

An excess of access: An examination of Part IIIA of the Australian Trade Practices Act

HENRY ERGAS¹

Abstract

Part IIIA of the Australian Trade Practices Act defines circumstances in which a facility owner may be required to provide a third party with use of its facility. This paper examines what Part IIIA might be doing from an economic perspective and criticises ‘monopoly leveraging’ arguments for third-party access. It argues that the transactions costs of access are potentially significant, and can exceed any efficiency gains third-party access permits. These contentions are corroborated by reference to the long-running dispute between the Fortescue Metals Group and BHP Billiton Iron Ore over access to rail track in the Pilbara region of Western Australia.

Introduction

In June 2004, the Fortescue Metals Group Ltd (‘FMG’) — then a purely speculative mining venture — applied under Part IIIA of the Trade Practices Act 1974 (‘the TPA’) for mandated access to a particular section of rail track in the Pilbara region of Western Australia.² This track, referred to as the Mt Newman line, was to be used by FMG to transport ore from a prospective ore body that was too small to justify a dedicated rail link of its own. The railway facilities in question are owned and operated by BHP Billiton Iron Ore (‘BHPBIO’), and are used to transport millions of tonnes of iron ore from inland mines to specially-

¹ ergas.henry@gmail.com. I owe a large debt of gratitude to Jonathan Pincus, Sabine Schnittger, Ed Wong, Jason Soon, Eric Ralph and Sam McSkimming. I am also very grateful to the Editor, William Coleman, for his comments and advice. I provided consultancy advice to BHPBIO in its filings to the National Competition Council in the period leading up to the declaration of the Mt Newman Line but am not involved in the current proceedings. That said, the views expressed here are entirely my own, as is the responsibility for any errors.

² There are a number of legal regimes in Australia that provide for third-party access. The focus here is on Part IIIA of the Trade Practices Act, which is generally referred to as the national access regime (reflecting the fact that it is not specific to a particular industry or state or territory but, rather, applies, potentially, to all entities subject to the Trade Practices Act).

built export ports on the Western Australian coast. BHPBIO's total rail system in the Pilbara consists of over 1000km of track, and is capable of transporting in excess of 350,000 tonnes of iron ore every single day of the year.³

Almost exactly five years later, FMG's application is still undetermined, with the matter scheduled to be heard by the Australian Competition Tribunal towards the end of 2009. The time the proceedings have taken, and the enormous costs incurred in those proceedings, have created burdens on, and material uncertainties for, all the parties involved. In itself, however, it is neither exceptional nor reprehensible that such proceedings should be complex and prolonged: after all, the tracks are owned by BHPBIO, and intruding on property rights is obviously a decision that should be taken with great care. But the proceedings have focused attention on the wide-ranging potential applicability of Part IIIA of the TPA, and the risks it creates for facility owners. It is therefore timely to examine the issue of third-party access, all the more so as Part IIIA is approaching the fifteenth anniversary of its enactment, meaning that sufficient experience has accumulated with its provisions to allow an assessment of its strengths and weaknesses to be made.

The objective of this article is to consider what it is that Part IIIA might be doing from an economic perspective, and what that tells us about when mandated third-party access might be efficient.

I start by considering the standard monopoly or 'bottleneck' explanations for refusals by a profit-maximising, vertically integrated monopolist to provide a more-efficient third party with access to a potentially shared input. I show that those explanations are rarely convincing, but that when they are, what is involved is a trade-off between productive efficiency gains the vertically integrated firm could make by 'contracting out' use of that input to the assumed more-efficient rival and a transactions costs loss associated with the monopolist's inability to write contracts that could claim for itself the resulting gains from trade. I suggest that when that trade-off results in no access being provided (that is, a corner solution in which the monopolist sets the supply of third-party access services to zero), the most likely explanation is not monopoly, but rather that the transactions costs associated with access exceed any potential efficiency gains.

Assuming nonetheless that cases exist in which the monopoly explanation might hold water, I then examine third-party access from the perspective of the economics of property rights. This examination highlights how stringent are the conditions that must be met for the cure to be better than the disease. A particular focus of my discussion is on the extent to which access prices can

3 BHP Billiton, Annual Report, 2008: 37.

be relied on to support efficient, decentralised, decisions by the access provider and the access seeker about use of the shared facility; I show that particularly when access is to a facility that sits within an interdependent complex of facilities, no set of access prices may exist that can play that role. In those cases, there will be an ongoing need for centralised, administrative control of that shared use, much as would occur within a vertically integrated firm. I suggest that as a general matter those control functions will be most efficiently provided by the vertically integrated firm itself, but that creates risks that the firm will discriminate against (that is, 'sabotage') the rival. Attempts by the regulator to prevent such discrimination impose costs of their own, and in the presence of those costs, a welfare-maximising regulator would set relatively high access charges (as high access charges make discrimination less or un-profitable). The result is to further reduce the social gains from access, which are obviously even smaller if the regulator does not implement the welfare-maximising solution. The result is that mandated access can reduce welfare rather than enhance it. This suggests that stringent tests should be met before third-party access is granted. In a companion piece,⁴ I examine Part IIIA in the light of these findings and conclude that it does not provide a sufficiently stringent 'filter' against unwarranted extensions of mandatory third-party access.

