next block down each of the streets, thus keeping each other updated as to who knew a way to get to any given place. If they received no reply from an officer who had previously been there, they would assume that that street had disappeared. By the same token, they would register replies from new officers as marking the addition of a street to the overall world map. Each time a traveller came by to ask the way, the police officer would direct that traveller on to the police officer at the next block who could then help them on the next step of the journey. So to get to a specific house in another town a traveller would first be directed (block by block) to get to that town, then to the street, and finally to the house on the street.

There is a marked contrast between this approach to traffic management and that adopted in the PSTN. In the PSTN, the signalling network[15] provides for centralised control, by the network, of the transport of traffic between the PSTN's end-points. End-to-end service quality is assured by the network's policing of the number of demands for connection it will accept, with each successful connection being assured access to a specified volume of network resources[16]. The Internet lacks any such centralised control by the network. Rather, transport management occurs on a decentralised, hop-by-hop basis, while end-to-end service quality is only assured by the end-points' monitoring of, and response to, congestion or other causes of packet loss.[17]

---

[15] The signalling network is that component of the PSTN which is responsible for the transmission of address, supervision and routing information between network end-points and switching systems and between switching systems, including any information needed for billing. The signalling network controls the establishment of circuits on demand between end-points in the PSTN.

[16] For example, once a telephony link is set up between two users, those users are assured access to a capacity of some 4KHz each for the duration of their session.

[17] It is for this reason that the Internet is generally referred to as a "best efforts" network. When a router accepts traffic onto the Internet, it does not know whether the network has sufficient resources to deliver that traffic to

This approach can provide significant efficiencies for two reasons. First, hop-by-hop routing allows for fuller utilisation of bandwidth[18], as it permits virtually continuous adaptation of the paths used to the load on the network at each of its nodes. Secondly, devolution of quality assurance functions to the end-points minimises the processing that needs to be carried out within the network, and exploits economies of scope in end-user systems between these functions and others.[19]

Importantly for the issues being considered here, these efficiencies are obtained at a cost in terms of the volume and type of information captured in the network.[20] In connection-oriented networks, the establishment, monitoring and termination of connections ("states") provides the basis for capturing detailed information about usage. No such basis exists in the Internet. Routers, particularly those used in large scale networks, are not designed as platforms for traffic accounting, as the processing overhead involved in capturing the types of information routinely gathered in the PSTN would very substantially degrade routing performance.

## D Access to and interconnection in the Internet

The many thousands of networks that comprise the Internet define a tiered or hierarchical structure. Internet Service Providers ("ISPs") connect end-users to the Internet, either through dial-up connections or through permanent links. For information to flow among these end-users, some form of network interconnection is needed.

Although the underlying technology of the Internet – the IP protocols – were devised to permit the linking of any number of diverse networks, no specification was ever established for the commercial terms on which the interconnection of ISP networks would occur. Consistent with the IP's public sector origins, the original assumption appears to have been that each ISP would bear its own costs, with the costs of the links between these being covered by government subsidies. As the commercial element in the Internet became more dominant, a pattern evolved centred on two forms of interconnection.

---

its destination. Nor does the network monitor whether that traffic ultimately is delivered. Rather, those functions are carried out by the source, notably through the use of the Transmission Control Protocol (TCP). TCP monitors the performance of the network and adjusts transmission and re-transmission accordingly.

18 Bandwidth is a measure of the volume of traffic that can be carried over a link or connection, and is usually given in terms of bits per second.

19 Thus, the PC on which this is being written is being used both as a Word Processor and (in running an electronic mail program) to operate TCP – that is, provide quality assurance for Internet transmission.

20 Additionally, the view has been expressed that the efficiency of information transmission over the Internet could be increased by making some use of connection-oriented switching. This is especially said to be the case with respect to applications (such as video-conferencing) that are intolerant to delay. This view underpins emerging technologies such as the Resource Reservation Protocol (Peterson and Davie, op. cit. at 456 and follows) and Multi-Protocol Over ATM (MPOA) (see G. M. Whalley et. al. "Advanced Internetwork Design" BT Tech. J. 16 (1998) 25 and P. Giacomazzi and L. Musumeci "Transport of IP Controlled-Load Service Over ATM" IEEE Network 13 (1999) 36).

The first is generally referred to as *peering*. In these arrangements, one ISP grants another ISP access to its network in exchange for the first ISP also gaining access to the second ISPs network. 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 – with each ISP bearing the cost of the other ISP's use of its network in exchange for the benefit of the use of the other's network.[21]

The second is generally referred to as *transit*. In this arrangement, a local ISP pays another ISP for use of the second ISP's network to provide the local ISP with connectivity to both the second ISP's clients and to the wider Internet. Transit functions are generally provided by service providers who operate backbone networks – that is, the links connecting networks located in different places. These firms are commonly referred to as *Internet Access Providers* ("IAP's").

## E Whether Telstra had market power in the supply of Internet access

The Competition Notice defined the relevant market as a market for "Access Provider Services":

the market for services providing access to and transmission of IP based data by means of carriage services by and between Internet access providers, and by and between Internet access providers and Internet service providers.

In *Queensland Wire Industries*, Deane J referred to the "variety of arguable markets" (at 195) and much the same could be said here. It seems simpler, and more faithful to the dynamic nature of the industry,[22] to define two markets, with the relevant allegations focussing on the second of these:

(a) a market for the provision of retail and wholesale Internet services ("the service provider market"); and

(b) a market for the provision of interconnection between Internet Service Providers ("the access market").

As noted above, each of these markets is hierarchically ordered. It is conventional in analyses of the Internet to refer to 'tiers' of service and access providers. Tier 3 providers typically operate on a relatively small scale in narrow geographical

---

[21] In theory, peering can occur on an "settlement for differences" basis. However, this is extremely uncommon, partly because it requires the measurement of bi-directional traffic flows (and hence does not provide for the saving of measurement costs which is one of the major reasons for peering). As a result, the term "peering" is used in this paper to mean "peering on a settlement free basis".

[22] In rapidly evolving industries, it is important not to base market analysis on a snap-shop view, but rather to see the market as a "moving picture of continuing commercial activity": *Tru Tone v Festival Records Marketing Ltd* (1988) 2 NZLR 352

markets. As a result, their networks have limited internal transmission capacity. Tier 2 carriers have a relatively wider scope of operations and provide and manage some capacity between places. However, they purchase access to the global Internet on a wholesale basis. Tier 1 providers generally have national and international networks, and provide and manage links capable of reaching all destinations in the global Internet.[23]

In late 1997, Telstra was the largest Tier 1 provider in Australia, with the other Tier 1 carriers being Connect.com (owned by AAPT), AccessOne (owned by OzEmail) and Optus. The Commission estimated that Telstra had 44 per cent and 53 per cent of domestic and international Internet bandwidth respectively. Other sources estimate that some 45 per cent of Australian ISPs directly obtained bandwidth from Telstra, while an additional 16 per cent obtained Telstra services by resale. However, of those ISPs directly or indirectly obtaining bandwidth from Telstra, nearly half also obtained services from another Tier 1 carrier, a practice generally referred to as "multi-homing".[24]

There are some economies of scale to the provision of Internet access, and the pattern of market shares may in part reflect those. These economies arise from two sources.