At points in this article, I will illustrate my propositions by reference to material submitted in the lengthy proceedings between BHPBIO and FMG. I should stress that my purpose in doing so is merely to give some concreteness to the points made. I should also stress that I rely on that part of the material BHPBIO has filed that is in the public record. But, of course, that material has not yet been tested in court, at least as far as its economic substance is concerned. The reader should view this material with that caveat in mind.

Part IIIA of the *Trade Practices Act* 1974

It is fundamental to our system of property law that a firm has a right to choose with whom it wishes to deal and on what terms. Thus, in *Colgate*, the United States Supreme Court stated that even a monopolist can 'exercise his own independent discretion as to the parties with whom he will deal'.⁵ In this vein, the ownership of property brings with it a concomitant bundle of rights which include the freedom to exclude others from the property, to derive benefit from

4 Ergas, Henry 2009, 'Access, Property Rights and Welfare: An Examination of Part IIIA of the Trade Practices Act', at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1428184.

5 U.S. v Colgate & Co 250 U.S. 300, 307 (1919).

the property, to retain possession, and obtain legal redress for violations of that possession. Of course, property rights have never been unqualified, and are often abrogated for a public or private purpose.⁶

Part IIIA of the TPA is just such an intrusion into the property rights of firms. Enacted in 1995, Part IIIA was one of several major competition law reforms recommended by an Inquiry chaired by Professor Fred Hilmer.⁷ Given that I discuss the detail of its provisions in the companion piece mentioned earlier, suffice it to say that Part IIIA provides for mandated access to the facilities of a firm where that access would meet a number of conditions and, in particular, promote competition in a relevant market. That is, in broad terms, where competition in a market is dependent upon access to a 'bottleneck' or 'essential facility', the Part allows for imposing a right of third-party access to that facility by means of 'declaration' of the service provided by that facility.

Obviously, declaring a service would seem of little use without concomitant rights as to price and quality, as the access provider, who was being coerced into providing access, would have incentives to increase the former and degrade the latter. Thus, Part IIIA also establishes a mechanism by which the Australian Competition and Consumer Commission ('the ACCC') arbitrates the terms and conditions of access to declared services, should the access provider and the access seeker be unable to reach a commercial agreement.

The stated objective of Part IIIA is to improve economic efficiency in the provision of infrastructure inputs and, by doing so, promote effective competition in upstream and downstream markets.⁸ Declaration, and the rights and obligations that flow from it, are the key instrument by which it seeks to achieve that objective.

To that end, what declaration does, in its substance, is to alter the relationship in terms of legal rights in the supply of infrastructure inputs between the access provider and the actual or potential access seeker. Thus, prior to declaration, the extent of that relationship, if any, is based solely on commercial negotiation, and the access provider has the option of entering or not entering into those negotiations and, having entered into them, of agreeing or not agreeing to access. In contrast, once the facility is declared, the access provider must enter into negotiations over the supply of infrastructure inputs and is no longer free to determine whether access will be provided and if so on what terms: in that

6 For example, all legal systems have some mechanism by which the state can compulsorily acquire privately owned land. Likewise, many legal systems, including Australia's, have a form of private redress whereby an owner of property can obtain an 'easement', by compulsion, over adjacent blocks — so as to provide access, drainage or similar.

7 National Competition Policy Review ('Hilmer Report'), 1993, Chapter 11.

8 The TPA, s 44AA.

sense, the option of access (that is, the right, but not the obligation, to make use of the facility) is shifted, at least in part, to the access seeker, albeit within the limits imposed by the outcome of any arbitration in the event of dispute.

As a result, the provisions of Part IIIA have two major effects. First, they create the scope to impose on owners of infrastructure inputs an obligation to supply third parties, with whom they might otherwise not have dealt, the services provided by those inputs. An owner and operator of a conveyor belt, which it uses in producing goods for sale to end-users, can, in other words, be obliged to supply 'conveyor-belt services' to a third party in circumstances where no such services might otherwise have been supplied.

Second, in imposing that obligation, Part IIIA replaces one *mechanism* or *process* generating outcomes (purely commercial negotiations between the potential access provider and the access seeker) with another mechanism or process generating outcomes (the 'negotiate/arbitrate' model that applies to declared services, in which the parties negotiate commercially for access, and should that fail, the ACCC can determine the terms and conditions of access through a compulsory arbitration).

What is Part IIIA doing?

As noted above, it is an important element of conventional property rights that they allow property owners to choose with whom they deal and on what terms. It is therefore worth standing back and asking exactly what Part IIIA might be trying to do, and whether that objective could be justified (and, if so, when).

In saying this, I recognise that to many readers, the answer will seem entirely obvious and, indeed, as having been stated above: Part IIIA is trying to address a problem of monopoly. The owner of the process referred to above — with the conveyor belt that it uses to produce goods (call them, as tradition requires, widgets) that it sells to final consumers — has, for some reason, a monopoly over conveyor belts (perhaps at a particular location); Part IIIA ensures that it does not unreasonably deny third parties 'the right' to use that conveyor belt; that is, that it does not use its monopoly in the upstream stage (conveyor-belt services) to exclude socially desirable competition in downstream supply (the sale of widgets). Seen in this perspective, the evil Part IIIA is designed to address is that of the incentive of monopolists to deny to otherwise efficient rivals access to 'essential facilities' that they control.

Of course, merely *assuming* that the widget supplier is a monopolist, simply because it refuses to supply conveyor-belt services, is completely unsatisfactory — for (as we will see below) there are many other reasons that might underpin

that refusal to supply, not least efficiency. And if it would be efficient to supply such services, then plainly a widget supplier that operated in a competitive market would have every incentive to do so: which means that if the market is competitive, then a refusal to supply can be presumed to be efficiency-enhancing. As a result, before falling back on the monopoly explanation, one would want to be convinced that there was indeed a monopoly facility there.

However, even if there was such a monopoly, the monopoly explanation is still unsatisfactory. In effect, no profit-maximising monopolist has a general incentive to refuse to supply: on the contrary, monopolists want to supply, albeit at a price that allows them to claim a greater share of the gains from trade than they could if buyers could turn to good alternatives. Why would a monopolist over conveyor-belt services refuse to sell those services at a monopoly price?

The answer that immediately suggests itself is that it would refuse to sell those services so as to support a monopoly price for the final good, in this case the widgets: put in the coloured language that is often used, the monopolist is seeking to 'leverage' its control over the essential facility into an unnecessary and socially undesirable monopoly in the potentially competitive downstream market. But this too seems puzzling.

It is puzzling because it suggests the monopolist forgoes potential gains from trade. Thus, one presumes that what we are interested in is efficient competition — after all, the objective of Part IIIA is to enhance efficiency, and if potential users of the conveyor-belt service are *less* efficient than the vertically integrated monopolist, then society can readily lose if that monopolist is forced to supply them with the service.⁹ But if the potential downstream rivals really are more efficient than the conveyor-belt monopolist in transforming conveyor-belt services into widgets, why wouldn't the conveyor-belt monopolist simply charge those rivals a price for conveyor-belt services that captured at least some of the surplus from that superior efficiency? Why wouldn't it, in other words, contract out to these more efficient 'transformers' the task of converting conveyor-belt services into widgets, claiming for itself as profit the resulting gain in productive efficiency?

One can construct situations in which the conveyor-belt services monopolist would want to exclude even efficient rivals from competing in parts of the widget market. The general structure of these situations can be illustrated by an example.

⁹ There could, of course, be a trade-off here. Society might gain from an allocative efficiency improvement, if the entry of a competitor were to bring prices closer to marginal costs. However, if the entrant is less efficient than the incumbent, there is a productive efficiency loss. Impacts on dynamic efficiency could go either way. Usually, a productive efficiency loss will be larger than an allocative efficiency gain (rectangles versus triangles), so the trade-off would not favour access.

Thus, assume that due to differences in willingness to pay between end-consumers there are profits to be made from price discrimination in the market for widgets but that the monopolist cannot devise a non-linear price schedule for conveyor-belt services that fully extracts those profits through the sale of conveyor-belt services to potential downstream rivals. Even then, however, it is not apparent that the profit-maximising monopolist would generally refuse to supply such services (thus reserving for itself the entire downstream market). Rather, the logic would require it to segment the market by setting a high price for services sold to downstream competitors (the 'access' charge) while setting a low widget price to customers with a low willingness to pay, whom it would only serve itself, along the lines of the Alcoa case discussed by Perry (1980). In that way, the more efficient 'transformers' could at least serve the high-willingness-to-pay segment, with the conveyor-belt monopolist claiming through the 'access' charge some part of the additional profits from their superior efficiency, while the monopolist reserved for itself the low willingness to pay part of the market (and so did not entirely forgo whatever willingness to pay those customers had).

What is happening in this example (and is generally true in cases where some kind of exclusion of no less-efficient rivals by control over inputs is at issue) is that the conveyor-belt monopolist faces two distortions that prevent it from extracting all the surplus that might be available:

- The first is its own productive inefficiency, relative to potential downstream rivals, which makes those rivals more efficient 'transformers' of conveyor-belt services into widgets. The result of this inefficiency is that the monopolist, when it serves the downstream market, gives up some rents a lower cost monopolist could have obtained.
- The second is a transactions cost inefficiency, reflected in the monopolist's inability to devise or implement the non-linear price schedule for the upstream input (in this case, conveyor-belt services) that extracts all the rents in the downstream market. (Transactions costs are the costs of search, information, bargaining, decision, policing and enforcement associated with rights and contracts that bear on rights.¹⁰) The result of this inefficiency is that when the monopolist allows rivals to serve that market, it loses some rents that a no-transaction-cost upstream monopolist would have obtained through the optimal (profit-maximising) access contract.

Given these two distortions, the profit-maximising monopolist selects its own level of output, relative to that of these more efficient 'transformers' in a way that

10 Following Dahlman (1979: 148), endorsed by Coase (1988a: 6). More specifically, transactions costs comprise the costs incurred by parties to a contract in order to find each other, exchange information, delineate the property rights to be transferred, draw up the conditions of the contract, signal preparedness to fulfil their obligations, monitor each other's performance and otherwise see to it that obligations are fulfilled, and take action if obligations are not fulfilled.

balances the cost of each distortion at the margin. In other words, it increases sales of conveyor belt services up to the point where the gain it makes from claiming some of the cost reduction effected by the more efficient 'transformers' is just balanced by forgoing a share of rents it can obtain from serving the market itself (albeit at higher cost). Clearly, for the profit-maximising outcome of this balancing to be a corner solution — in which the monopolist sets sales of conveyor-belt services to the more efficient 'transformers' at zero, thus entirely forgoing the profits achievable through their superior efficiency — special assumptions are needed (for example, that the efficiency rents the monopolist can extract from the more efficient 'transformers' are not large enough to cover the product-specific fixed costs of supplying access to its conveyor belt as a marketable service: which suggests that requiring it to sell those services may not be efficient, a point I return to below).