First, there is some evidence that the cost function of core routing is sub-additive.[25] Holding traffic and number of subscribers constant, the total industry-wide expenditure on core routers increases when the number of access providers increases by one, as all routing tables need to be extended. Additionally, the expertise needed to maintain and operate core routers is in the nature of a fixed cost, and hence may give rise to declining unit costs. Hence a standard authority concludes that "core systems work best for Internets that have a single, centrally managed backbone".[26]

Second, there are economies of scale to transmission. "Larger" circuits can be derived without additional outlays on trenching, housing and installation. As a result, unit prices typically fall as the capacity of a circuit rises.[27] Additionally, with random traffic loads, larger circuits can be utilised more heavily without degrading service quality.[28]

---

23 In practice, this involves providing capacity to and from major Network Access Points in the United States.

24 This information was obtained from a data base developed and maintained by the Sydney-based Internet market research firm, www.consult.

25 Sub-additive cost functions imply that it is cheaper for one firm to produce any given level of output than it is for two or more firms to produce that level of output. Core routers are those that hold entries in their router tables for all other routes in the Internet.

26 D. Comer Internetworking with TCP/IP (1995) volume 1 at 240.

27 In the highly competitive US market, for example, charges for circuits are typically proportional to the bandwidth of the circuit raised to a power in the range of 0.5 to 0.7. The scale elasticity is, in other words, between 1.4 and 2. P. C. Fishburn and A. M. Odlyzko "Dynamic behaviour of differential pricing and Quality of Service options for the Internet" (mimeo, 1998) at 9.

28 While the extent of this effect will depend on the nature of the stochastic process generating the observed traffic load, the effect itself will be observed across a wide range of stochastic processes. M. Tanner Practical Queuing Analysis (1995) at 326 and follows.

While these factors may give rise to scale economies, it is questionable whether they generate material barriers to entry, at least in the Australian market, or in any other way might have entrenched Telstra's position. This is for three reasons.

First, the sunk costs involved in entry are low relative to the scale of the market. Circuits, both domestic and international, can be obtained under a variety of arrangements and are readily traded[29]. While there have been occasional shortages of capacity, supply has adapted remarkably quickly to demand, and indeed has tended to outstrip it.[30] Routers are not high cost items of equipment, especially when set against the outlays routinely made by companies such as Optus and AAPT, and can also be resold.

Second, the market is growing very rapidly, with the result that new entry, even on a substantial scale, need not depress output prices.[31]

Third, purchasers of access are relatively footloose. The widespread practice of multi-homing (discussed above) allows purchasers of access to keep their options open while shifting their load across competing carriers.

These factors have been reflected in the strong expansion in the Internet capacity of Telstra's competitors. While Telstra accounted for virtually all the bandwidth used for Internet access in late 1995, by late 1997 its market share had been halved. From November 1996 on, the increment to capacity provided by Telstra's competitors exceeded that provided by Telstra itself.[32] Important new sources of competition had expanded very rapidly: Optus, for example, increased the capacity of its Spinnaker Internet Access Service from zero to 120 Mbps (equivalent to some 18 per cent of market demand) in just twelve months.[33] Further competition was coming from satellite, notably with the entry of PanAmSat into the Australian market (which brought an additional 150Mbit/s of capacity into the pool of available supply). Given these developments, it seems difficult to see how supply-side considerations could support the inference that Telstra had substantial market power.

In its Competition Notice, the Commission seemed to rely on demand-side scale economies as a source of market power. In particular, the Commission said (at §15 of the Notice) that as a result of "the amount of content, number of end users, and ISPs

---

29 The access provisions set out in Part XIC of the Act also conferred on the ACCC substantial control over the terms on which the inputs needed for entry could be obtained, notably with respect to local termination and domestic transmission.

30 Measured in 64kbit/s equivalents, trans-Pacific transmission capacity increased from 190,890 in 1997, to 311,850 in 1998 and 1,763,370 in 1999.

31 Over the period from December 1996 to July 1998, the number of ISPs in Australia increased from 300 to 680. (See www.consult 7th IAP Report: Internet Access in Australia (November 1998)). Over the period from October 1996 to January 1998, the bandwidth purchased by Australian ISPs increased from 122 Mbit/s to 696 Mbit/s. (See www.consult International Internet IP Bandwidth (January 1998) at 9).

32 www.consult 7th IAP Report: Internet Access in Australia (November 1998) at 16.

33 http://spinnaker.optus.net.au/ (26 April 1999).

located on Telstra's network", ISPs faced little or no choice but to use Telstra's services. This argument seems weak for three reasons.

First, it rests on the assumption that others cannot contest and attract away from Telstra "the amount of content, number of end users, and ISPs located on Telstra's network". In other words, its effectiveness relies on some (unspecified) obstacles to competition, presumably on the supply-side.[34]

Second, it seems to refer not to the Access Market (which hosts no content and has no end users) but rather to the Service Provider market. In this market, which was relatively fragmented, Telstra's market share was low, and it was not in any position to refuse access to the content it hosted or the end-users linked to it[35].

Third, it sits uneasily with the fact that over a third of all ISPs (and a significantly higher proportion of those using high capacity links) obtained no bandwidth from Telstra.[36] Moreover, these ISPs seemed to be expanding more rapidly than the market as a whole, suggesting that there were no material penalties associated with using non-Telstra sources of supply.

In short, the available evidence suggests that:

(1) Though there were some economies of scale to the provision of access, the barriers to entry into the Access Market were low.

(2) Reflecting this, while Telstra had and continued to have a substantially larger network than its competitors, its share of the Access Market had been diminishing rapidly, as competitors expanded.

(3) Considerable doubt must therefore be cast on the Commission's inference (at §16 of the Notice) that Telstra had substantial power in the relevant market.

## F Whether the behaviour involved taking advantage of market power

Even assuming that Telstra had market power, a contravention of the Competition Rule requires proof that the conduct "takes advantage" of market power. The central test here is whether that conduct would have been possible or profitable absent the market power at issue.

---

34 This is a variant of the error commonly made in diagnosing alleged network externalities of failing to take account of the scope for their internalisation. See generally S. J. Liebowitz and S. E. Margolis "Are Network Externalities a New Source of Market Failure?" 17 Res. In Law and Econ. (1995) 1.

35 While the Herfindahl index for the Access Market was in the order of 0.31, for the Service Provider market it was 0.15.

36 www.consult 7th IAP Report: Internet Access in Australia (November 1998) at 56.

In the Competition Notice, the Commission said:

> §21 By engaging in the Conduct (..) Telstra has at all material times taken and is taking advantage of its substantial degree of market power in the Access Provider Market, in that in the absence of such market power Telstra could not engage in the said Conduct.