As a result, the claim that the refusal to supply is an exercise of monopoly power should not be accepted without very close scrutiny. There is an alternative explanation, which in many situations will be more compelling (and require fewer special assumptions): the transactions costs of selling access exceed the potential efficiency gains.

Thus, all else being equal, wealth-maximising agents have incentives to replace higher transactions-cost contracts with lower transactions-cost contracts.¹¹ The Nobel laureate Ronald Coase famously explained the formation and organisation of firms in these terms. Viewed through the prism of Coase's formulation (Coase 1937), firms form when there is scope to replace high transactions-cost contracts for intermediate *outputs* (for example, a contract to operate a conveyor belt from A to B) with lower transactions-cost contracts for the primary *inputs* (including labour services) required to produce those same outputs. Once the firm has acquired the property right to those primary input services, it can direct the use of those services as flexibly as the contract for those inputs permits, which is usually assumed to be more flexibly than would be possible under an (arm's length) contract for the intermediate services those inputs can be used to produce. This flexibility increases the wealth that can be produced from these inputs and in that sense creates a gain from trade.

However, primary input contracts also give rise to their own types of transactions costs, most notably principal/agent problems, with a classic example being shirking under labour contracts. This reflects the inherent incompleteness of contracts, which always leave some margins of use (or dimensions of the resources being traded) not fully specified and, in that sense, in the public domain. One

¹¹ Of course, this is provided that it does not give rise to other costs or reduce output. In other words, firms will economise on transactions costs in exactly the same way as they economise on other forms of costs, equalising the gains from doing so at the margin of cost-reduction efforts.

would predict that the bounds of firms' activities will extend to the point where value of the increased flexibility from replacing higher transactions-cost contracts with lower transactions-cost contracts equals the increased costs associated with the 'unexploited margins' inherent in input contracts — see Barzel (1989) and Demsetz (1997).

It follows that if firms use primary inputs (such as labour, capital and raw materials) to produce intermediate input services (such as conveyor-belt services) that in turn serve to produce final goods (widgets), this may well be because producing those intermediate input services internally avoids transactions costs that would be incurred in buying them on the market — an inference that receives strong support from the empirical literature on the determinants of vertical integration, as surveyed by Lafontaine and Slade (2007). As a result, a candidate explanation for the refusal to sell 'conveyor-belt services' to third parties is that doing so would incur transactions costs penalties that exceed any potential efficiency gains.

However, even assume for the moment that the monopolist is in fact refusing to sell conveyor-belt services so as to support monopoly prices for widgets. This is still at best incomplete as an explanation or justification for Part IIIA.

In effect, the question that must be asked is why society would deal with the widget monopolist by forcing it to supply conveyor-belt services rather than by regulating the (assumed otherwise too high) retail price of widgets. There must be some efficiency gain from intervening in the upstream market, above and beyond whatever efficiency gains might be obtained by regulating downstream supply.

Presumably, this additional gain can take two forms. First, intervening in the upstream market may allow greater efficiency in the supply of 'transformation' services — that is, in the services of transforming conveyor-belt services into widgets. Second, it may be that it is easier to regulate the terms and conditions of supply for conveyor-belt services than for widgets — in other words, that the risk and cost of regulatory error is lower in setting prices and other supply conditions for conveyor-belt services than for widgets.

Whether either of these gains would be substantial is an empirical question. But one might well be sceptical that they would be. Starting with the social gains from allowing in more efficient 'transformers', surely the monopolist has every incentive to reduce the costs incurred by its own 'conveyor-belt services' division as much as possible. Moreover, if despite those incentives that division is materially less efficient than rivals, that would tend to make any refusal of supply unprofitable, as such a refusal would mean a greater sacrifice of potential profits. As a result, all else being equal, where there really are large differentials

in the efficiency with which ‘transformation’ occurs, refusals to supply would have to be unusual. This leads one to suspect that the extent of the first type of gain may not be all that great, at least where firms are indeed profit-maximisers.¹²

As for the ease or difficulty of getting regulated upstream prices ‘right’ relative to the ease or difficulty involved in regulating the downstream price, typically, the supply of intermediate inputs produced within a vertically integrated firm involves higher transactions costs and contracting difficulties than the supply of that firm’s final goods and services. This follows directly from the standard explanation for the existence of firms just discussed. But if the firm is then obliged to sell those intermediate input services to third parties, presumably those transactions costs will need to be incurred (offsetting, and at least in some instances entirely negating, any gains from the scope for more efficient ‘transformers’ to enter the market). In contrast, the firm already sells widgets, and hence regulating the price of widgets (whatever other costs and distortions that may impose) will not require those transactions costs to be incurred.

Moreover, the risks and costs of regulatory error are not likely to be any smaller in regulating the supply of conveyor-belt services than in regulating the price of widgets — rather, the exact opposite would seem probable. Thus, with conveyor-belt services, the regulator will likely have less information and fewer benchmarks to go on, not least because those services are not traded. Moreover, the optimal charging structure for such services is likely to be complex, involving multipart prices that reflect the underlying structure of costs and of the downstream profit functions of the competing users.¹³ Devising both the level and the structure of such charges poses obvious challenges. And any errors in setting prices can cause social losses that are both high in themselves and greater than those arising from errors in setting final prices. This is because above and beyond any allocative inefficiency, which would likely also arise from price regulation of the final market, there will be an additional productive inefficiency loss across *all* of output when prices for intermediate inputs are incorrectly set (that is, a rectangle as well as the conventional triangle), as well as potential losses in dynamic efficiency from distorting the incentives for investment by both the access provider and the access seeker.

Additionally and importantly, mandated third-party access redistributes income between market participants — most obviously between the access provider (whose shareholders will generally lose from the mandated access requirement)

12 If firms are not profit-maximisers, as might be the case for government-owned businesses, then presumably the refusal to provide access is unlikely to be related to protecting monopoly profits downstream. That said, it might be due to protecting rents that are being taken in the form of unduly large ‘empires’, cross-subsidies to politically favoured constituencies or featherbedding in the labour market.

13 The downstream profit functions are relevant to determining the Wicksell-Lindahl prices that efficiently recover the common costs of the facility, assuming it has public or club good characteristics.

and the access seekers (which at least in the short run can hope to gain). The scope for such redistribution can create formidable incentives for 'rent seeking': that is, the making of investments aimed not at expanding society's capacity to produce, but rather at securing for oneself a larger share of any economic rents that may be available. These incentives are especially strong where the redistributive effects involve narrow groups, as will be the case in instances involving third-party access (since at least in the first instance, the redistribution third-party access effects is a small numbers situation involving the access seeker(s) on the one hand and the access provider on the other), for then the risk of free-riding does not undermine the ability of the affected parties to mobilise resources and make rent-seeking investments. However, since those investments are a cost (they absorb resources that could have been put to productive use elsewhere), but themselves do not increase output, they amount to a pure waste of society's scarce resources. Economic theory predicts that such rent-seeking can waste up to (and in some cases even more than) the available rents.

As a result, while all forms of regulation impose costs — so that the distortions arising from 'government failure' may be no less than those of 'market failure' — the costs of access regulation can be especially high. It would consequently be wrong to presume, much less act on the assumption, that access regulation will deal with any market failures at lower social cost than would arise from not regulating at all or (should regulation be absolutely required) from regulating only in the final product market.

In short, monopoly explanations, to paraphrase George Stigler, have powerful sway over untutored minds. But looked at closely, the monopoly explanation for Part IIIA is not all that convincing, and many assumptions need to be made before it seems likely that profit-maximising monopolists would refuse access entirely, and do so in circumstances where mandatory access would provide an efficient remedy.

Another view of the cathedral

As noted above, Part IIIA does not merely make it possible to impose on owners of infrastructure inputs an obligation to supply third parties, with whom they might otherwise not have dealt, with access to the services provided by those inputs. Additionally, in imposing that obligation, Part IIIA also replaces one *mechanism* or *process* generating outcomes (purely commercial negotiations between the potential access provider and the access seeker) with another mechanism or process generating outcomes (the negotiate/arbitrate model that applies to declared services).

As a result, another way of looking at what Part IIIA is trying to do (and whether and when it might be justified) is to ask: 'Under what conditions will purely commercial bargaining between the potential access provider and the access seeker fail to achieve efficient outcomes, and justify being replaced by a mechanism based on some form of collective fiat?'

Now, the purely commercial bargaining between the potential access provider and the access seeker is based on the access provider's property rights, including (to continue my earlier example) over its conveyor belt, and involves the willingness or otherwise of the access provider to sell some part of that property right to the access seeker. Whether that bargaining will yield efficient outcomes therefore depends on the factors that affect the efficient allocation and reallocation of property rights.

In considering these factors, it is useful to start from the economics of property rights, as it has emerged from the work of Ronald Coase.¹⁴

From an economic perspective, a property right is a right to determine the use to which an asset will be put, including the right to transform or transfer the asset. In addition, it entails the right to the residual income generated by that asset. Clear property rights bring many benefits, most significantly the incentives to maximise the net wealth generated by an asset through expansion, contraction or disposal decisions (Demsetz 1967; Alchian 1977; Barzel 1989). At the same time, by vesting decision rights in an owner (or a clearly defined group of owners), properly specified property rights allow wealth-maximising decisions to be taken quickly as economic conditions change. They also create the conditions for rights to be traded, which permits collections of related assets to be brought under common ownership, minimising 'spill-overs' or unpriced interactions between separate rights owners (which would otherwise give rise to inefficient decisions, as decision-makers would not face the full costs and benefits of their decisions). The scope for such aggregation of rights is also crucial in permitting production at efficient scale and scope.

Following Coase (1960), if in a two-party framework a property right is well defined and the initial holding of that right is clear, then the right will end up with the party able to achieve the highest net wealth, regardless of whether that wealth-maximising party is the initial holder or someone who buys it from the initial holder, provided that there are no transactions costs and no wealth effects.¹⁵ This result is quite independent of market structure: a monopolist

14 Coase compared a situation in which property rights were poorly specified, or (for particular aspects of a transaction, such as the sparks generated by a train passing on rail lines) not defined at all, with a situation where property rights were fully specified. Calabresi and Melamed (1972) extended the discussion to deal with different types of property rights.