This conclusion seems to rest on an analysis set out in an undated paper by the Commission's staff, titled "Allegations Received by the Commission and Staff Analysis for the Purposes of s.151AJ(2)". In that paper, Commission staff say:

> §31 The conduct is Telstra's refraining from negotiating or, in the alternative, refraining from agreeing with other IAPs for a mechanism whereby they will be compensated for use of their networks. In a competitive market IAPs would not supply traffic to an IAP that refused to negotiate or agree to reciprocal financial arrangements.
>
> §32 Information provided by competing IAPs supports this view. These similar-sized IAPs all lack market power and have signed reciprocal compensation agreements with each other.

These statements cover two distinct issues. The first, referred to in §32 of the Staff paper, involves peering – that is, agreements in which providers allow use of each other's network on what is essentially a barter basis. The second, referred to in §31 of the Staff paper, involves the terms on which transit is supplied – and in particular, whether the transit carrier owes some compensation to the service provider to whom a transit service is being supplied.

## (A) Peering

In practice, peering is an arrangement in which parties provide each other with unmetered access to one another's resources. From an economic point of view, it transforms each of the networks being thus connected into a common property resource for the others.[37]

As a general matter, common property is an inefficient form of organisation. Because the marginal private costs of access to each access-seeker do not reflect the social costs access imposes, the resource will be over-used in the short-term.[38] Additionally, long-term investment will be too low, as investors cannot anticipate capturing the

---

[37] A resource is a "common property" resource if use of that resource is rivalrous (*ie* use by one party displaces use by another) but not subject to exclusion. The classic example of a common property resource are the ocean fisheries. See H. S Gordon "The economic theory of a common property resource" 62 J. of Pol. Ec. (1954) 124, A. A. Scott "The Fishery" 63 J. of Pol. Ec. (1955) 116 and generally, G. D. Lidecap Contracting for Property Rights (1989).

[38] This phenomenon has been famously termed "the tragedy of the commons" by G. Hardin in "The tragedy of the commons" 162 Science 1243.

resulting returns. Further, because rights in common property are poorly defined, there is no guarantee that the resources will be used by those who value them most highly. Finally, any form of common property is likely to divert some resources from productive uses to socially wasteful investments in defensive and/or predatory activities. As a result, placing a resource into common property may (and usually will) lead to substantial rent dissipation.[39]

These concerns are of clear relevance to Internet interconnection. In a peering arrangement between parties with over-lapping networks, each party has an incentive to shift traffic from its own network to that of its peer. As this happens, the party which shifts the least traffic suffers congestion, as its links are now carrying its peer's traffic as well as its own.[40] If this process continues, each of the networks will progressively shrink its output and investment, with there being no assurance that it is the most efficient network that will survive.

A formal model, set out in an Appendix to this paper, illustrates these propositions, and shows that peering arrangements do not define a sustainable equilibrium for interconnection agreements between profit-maximising ISPs.[41]

Despite this, peering arrangements can occur and be sustained where there are strong constraints on free-riding. Two circumstances are especially significant.

First, where the networks at issue do not overlap, the scope for one network to shift the costs of transport onto the other is limited.

A second circumstance arises when the networks are of approximately equal size. In this case, each of the parties may feel that the resources to which it is obtaining access are of about equal worth to those to which it is granting access. Moreover, any significant free-riding by the peer network is likely to be detected rapidly, which need not be the case when there is a significant asymmetry in network size.[42] Free-riding can therefore only be of relatively short duration.

---

39 Rent dissipation can be defined as the use of resources in the attempt to secure economic rents. The general theory of, and evidence for, rent dissipation under common property is set out in T. Eggertson Economic Behaviour and Institutions (1990) at 83 and follows.

40 Although TCP provides a means of controlling congestion, IP networks are relatively vulnerable to quality degradation as a result of excess loads. This is partly because only 20 per cent of the traffic on the Internet is directly rate-adaptive, the rest using non-adaptive protocols such as UDP. It is also because TCP itself can be unstable when loss occurs simultaneously across a broad range of uses. This creates a condition known as *global synchronisation* in which vast numbers of parties attempt simultaneous retransmission. As a result, unplanned traffic shifts can impose major costs.

41 In a recent paper Pio Baake and Thorsten Wichmann ("On the Economics of Internet Peering" 1 Netnomics (1999) 89) set out a model in which decisions about peering have competitive effects. However, the model is deficient in numerous respects. To begin with, its formulation effectively assumes away free-riding. Even more importantly, the effects it obtains depend on the assumption that peering provides higher *quality* than transit. There is no obvious reason for this assumption.

42 This can be seen by considering two networks, A and B, of even size. A significant shift in traffic from A to B will substantially increase the load on B and hence is not likely to be mistaken for a random departure from normal traffic loads. However, if A is substantially smaller than B, it could shift much of its traffic to B

In these circumstances, the saving that peering effects in metering and billing costs may outweigh the risk associated with common property resources. However, too much weight should not be put on these cases. The US experience in particular shows that peering arrangements, even between relatively large carriers, are fraught with disputes and can be extremely unstable.[43]

Given this, the Commission appears to have erred in two respects in its analysis of the outcomes that would prevail in the absence of market power.

First and most important, it mis-specified the relevant counter-factual. As the Privy Council noted, the question is whether the firm "acts in a way which a person not in a dominant position *but otherwise in the same circumstances* would have acted."[44] By basing its analysis on the behaviour of a group of small, equally sized, ISPs the Commission failed to maintain the other circumstances of the case unchanged.

In particular, the relevant fact – as the Commission itself emphasized in its analysis of market power – was that Telstra's network was substantially larger than that of the other participants in the Australian market. As a result, the relevant question was whether a firm without market power, but substantially larger than another firm, would choose to enter into a peering arrangement with its smaller counterpart.

Here the evidence was quite at odds with that the Commission's staff cited. Even prior to the Notice being issued, Optus made it clear that it had no intention of entering into peering arrangements with smaller carriers[45] – a view since reaffirmed on numerous occasions[46]. Connect.com.au showed a similar reluctance, and both connect.com.au's and Optus' behaviour in this respect had been the subject of complaints to the Commission by smaller IAP's. Short of assuming that Optus and connect.com.au themselves had substantial market power, the inference the Commission drew in respect of Telstra (that its behaviour reflected market power) was poorly based.

Second, even had the counter-factual assumed by the Commission been correct, it could not support the conclusion the Commission drew. Thus, a careful analysis would have shown that peering, even among firms of approximately equal size, can have substantial costs. A refusal by a firm to bear costs on behalf of competitors does not seem like a reasonable basis for inferring a taking advantage of market power.

---

without the load on B going greatly outside the confidence intervals associated with the normal randomness in traffic load.

43 See, for example, D. Bushaus "Peering into Future Payments – Growing Internet Traffic Forces ISPs to Look at Settlement Systems" 309 Tele.com August 1, 1998.

44 Telecom Corp of NZ Ltd v Clear Communications Ltd (1995) 1 NZLR 385; (1994) 6 TCLR 138; 5 NZBLC 103,552, emphasis added.

45 See "Telstra to Offer Internet Plan Mark II", *Exchange*, March 6th, 1998, quoting Optus spokesperson Amanda Wallace.

46 See for example, B. Jew and R. Nicholls "Internet Connectivity: Open Competition in the face of Commercial Expansion", paper presented to the Pacific Telecommunications Conference, 17-20 January, 1999.

## Transit

The second leg of the Commission's claims involved an alleged asymmetry in payment between Telstra and other ISPs and IAPs. Such an asymmetry is independent of whether peering occurred, and hence needs to be addressed separately.

The Notice puts matters in these terms:

> §9. Save as to OzEmail Ltd, Telstra charges for Access Provider Services provided to other Internet Access Providers while at the same not time not paying for or otherwise compensating other Internet Access Providers for Access Provider Services it receives from other Internet Access Providers.(..)
>
> §18. Telstra, because of its market power, is able to charge other IAPs for the supply of Access Provider Services while not paying or otherwise compensating for Access Provider Services supplied by other IAPs (save as to OzEmail Ltd).
>
> §19. In a competitive market, Telstra would either pay for Access Provider Services supplied to it by other IAPs or enter into reciprocal financial arrangements with other IAPs (..).

In essence, the Commission is claiming that Telstra received services from, as well as provided services to, the IAPs at issue, but did not compensate them for the services it obtained. The Commission further asserts that in a competitive market, such compensation would be forthcoming.

This claim confuses an issue about the *level* of prices with one about price *structures*. When a customer buys a table from Ikea (which sells tables in disassembled condition, to be collected at the point of sale), as against buying a table from Designer Warehouse (which delivers an assembled table), Ikea is entering into an implicit contract with the customer for the customer's services in transporting and delivering the table. However, there is not, and does not need to be, an explicit payment for those services. Rather, the amount of any costs the customer allows Ikea to avoid is fully reflected in the price the customer pays Ikea for the table. In a competitive market, in other words, the *level* of prices will reflect the bilateral flow of services, but the *structure* of charging (in the sense of the nature and direction of settlement) will reflect other considerations.

These considerations relate to the costs of contracting, and notably of billing and metering. In an efficient industry configuration, measurement will be undertaken by the party that has the lowest cost access to information and requires the least resources to carry out the measurement task. Moreover, output (both in terms of volume and quality) will be measured at those points in the process of production, consumption and exchange where the measurement can be done with the least use of resources. Finally, the measurement signal will be structured in such a way as to allow effective monitoring of performance, thus providing incentives for performance commitments to be met.[47]

In this instance, the parties were purchasing transit services from Telstra – that is, Telstra was providing them with backbone transmission between places in Australia and/or between Australia and overseas. Parties were billed on the basis of the volume of data *received* from Telstra by the Australian client. This type of charging, also referred to as charging on a *return traffic* basis[48], reflects three factors.

First, charging on the basis of data received reflects underlying cost causation. Traffic flows for Australian ISPs are strongly asymmetric, with traffic received exceeding traffic sent by a factor of 2 or more. As a result, the network is dimensioned – notably at its points of interconnection in the US[49] – to carry traffic flowing to, rather than that flowing from, ISPs. Given that variable costs are negligible (as it is the in-bound flows that trigger congestion), it is the volume of traffic to ISPs that should be used as the price signal.

Second and related, charging on the basis of traffic flowing to ISPs has socially desirable incentive effects. Most obviously, it encourages ISPs to economise on international transmission, notably by using web caching.[50] As some 70 per cent of ISP traffic involves downloads from the World Wide Web, and as 50 per cent of material down-loaded is typically accessed frequently, storage at a site near the end-user can cut external transport needs by some 20 per cent.[51]

As well as inducing efficient caching choices by ISP's, charging on the basis of traffic flows to ISP's provides economically rational incentives to the provider of transit services. In particular, it is not easy for an ISP and a transit carrier to contract for a predefined grade of service.[52] However, the grade of service an ISPs customer receives is significantly influenced by the capacity and speed of that ISPs links to the global Internet. By only rewarding the transit carrier for traffic it actually delivers to the ISP, a charging mechanism based on return traffic flows automatically penalises

---

47 See Y. Barzel "Measurement costs and the organisation of markets" 25 J. of Law and Econ. (1982) 27.

48 Two directions are commonly distinguished in networks: the *forward* path, which runs from the customer (or more generally end node) to the network, and the *return* or reverse path, which runs from the network to the customer.

49 Since international traffic accounts for over 70 per cent of the total, with some 80 to 85 per cent of this coming from the United States, the costs associated with interconnection in the US represent a very large share of attributable costs.

50 "Web caching" refers to the storage, in a server on or close to the ISPs premises, of the results of recent downloads from the World Wide Web.

51 The actual reduction in transport needs effected by caching depends on the efficiency of the caching algorithm, notably in determining when to store and when to discard downloaded information. The assumption used here is that of 50 per cent efficiency.

52 As Huston notes: "The sheer size of today's Internet effectively precludes any QoS approach that attempts to reliably segment the network on a flow-by-flow basis." See also more generally Huston op. cit. at 315 and follows.

the transit carrier for bit loss, and can hence support efficient decisions about network dimensioning.[53]

Third, charging on the basis of traffic on the return path is less vulnerable to manipulation, and hence reduces the costs that must be incurred in resolving billing disputes.

All Internet traffic is vulnerable to routing manipulation. In contrast to the PSTN, dummy traffic can be generated readily, and will not be detected as such.[54] Moreover, *loose* and *strict* source routing[55] options allow the user of the addressing protocols to specify locations through which a packet must transit on its way between terminal points. It follows that any scheme of traffic related charging is liable to being abused. However, return traffic is significantly less vulnerable to abuse than traffic on the forward path. This is for three reasons.

(1) A transit carrier could not use loose source routing to distort the flow of traffic to an ISP lacking a transit network.

(2) Capacity is dimensioned on the basis of the return path, since traffic flowing from the network to the terminating ISP is almost always much greater than that flowing from the terminating ISP to the network. As a result, a transit carrier seeking to generate dummy traffic would in all likelihood quickly run into capacity constraints. In contrast, an ISP seeking to send dummy traffic loads to the transit carrier would have ample scope for doing so.

(3) Attempts by a transit carrier to generate substantial volumes of dummy traffic on to the terminating ISP's network are likely to be identified far more quickly and surely than are dummy flows in the opposite direction. This is because the transit carrier, in this case Telstra, is substantially larger than its terminating ISPs. A slight increase in the traffic load from Telstra to the terminating ISPs will involve a large percentage increase in the load on each of the terminating ISP networks. As a result, it will be readily detected, and subjected to analysis. In contrast, even a substantial absolute increase in the inbound traffic from the terminating ISPs to Telstra will be small relative to Telstra's traffic load. As a result, it could readily fall within the normal range of statistical load fluctuation.[56]

---

[53] TCP is designed to tolerate some bit loss, so that the efficient level of congestion in an IP network with some TCP being run over it is not zero. For a given price for delivery of transit traffic, the transit carrier can expand capacity to the point where the marginal cost of further capacity additions equals the transit price.

[54] Indeed, dummy traffic flows are an integral part of network management in the IP, as they are used to monitor the status and performance of the network.