15 This is the celebrated Coase theorem, put broadly in the form framed by Stigler (1966) and endorsed by Coase (1988a: 157ff). It follows that (assuming no transactions costs and no wealth effects) the use of the assets

has every bit as much incentive to seek more efficient uses of resources as has a competitive firm and in a world with fully specified property rights and no transactions costs, full efficiency would prevail, regardless of market structure.

In practice, however, transactions costs are never zero. The presence of transactions costs means that contracts can never be 'perfect'. It also means that property rights are never fully specified: there are always some goods or attributes that are left 'in the public domain', either because there is no practical way of metering them or because the costs of establishing and enforcing property rights over them exceed the benefits.¹⁶ These 'unpriced margins' (in the sense of goods that individuals can claim for themselves without facing the opportunity cost of their doing so) are then allocated by non-market means, such as the 'first come, first serve' competition for the best seats at a cinema. The resulting allocations — such as those that occur through queuing — may be inefficient, though there remains the question of whether the gains from attempting to cure the inefficiency by creating a property right over that 'unpriced margin' would exceed the costs (see generally Parish 1980, and Barzel 1982 and 1989). 'Externalities' and other third-party effects are merely one form of such an unpriced margin.

Transactions costs and unpriced margins can result in forgone opportunities for efficiency-increasing reallocations of rights. Familiar instances include the difficulty of negotiating to a wealth-maximising outcome in the presence of diffuse externalities, such as air pollution, noise pollution or road congestion. In these instances, the unpriced margins are dissipated by over-production, and involuntary over-consumption, of the relevant external effect, with transactions costs being too high to allow coalitions of the producers and consumers of those effects to negotiate their way to a more efficient use of resources. In conventional monopoly pricing, a related but somewhat different form of transactions cost leads to the inefficiency: the inability of the monopolist to determine and impose on customers prices that extract the entirety of their willingness to pay for the good the monopolist controls. The result, most evident when the monopolist is constrained to linear (that is, uniform) prices, is under-use and a deadweight loss, with some potential gains from trade being forgone and too little consumption of the monopolised good.

This suggests that there are two possible reasons why purely commercial negotiations between an access provider and one or more potential access seekers could fail to achieve an efficient outcome.

to which the property right pertains will be the same, irrespective of the initial holding (Werin 2003: 78).

¹⁶ In respect of some goods — such as one's own body and elements thereof — there are also ethical (that is, non-consequentialist) reasons for not creating conventional property rights.

The first is that the parties at issue — the potential access provider and the aspiring access seeker — may not bring to the negotiation all of the social benefits and costs of a potential reallocation of property rights. There may, in other words, be third-party effects that are unrepresented in the negotiation, and that drive a wedge between the interests taken into account in the bargaining and society's interests overall. The most obvious such third-party effect arises when the benefits of access would flow to final consumers, in the form of lower prices or higher quality, so that the total gains from access would exceed the gains accruing to the negotiating parties. Those parties might then fail to reach agreement (or agreement on reasonable terms), even though such an agreement would yield social benefits in excess of its costs. Of course, whether such additional gains are there to be had depends on all the issues reviewed above in discussing the 'monopoly' view of third-party access.

Second, even if there are no third-party effects, so that the interests represented in the bargaining capture all of the relevant gains and losses, other transactions costs may nonetheless prevent an efficient outcome being achieved. The transactions costs most relevant here are likely to be those associated with information asymmetries. These arise when one party to a (potential) bargain has more or better information than another. Information asymmetries will, as a general matter, limit the efficiency of bilateral bargaining (Wilson 1998). More specifically, the respective parties may not know the costs and benefits the other would secure from an agreement, and have incentives to misrepresent those costs and benefits in ways that impede, and can entirely prevent, agreement from being reached (Roth 1985). It is therefore not uncommon for 'money to be left on the table' in commercial negotiations, with the resulting losses taking the form of strikes, costly legal battles, or simply lost opportunities.

The question of how society might deal with these inefficiencies — when they engage wider social interests — was central to an important extension of Coase's work on property rights by Calabresi and Melamed (1972).

In that article Calabresi and Melamed defined a 'property rule' as a rule which granted the owner of property the right to preclude others from impinging on that property. Thus, the beneficiary of such a rule would have a right to prevent (for instance, through injunction) another party from engaging in conduct that would impose on that property a cost or other adverse effect. Given that right, the owner would then have the option (but no obligation) to negotiate with the other party for compensation in exchange for giving that party a more or less limited right to impose that third-party effect; it would, in other words, have the ability to sell, in whole or in part, its right to be free of the third-party effect, but

would be under no compulsion to do so. If it did so, and set the compensation efficiently (so that, at the margin, it signalled the cost or benefit associated with the externality), this would align the parties' respective interests.¹⁷

However, as Calabresi and Melamed noted, the consequence of relying solely on such property rules is that some potentially wealth-maximising reallocations of rights might not occur, because the transactions costs to the parties of negotiating and implementing those reallocations are too high.

One response to this, they suggested, is to replace the property rule (which allows the party owning the property to enjoin the other party's non-consensual use of its asset) with a 'liability rule'. Calabresi and Melamed defined such a 'liability rule' as one which allows one party to act in a way harmful to another in exchange for a compensating payment as determined by a third party, such as a regulatory agency or a court. The option of imposing costs is then shifted to the potential infringer, but subject to a mechanism which requires a compensating payment; that is, faces that infringer with the equivalent of a market price for the cost-imposing action (such as emitting pollution, or running trains that set off sparks, occasionally causing fires in adjacent pastures). They then argued that such liability rules should, from an efficiency perspective, be relied upon when at least two conditions are met:

1. Transactions costs will prevent wealth-maximising reallocations of rights from occurring under an approach based on property rules alone; and
2. A liability rule, and especially the conditions associated with that rule (including payments for costs imposed on the rights-owner), can allow those transactions to occur in a way that, given the costs and benefits of the rule, on balance increases society's wealth.

In other words, as Calabresi and Melamed stressed, the decision to rely on a liability rule requires an assessment of which of market transactions or collective fiat is most likely to bring the allocation and re-allocation of rights closer to the Pareto-optimal result the 'perfect' market for rights would reach.

Seen in the perspective of Calabresi and Melamed, third-party access regimes, such as Part IIIA, convert a property rule — under which the owner of a facility can choose whether or not to allow third-party use of that facility — into a liability rule.¹⁸ Under that liability rule regime, the option of using a facility is transferred from the property owner to one or more potential users

¹⁷ An important extension of Calabresi and Melamed that explains the shifting of rights that occurs through liability rules in terms of options contracts is in Ayres (2005).

¹⁸ See Ergas 2008.

under conditions that include liability to make a payment for that facility's use. Such conditions would also include the procedure or mechanism whereby the payment is (in the event of a dispute) determined by an independent third party.

As a result, the question in assessing third-party access must be whether the conditions Calabresi and Melamed identified as justifying the imposition of such a liability rule are in fact met.

'Market failure' versus 'regulatory failure'

Before turning to consider that question in the specific context of Part IIIA, it is worth paying closer attention to the second leg of the Calabresi–Melamed conditions: whether a liability rule, managed by a regulator or a court, can be expected to achieve the potential Paretian improvements in the allocation or usage of property rights when private bargaining fails to do so. The question to be addressed here, in other words, is: 'Can the liability rule fill the gap efficiently?' It is easy to blithely assume that regulatory processes, despite their recognised imperfections, can 'clean up the mess' in situations where private bargaining fails; but the reality is that numerous and important obstacles may prevent this from occurring.

The obvious obstacle to the liability rule succeeding in improving on private bargaining is the fact that information is inherently imperfect and asymmetric as between the regulator and the owner of the facility to be regulated. Indeed, though the regulator may be able to rely on its powers to compel the production of information, the underlying differences in knowledge and understanding between the regulator and the regulated are likely to be at least as great as those between the commercial parties.

It is a familiar result in the modern theory of regulation that from an efficiency perspective, an imperfectly informed regulator, operating in circumstances where a high weight must be placed on ensuring the regulated firm continues to produce, should set the cap that controls the firm's prices in such a way as to allow the regulated firm the expectation of substantial economic profits (see Gasmi *et al.* 2002). Gasmi *et al.* also show (Table 7.12) that when the regulator is not certain as to costs (including the appropriate rate of return), price caps are the optimal regulatory mechanism, even though they allow firms to earn above normal earnings (p.180; a price cap with a profit tax is superior to a price cap on its own, but the main effect of the profit tax is to increase consumer welfare: it reduces the firm's profitability only moderately). And the greater the gap in information between the regulator and the regulated, the greater the 'information rent' it is efficient for the regulator to allow the regulated firm.

For reasons noted above, these issues of information imperfections are likely to be especially acute in respect of intermediate input services, such as the conveyor-belt service our hypothetical widget monopolist supplied to itself. This sets inherent limits on the extent to which regulation can eliminate the vice of monopoly pricing in respect of such services. Moreover, it is plain that a reluctance by the regulator to allow those ‘information rents’ to be claimed by the regulated firm can be even worse than allowing the firm to set the unregulated, monopoly, price, as it can lead to regulated prices that undermine the incentives to produce and invest (and hence sacrifice both the consumer and producer surplus).

Limitations on the price mechanism

However, the difficulties facing the liability rule may be even more serious than conventional discussions of information imperfections suggest.

Especially important in this respect are the inherent limitations of price systems. Of course, economists are (quite properly) great admirers of the price system, and usually assume that reliance on price signals can allow decentralised decisions to secure efficient outcomes. That prices should be able to do so is central to the efficiency of the liability rule, as the rule essentially involves setting a ‘price’ that signals to the cost-imposer the social costs of its actions. For example, when a liability rule is imposed on a polluter, requiring that polluter to pay those adversely affected a ‘price’ equal to the marginal social cost of pollution, this allows the decentralised decision-making by the polluter on the one hand, and the victims of pollution on the other, to achieve efficient levels of pollution, thus securing the potential gains from trade in pollution rights. Equally, an efficient price for third-party access to an ‘essential facility’ supports decentralised decision-making about the supply and use of that facility, allowing the access provider to take efficient decisions about the level of investment in access-related facilities, while allowing the access seeker to determine its level of use of those facilities in the light of its own costs and downstream demands (which it knows but the other parties do not).