[55] *Loose* and *strict source routing* are options in the IP version 4 protocol. In loose source routing, specific intermediate destinations are specified. In strict source routing, every router that the packet must transit on the way to its destination is specified. See J. W. Stewart BGP4: Inter-Domain Routing in the Internet (1999) at 8, T. Russell Telecommunications Protocols (1997) at 164-165.

[56] This is on the plausible assumption that volatility increases with mean load, as it does on IP networks.

In short, there are compelling reasons for charging to be carried out on the basis of the flow of traffic on the return path – and this is the primary means of charging in transit arrangements worldwide.[57]

In contrast, the ACCC seemed to assume that some form of "traffic accounting" was possible, in which different types of traffic flows would have been monitored, and transit and terminating traffic distinguished and charged for separately. However, while this can be done on relatively small networks, the technology at issue scales poorly and hence imposes substantial additional costs when implemented in a major backbone network.[58] Here too, the Commission therefore fell into two errors: it compared Telstra's behaviour with that of a small network, thus using an irrelevant counter-factual[59]; and even then, it failed to properly consider the economic factors at work in the design of Internet charging arrangements.

## G Whether the behaviour substantially lessened competition

The analysis set out above concludes that Telstra's conduct had strong efficiency justifications. This in and of itself casts doubt on the claim that it could have the purpose or effect of substantially lessening competition.

With respect to purpose, the pursuit of efficiency, or the avoidance of inefficiency, are not considered to fall among the purposes proscribed by s46[60]. With respect to effect, it would seem contrary to the goal of the competition policy to penalise a firm, even one with great market power, on the grounds that it was seeking to secure or preserve a better use of resources.[61]

---

57 There are some instances in which charging is carrier out on a "traffic accounting" basis, that is, through the monitoring of individual traffic flows and their usage of network resources. However, the traffic metering required scales very poorly and hence is only really feasible in relatively small networks.

58 On a rough estimate, routing efficiency is reduced by 10 to 20 per cent by detailed traffic accounting. Additionally, costs need to be incurred to store and process the information.

59 The Commission, in other words, did not compare the efficient behaviour for a network of Telstra's size with and without market power. Rather, it compared the efficient behaviour for a network of Telstra's size with the set of feasible behaviour for a network substantially smaller than Telstra's. This comparison is obviously irrelevant to the issue at hand.

60 A firm and unambiguous view on this issue is expressed in *Clear Communications Ltd v Sky Television Network and Others* (1994) New Zealand High Court, see especially at 68, where Gallen J. and Dr. M Brunt consider and reject the argument that efficiency effects should only be canvassed in the context of authorisation. Rather, they conclude that these effects are central to the assessment of competitive effects and purpose. From a purely legal point of view, it is possible that the conduct involved many purposes, only one of which was the pursuit of efficiency, with one or more of the others being proscribed purposes under s36. However, such an argument would be speculative, and seems inconsistent with the strength of the efficiency defences. See relevantly, *General Newspapers Pty Ltd v Telstra Corporation* (1993) 45 FCR 164.

61 See for example P. Areeda and H. Hovenkamp Antitrust Law (1996) at §658 and classically, *Olympia Equipment Leasing v Western Union Tel Co* 797 F 2d 370 (7th Cir 1986) cert. denied, 480 U.S. 934 (1987) at 379.

The Commission's analysis, however, disregards these efficiency effects. Rather, the Notice states:

> §22. The effects or likely effects of the Conduct on competition in the Access Provider Market are:
>
> (a) By engaging in the Conduct, Telstra substantially raises the costs of rival IAPs;
>
> (b) The higher costs of rival IAPs substantially hinders their ability to attract ISPs, end users, and content providers to their networks, which further limits their ability to generate additional revenue and to compete with Telstra;
>
> (c) The Conduct threatens the viability of at least some IAPs;
>
> (d) The higher costs of rivals IAPs are reflected in higher prices to ISPs and these higher prices are reflected in higher prices to end users and content providers than would exist in a competitive market; and
>
> (e) the likelihood of experiencing higher costs for the provision of Access Provider Services in Australia acts as a substantial disincentive for entry by potential entrants into the Access Provider Market.

These claims appear somewhat mis-stated[62] but even more importantly, fail to distinguish impacts on competitors from impacts on the competitive process. The confusion between these is even more striking in the Staff paper, which says:

> §37. A staff analysis [found that] .. If Telstra's charges were reduced to 5 cents, the result would be a dramatic decrease in cost for competing IAPs, promoting both competition and investment.(..)
>
> §40. [An] IAP has argued that the effect of Telstra's refusal to enter into reciprocal financial arrangements is to squeeze margins to the point where IAPs are forced from the market. This IAP has quantified its loss stemming from the lack of reciprocity at over $700,000 (..)
>
> §41. If Telstra agreed to a reciprocal financial arrangement, the cost to rival IAPs would decrease significantly. Two IAPs have indicated that with a reciprocal financial arrangement for domestic traffic, each would only pay 10 per cent of its present fees to BPD [Big Pond Direct].

62 Claim (d), for example, which states that prices are higher than they would be in a competitive market, seems both irrelevant (as the issue is not whether prices are higher than they would be in a competitive market, but rather whether than they are higher that they would be in the absence of the alleged conduct) and when appropriately recast, inaccurate (given that the conduct at issue increased efficiency).

It does seem plausible that if Telstra reduced its transit charges – in the extreme, say, to zero – and yet continued to provide transit service, purchasers could be better off.[63] However, whether competitors would be advantaged by a reduction in input prices is not the relevant test. Rather, what is of relevance is whether charges would be materially lower, and competition materially more vigorous, in the absence of conduct that relied on market power. It is this that the Commission failed to consider or demonstrate.

## H The remedy and its effects

As a result of its analysis, the Commission effectively required Telstra to enter into peering arrangements which it might otherwise not have accepted. This has had at least three effects in terms of resource allocation.

First, it has imposed additional resource costs on Telstra.[64] This is because the Commission has effectively required Telstra to accept arrangements in which it provides both peering and transit to an IAP. As a result, Telstra has had to deploy equipment for traffic accounting which is both costly in itself and has an impact on network performance.

Second, it has conferred substantial rents on the immediate beneficiaries of the requirement to peer.

Third, it has altered market behaviour in important respects. There are, in particular, strong signs that competing IAPs have restricted their investment in transport outside of the main Eastern metropolitan areas, relying instead on the Telstra backbone.[65] As a result, the Commission's actions have tended to lessen, rather than enhance, diversity and competition in the provision of Internet backbone services.

While these consequences are undoubtedly material, they are perhaps less important than the cloud of uncertainty which the Commission's actions cast over the market-place. In the process of negotiation and then in the Notice itself, the Commission did not articulate any clear rule (much less a bright-line test) by which Telstra itself and its competitors could distinguish permissible from impermissible conduct. Rather, it made statements which were at best inconsistent[66] and which, taken as a whole, left

---

[63] Note, however, that even in this implausible scenario, social surplus would fall relative to a situation where charges were based on costs.

[64] That is costs above and beyond financial transfers.

[65] See for example www.consult 7th IAP Report: Internet Access in Australia (November 1998), especially at pages 18 to 21.

[66] Thus, the Notice itself limited Telstra's alleged impropriety to conduct towards a small number of designated competitors; but the Commission's claim was explained to *The Australian* newspaper by ACCC Commissioner Shogren in these terms: "Our concern has been that reciprocal pricing had to apply to all major players in the marketplace, not just a couple of sweetheart deals. *We want these sort of agreements to spread across the industry*". (Nicole Manketlow "First Blood" in The Australian, June 2nd 1998; emphasis added.) Needless to say, the Commission never set out, in plain terms, what it meant by "major players" or quite what "reciprocal pricing"

extremely vague the rights and obligations flowing from its decision. This has made it difficult or even impossible for subsequent commercial negotiations to rectify the damage the Commission's decision caused.[67]

## I Conclusions

Accuracy in law enforcement and adjudication is valued both because of its importance to fairness and because of its influence on the efficiency of the legal system.[68] More accurate enforcement increases compliance for any given level of penalty. It reduces the need to impose socially costly penalties.[69] And it avoids over-investment in precautions against prosecution, and hence limits the risk that socially desirable conduct will be deterred.

In the context of the competition laws, and notably of provisions against unilateral conduct, accurate enforcement is made all the more important by the danger of deterring vigorous competition on the merits. The fact that the Competition Notice effects a shift in the burden of proof, and hence increases the risk that innocent conduct will be punished, should underscore this concern.[70]

Through its intervention in the Internet peering issue, the Commission imposed an access regime which (1) applied solely to Telstra's facilities and (2) provided a small and select group of competitors with access at charges well below those faced by other firms. Both of these points deserve some comment.

First, the dangers involved in imposing third party access are well-recognized. Thus, as Justice Breyer of the US Supreme Court noted in a recent case:

> "Rules that force firms to share every resource or element of a business would create, not competition, but pervasive regulation, for the regulators, not the marketplace, would set the relevant terms".[71]

---

actually meant. The errors and inconsistencies in the Commission's analysis of whether the conduct involved making use of market power, and if so why, only aggravated the resulting uncertainty.

67 As is well-known, legal decisions which create uncertainty about property rights are especially costly. This is because negotiations cannot then readily re-allocate the rights to the party which can put them to most effective use. Rather, uncompensated harms will persist, even when an alternative allocation of rights would increase efficiency. See Lidecap, op. cit. and J. M. Buchanan "Opportunity costs and legal institutions" in The New Palgrave Dictionary of Economics and the Law (1998) II, 715.

68 See generally L. Kaplow "The value of accuracy in adjudication: An economic analysis" 23 J. of Legal Studies (1994) 307.

69 This is because more accurate enforcement first, avoids the imposition of sanctions on the innocent. Second, for any given level of deterrence, total sanctions imposed on the guilty can fall, as the deterrent effect of sanctions depends on the *difference* in likely treatment between the innocent and the guilty.

70 For any given degree of adjudicative accuracy, shifting the burden of proof towards the defendant must raise the likelihood of punishing innocent conduct relative to the likelihood of excusing guilty conduct. See generally L. Kaplow and S. Shavell "Accuracy in the determination of liability" 37 J. of Law and Economics (1994) 1.

71 *AT&T Corp. et. al. v. Iowa Utils. Bd. et al.*, S. Ct. 721 (1999) at 754.

Even in the EU, where a version of the "essential facilities" doctrine has received relatively wide application, recent decisions have struck a significant note of caution. Advocate General Jacobs, for example, has recently said:

> "First, it is apparent that the right to choose one's trading partners and freely to dispose of one's property are generally recognised principles .. Incursions of these rights require careful justification. Secondly the justification in terms of competition policy for interfering with a dominant undertaking's freedom to contract often requires a careful balancing of conflicting considerations. In the long term it is generally pro-competitive and in the interests of consumers to allow a company to retain for its own use facilities which it has developed for the purpose of its business. For example, if access to a production, purchasing or distribution facility were allowed too easily there would be no incentive for a competitor to develop competing facilities. Thus while competition was increased in the short term it would be reduced in the long term. Moreover the incentive for a dominant undertaking to invest in efficient facilities would be reduced if its competitors were, upon request, able to share the benefits".[72]

Reflecting these concerns, the Court of First Instance now imposes a high hurdle before such access is required, saying that:

> "a product or service cannot be considered necessary or essential unless there is no real or potential substitute .. (that is) there are no viable alternatives".[73]

The contrast with the threshold set by the ACCC in the Internet peering case is apparent.[74]

Turning now to the terms and conditions on which access was provided, it is striking that these appear to have been set without any seeming consideration of underlying costs. It is well-established that even a monopolist should not be required to provide third parties with services on terms more advantageous than those on which it provides those services to itself.[75] In contrast, the ACCC, by ignoring the scope for free-riding, allowed the immediate beneficiaries of its decision to seriously under-compensate Telstra for the substantial costs they caused Telstra to incur. The fact that these beneficiaries were only a selected few out of the larger number of actual and potential access seekers makes this all the more questionable.

---

[72] Case C-7/97, Oscar Bronner GmbH & Co. KG v Mediaprint Zeitungs-under Zeitschriftenverlag GmbH & Co. KG and Others, delivered on 28th May 1998.

[73] Joined Cases T-374/94, T-375/94, T-384/94 and T-388/94, European Night Services and Others v. Commission, Judgement of September 15, 1998, not yet published, at para 285 and 207-9.

[74] Compare also to the recent decision in *Paddock Publications v Chicago Tribune Co et al* 103 F. 3rd 42 (7th Cir. 1996) rehearing den. 103 F.3rd 42, cert. den. 117 S. Ct. 2345 (1997), asserting that where there are three competing facilities, even though of unequal size, there can be no justification for an obligation to deal.

[75] Implicitly recognising the risk of free-riding, the formulation in *United States v Terminal R.R. Association of St. Louis* 224 U.S. 383 (1912) is "on a plane of equality in respect to benefits *and burdens*" (emphasis added).

The use of the Competition Notice system in this way stimulates two concluding remarks.

First, it is by no means apparent why the ACCC considered these issues under its Part XIB powers rather than relying on the access provisions of Part XIC. Those provisions would have imposed greater discipline on the process, in terms of neutrality of treatment and of rigour in the setting of access charges. Moreover, had access regulation been warranted (and the factors set out above suggest it was not), a formal access regime would have provided for greater effectiveness in determining, monitoring and revising the conditions of access.[76]

Second, the Competition Notice system vests very substantial powers in the Commission. The dangers this creates are all the greater because seeking recourse to the system is essentially a "one way bet" as far as Telstra's competitors are concerned. Their costs are socialised; the worst that can happen is that the Commission decides not to proceed with a complaint. As a result, the system exhibits none of the effects that the English rule of cost allocation has in deterring low-probability claims under s.46.[77] The ACCC's failure to articulate any clearly defined tests for determining the merits of complaints makes a strategy of complaint all the more attractive.[78] Given the rapid pace of change in the industry, and the resulting difficulty involved in assessing the merits of complaints, one would expect the Commission to respond to these incentives for the system to be manipulated by exercising great case in using its Competition Notice powers.

However, close scrutiny casts serious doubt on the manner in which the Commission has exercised these powers, at least in this instance:

(a) The underlying analysis appears to have been inadequate in important respects.

(b) The remedies effected seem contrary to the purposes of the competition policy. And

[76] Indeed, the supposed greater effectiveness of formal access regimes relative to provisions such as those of s46 is a central plank in the argument commonly put for Pt IIIA and Pt XIC of the TPA.

[77] Under the English Rule, costs are borne by the losing party. The deterring effects of such a Rule on un-meritorious claims are examined in S. Shavell "Suit, Settlement and Trial: A Theoretical Analysis Under Alternative Methods for the Allocations of Legal Costs" 11 J. of Legal Studies (1982) 55. While there is a lively debate as to the extent of this effect, the relevant literature leaves no doubt that a rule that removes any risk of loss from the complainant will reduce the quality of claims.

[78] Had the Commission set out bright line tests, those considering initiating a complaint might have been more cautious, for fear of acquiring a reputation with the Commission of being merely vexatious. In fact, the Commission's Telecommunications: Competition Notice Guideline (1997) are merely procedural, while its Information Paper Anti-competitive Conduct in Telecommunications Market is vague and seems designed not to limit the Commission's discretion.

(c) By not articulating any clear test as to which conduct was acceptable and which was not, the Commission created uncertainties which have weighed on subsequent commercial processes.

During the recent Senate Select Committee Hearings into Telstra's further privatisation, several parties argued for an expansion of the Commission's powers under Part XIB of the Act. The desirability of conferring any further powers on the Commission under Part XIB, which will allow the Commission to act more quickly and on an even weaker basis of fact and analysis, certainly does not emerge from an assessment of the case at hand.

# ATTACHMENT 1 ISP Peering

## 1. Introduction

The following is a simple stylised model that tries to capture some key aspects of the ISP peering problem. In the model, firms each provide some capacity, which under settlement-free peering is available to them both. Firms compete for customers through two-part tariffs. Their usage sensitive prices dictate the demand for capacity. We first characterise the welfare maximising choice of prices and capacity. In this solution, we show the social planner would choose usage prices above the marginal cost each firm faces. These prices take into account the extra cost 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[79].

In contrast, we show that the competitive equilibrium under settlement free peering will involve congestion: firms, through competition, price at a level in which demand for capacity exceeds the supply. An 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 characterised by each firm's usage prices set equal to marginal cost. In addition to under-pricing, we show the equilibrium also suffers from a free-rider problem, in which each firm would like to provide 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 explore a possible remedy. We show that if firms can charge for use of their facilities then 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 marginal cost of the supplied infrastructure. Under this rule, we show that usage prices and the level of infrastructure will be set at the same levels as the central planner would choose. 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. While ex-ante, under some specifications of infrastructure costs, firms have a mutual incentive to reach such a settlement agreement, ex-post each firm has an incentive to try to re-negotiate terms of settlement agreement in its favour. An alternative arrangement to solve the free-rider problem, if both firms face the same cost of providing infrastructure and equal demand for usage, is for each firm to only interconnect with the other if the other

---

[79] In a more complex model, such as one involving time-varying usage, some congestion would be optimal. In the case our results apply to the degree of congestion, rather than whether there is congestion or not.

firm provides capacity to cover its own usage. In this case, both firms provide equal facilities and have equal demand. Ex-post there is no need for settlement contributions in this symmetric case. Both firms take into account the capacity constraint even when pricing in competition.

## 2. Model set-up

We suppose there are two ISP's, denoted 1 and 2. Each ISP (or firm) provides customers access to a common infrastructure (the Internet) and charges customers for usage of this infrastructure.[80] 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 $k_1$ of the infrastructure and firm 2 provides $k_2$ of the 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 $f$.

Each customer of firm i uses $q_i$ units, where each unit can be thought of as a megabyte of data transfer.[81] This level of usage generates utility $u(q_i) + v_0 + \boldsymbol{q}_i$, 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, while $\boldsymbol{q}_i$ measures the additional costs and benefits from belonging to a particular network i (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 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

---

[80] 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.

[81] Unlike telecommunications, 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).

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

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 additional services provided by one of the ISP's or because of the greater reputation that one firm has developed.

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.[82] Each firm charges a per-unit price $p_i$ and a lump-sum fee (or rental) $r_i$. The share of customers that belong to firm 1 is then easily shown to be

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

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

$$sq(p_1) + (1-s)q(p_2) \le k_1 + k_2.$$

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

This says the weighted-average usage has to be within the capacity provided. Of course, customers can attempt usage above capacity levels. In practice, this causes some usage to be delayed, thus lowering customers' utility. For simplicity, we model any usage above capacity as generating no utility.[83] 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 maximise 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 $sq_1 + (1-s)q_2 \le k_1 + k_2$.[84] After some manipulation the following first order conditions can be derived (the derivations of this and all other results are contained in the appendix)

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

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 calls that are made. 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

$$s = \frac{1}{2} + \frac{b}{2} .$$

---

[83] 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.

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

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

The solution above determines the level of total capacity provided. 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, it would make sense to have one firm provide the entire capacity (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 decentralised solution under peering

Under a peering approach neither network owner pays for use of the other's network. In this case, firms first choose their investment in capacity with firm i producing capacity $k_i$, and given this they then choose their prices. Both decisions are 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. The Bertrand equilibrium in prices (where each firm sets its two-part tariff to maximise its profit, taking as given what the other firm is doing), implies

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

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 the firms' market shares are independent of $k_1$ and $k_2$. 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

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

Maximising each of these, taking $k_1$ and $k_2$ as given, we get the first order conditions

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

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 (we show this formally in the appendix). Firms in competition, under settlement free peering, under-price usage.

In fact, we can show 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 their no incentive to under-price. As just shown, this involves rationing.

### 4.3 Rationing equilibrium

In this subsection we characterise the equilibrium with rationing. We suppose that usage above capacity is rationed in a random fashion that neither firm can control (other than through pricing). 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 idea that with excess demand for capacity, there will be delay for all consumers and lower utility from this 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

$$q_1^r = gq_1$$
$$q_2^r = gq_2,$$

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[85], the share of customers that belong to firm 1 is

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

where

$$w_i = z(p_i)-r_i$$

[85] 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.

is the net surplus offered to firm i's consumers and

$$z(p_i, g) = \max_{q_i}\{u(gq_i) - gp_i q_i\}.$$

The firms' profit functions are now

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

The first order conditions are then

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

Substituting the first order conditions back into the profit functions implies

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

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 will 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 will be 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 decentralised and simultaneous way, it is unlikely they will 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. 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.

## 5. Settlement payments

In this section, we continue to assume firms first choose their investment in capacity, and given this they then choose their prices. Both decisions are made non-cooperatively. However, we also suppose there are settlement rules put in place before firms 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 this rule leads to the social planner's efficient outcome discussed in section 3.

Whether firms have a mutual incentive to implement these rules ex-ante, depends on the specification of the cost of building capacity. For some specifications, the adoption of the rule leads to higher profits for the firms. However, once the investment is in place, the net-users of the infrastructure have strong incentives to try to renegotiate the rule to obtain cheaper access. Under the 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.

### 5.1 Case 1: The asymmetric solution

Suppose, without loss of generality, that 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

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

while firm 2's profit is

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

Firm 1 chooses $p_1$ and $r_1$ to maximise $p_1$ 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 maximise $p_2$. The first order conditions are

$$\begin{aligned}
p_1 &= c + l \\
p_2 &= c + t \\
r_1 &= f + \frac{s}{s} + (t - l)q_2 \\
r_2 &= f + \frac{1-s}{s} \\
k_1 + k_2 &= sq_1 + (1-s)q_2.
\end{aligned}$$

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

$$\begin{aligned}
p_1 &= \frac{s^2}{s} + lk_1 - f(k_1) \\
p_2 &= \frac{(1-s)^2}{s} + lk_2 - f(k_2).
\end{aligned}$$

If each firm set $k_1$ and $k_2$ independently, it will choose

$$\begin{aligned}
l &= f'(k_1) \\
f'(k_1) &= f'(k_2)
\end{aligned}$$

Thus, ignoring the rentals, the equilibrium satisfies all the conditions of the central planner's solution

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

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 $(1-s)q(c + f'(k_1))f'(k_1)$. However, this ignores the possibility that firm 1 may then be unable to break-even. This will be the case when average costs exceed marginal costs by a sufficient degree. Algebraically, this is represented by the condition that

$$f'(k_1)k_1 - f(k_1) < -\frac{s^2}{s}.$$

As the degree of competition increases towards perfect competition ($s$ tends to infinity), this becomes increasingly likely. Then the appropriate settlement rate tends towards the average cost of capital provided, rather than the marginal cost used above. In particular, with perfect competition,

$$t = \frac{f(k_1)}{k_1} \text{ rather than } t = f'(k_1).$$

This is the lowest settlement rate that will ensure firm 1 can break-even, thus implying the least distortion to usage prices.

Unlike the solution studied in section 4.2, the solution here is indeed an equilibrium. Neither firm has an incentive to under-price, given what the other firm prices. 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. The firms acting independently set rentals equal to

$$r_1 = f + \frac{s}{s}$$
$$r_2 = f + \frac{1-s}{s}$$

While the level of these fixed charges was not determined by the central planner's solution, we did find they should be equal. Here, if there is asymmetry in the market ($b \neq 0$), firm 1 sets different rentals from firm 2. The inefficiency that results is general to competition with differentiated goods and not specific to a model of ISP peering. The idea is that, in equilibrium, a firm with a less popular product (but equal costs) will tend to price below the level of a firm with a more popular product. Some customers switch to the less desirable product, because of the lower prices. This is inefficient, since given the costs are the same, it is efficient for everyone to choose the product they prefer at equal prices. In the case here, usage prices are equal but rentals may differ. This draws some customers away from their otherwise preferred provider.

For example with $b > 0$ (so that firm 1 is the preferred provider at equal rentals), the share that firm 1 will receive in equilibrium is

$$s = \frac{1}{2} + \frac{b}{6} \ ,$$

rather than

$$s = \frac{1}{2} + \frac{b}{2} \ .$$

Firm 2, the firm with low market share, charges a lower rental than firm 1, causing some customers to switch from firm 1 to firm 2. This switching is inefficient.

An important question is would firms want to adopt such a settlement regime ex-ante? Comparing the profits above with those from section 4.3, the answer is yes if

$$f'(k_1^s)k_1^s - f(k_i^s) > -f(k_i^r) \ .$$

The superscript s denotes the level of capital that we found above (under settlement peering), while the superscript r denotes the level of capital under the settlement free rationing equilibrium. Since there is less capacity provided under rationing ($k_i^r < k_i^s$), it is possible that this condition will not hold. Assuming the condition does hold, both firms are better off adopting the rule. The real problem arises ex-post, where the access user has an incentive to try to re-negotiate the terms of settlement in its favour with the access provider. For instance if firm 2 is the access seeker, then we can derive that

$$\frac{\partial p_2}{\partial t} = -2(1-s)q_2 + k_2 < 0 \quad (\text{since } (1-s)q_2 > k_2),$$

so that firm 2 would like to re-negotiate towards a lower settlement rate. Similarly firm 1 would like to charge a higher settlement rate. To avoid the incentives for re-negotiation, and the associated implications for ex-ante investment, another solution is for firms not to be net users of each other's facilities. We consider this case next.

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

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

while firm 2 maximises

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

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

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

Maximising profit with respect to each firms choice of capital provision, implies

$$l_1 = f'(k_1)$$
$$l_2 = f'(k_2)$$

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. Appendix

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

## 6.1 The central planner's solution

The central planner chooses $k_1, k_2, p_1, p_2, r_1, r_2$ to maximise 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 of the firm. 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 will be 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 should be 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.

### 6.2 Is the capacity constraint binding?

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

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

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

So the first order conditions are

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

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.

### 6.3 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

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

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

$$\begin{aligned} dp_1 &= s(p_1 - dp_1 - c)(q_1 + dq_1) - s(p_1 - c)q_1 + sdr_1 \\ &= -scdq_1 - sp_2 dq_2 + s[u(q_1 + dq_1) - u(q_1) + u(q_2) - u(q_2 - dq_2)] \end{aligned}$$

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

$$\begin{aligned} dp_1 &= s(p_1 - c)dq_1 \\ &= s\, l_1 dq_1 \end{aligned}$$

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, will underprice usage.

### 6.4 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 transform the problem into the firms picking 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

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

$$T_1 = \frac{cs}{su'(q_1^r)} + cq_1^r + f$$

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

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

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

## 6.5 Settlement payments: case 1

The profit functions (re-written) are

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

and

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

Firm 1 chooses $p_1$ and $w_1$ to maximise $p_1$ 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 maximise $p_2$. Differentiating we get the following conditions

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

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

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

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

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 yield

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

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

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

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

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

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 will choose

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

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 will be

$$s = \frac{1}{2} + \frac{b}{2} + s\,[(v_1 - r_1) - (v_2 - r_2)] = \frac{1}{2} + \frac{b}{2} + (1 - 2s)$$

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 may not be the same across firms.

$$p_1 = c + f'(k_1)$$
$$p_2 = c + f'(k_2)$$
$$f'(k_1) = f'(k_2)$$
$$k_1 + k_2 = q(c + f'(k_1))$$
$$r_1 = f + \frac{s}{s}$$
$$r_2 = f + \frac{1-s}{s}$$

## 6.6 Settlement payments: case 2

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

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

while firm 2 maximises

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

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 maximise its profit we get the result in the main text.

## 7. References


Armstrong, Mark. 1998. "Network Interconnection" The Economic Journal 108: 545-564.

Carter, Michael, and Julian Wright. 1999. "Interconnection in Network Industries" Review of Industrial Organization, 14, 1-25.

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