However, it is an obvious mistake to imagine that such an efficient ‘price’ mechanism always exists. Indeed, it follows from Coase (1937) that the fact that activities are organised within a firm — that is, through a system of contracts for input services, rather than through contracts for outputs — in itself suggests that devising a price system that could efficiently coordinate those activities is uneconomic. As an influential text on organisational control (Sunder 1997: 45) notes, this explains the very limited use firms make of transfer prices as a way of coordinating their internal activities:

The transfer pricing problem is often stated in a manner that ignores the circumstances that give rise to the problem in the first place. Integration of the two divisions [used in setting out the simplest version of the problem] into a single firm is predicated on the cost of market exchanges between them being greater than the administrative and agency costs. After integrating two divisions under this rationale, one cannot then turn around and wish the administrative and agency costs aside through an appeal to a costless but non-existent system of market exchange.¹⁹

Part of the problem is that where demand and cost structures have particular characteristics, it can prove impossible to define a resource allocation process that is both:

(a) decentralised, in the sense of requiring only:

i. that each firm have information about its own production possibilities, as against also knowing the production possibilities open to other firms; and

ii. that each firm's message at any step should concern its own proposed actions at that step, as against also concerning the proposed actions of other firms at that step; and

(b) will result in efficient use of the resources being allocated (Heal 1973: 142ff).

Thus, it is well known that where production systems involve chains of processes, with each link in the chain being 'lumpy', in the sense that investments involve minimum increments of fixed, often large, size, efficient investment and use decisions may not be made without some degree of coordination, above and beyond that provided by price signals (Heal 1973: 142ff; Schelling 1983; Roberts 2004: 51ff; and, for a specific application to rail networks, Quinet 2003).

This is because decisions taken under those circumstances depart from two crucial assumptions conventionally made in economic models of decentralised choice of output and investment:

- that the set of alternatives is convex, which means that if two choices are available, any combination of those choices (that is, any intermediate point) is also available; and
- that the objective function (that is, the function that measures how outcomes, such as profits, vary depending on the decision taken) is concave, which means that if two choices lead to the same outcome, then a combination of

¹⁹ Obviously, firms use transfer prices for other purposes, such as determining tax liabilities. However, according to Sunder, they play a limited role in guiding resource-allocation decisions.

these choices would lead to a higher outcome (so that the objective function looks like a smooth climb to, and descent from, a single peak).

As Roberts (2004) explains, when both these assumptions hold, there is a single optimum. Moreover, that optimum can be found by exploring small variations around the current point, as a simple rule can be followed: change any dimension of performance, and if that change improves performance overall, continue to change it in the same direction until the improvements become smaller than the costs. As a result, it is easy to decentralise decision-making through the price system, allowing each unit to explore the options and assess the impact on its performance, without losing overall efficiency.

However, where investment decisions are indivisible (one cannot continuously vary the number of tracks, mines or ports), the set of alternatives is non-convex. For the same reason (and because of increasing returns to scale associated with redundancy in capacity), the objective function is non-concave. Put in practical terms, this has two implications.

The first is that there may be multiple outcomes that are ‘local’ equilibria, in the sense that small variations around those outcomes will result in worse performance. As a result, any decentralised process that involves individual decision-makers ‘exploring’ whether performance would improve by making slight changes around an initial position will not be able to identify the global (that is, overall) optimum — that is, the configuration of assets that maximises the value of the system as a whole.

Second, changes in any one dimension of the system at a time may not suffice to indicate whether that dimension should in fact be varied; rather, finding the optimal configuration requires changing (or at least modelling changes in) all aspects of the chain at once.

As a result, there is a need for some non-price-based mechanism that coordinates investment, production and use along each of the links in the chain. It is this that administrative coordination through vertical integration within the firm provides.

Contractual incompleteness

Moreover, these difficulties in determining prices that can support decentralised decision-making are compounded by problems of contractual incompleteness. This refers to the difficulties of devising and enforcing agreements that would manage shared use of the firm’s common resource (that is, our hypothetical conveyor belt) efficiently. Three problems are likely to be especially acute:

(a) Any agreement may not be able to fully *control for the quality* of the access seeker's use of the facility — the access seeker may, in particular, have the incentive and ability to reduce its own costs at the expense of degrading the facility or otherwise increasing costs to the facility owner.²⁰ The impossibility of negotiating and enforcing complete or 'perfect' contracts and of setting fully efficient decentralised prices can, in other words, give rise to a 'tragedy of the commons', in which access seekers exploit any unpriced margins in a way that reduces the aggregate value of the shared assets (Barzel 1989).

(b) The relation between the access seeker and the facility owner may give rise to *hold-up risks*, particularly if the access seeker can impede the facility owner from responding to changes which require alteration to the pattern of facility use (Pirrong 1993). These hold-up risks can invite otherwise inefficient 'counter-measures' by the facility owner, such as investing in excess capacity (so as to have more of a 'shock absorber' in the event of disagreements about facility use) or, conversely, restricting capacity below efficient levels, if that allows the access seeker to be 'rationed off' the system.

(c) Further hold-up risks can arise in *the investment process*, to the extent to which that process becomes vulnerable to being hindered (for instance, through regulatory delays) by the access seeker. These risks are a form of what is now often referred to as the tragedy of the 'anti-commons', in which vesting veto rights in parties, each with incentives to act opportunistically in seeking for themselves a greater share of the joint pie, prevents efficiency-enhancing transactions from being undertaken (Heller 2008).

Controlling these risks, too, requires both considerable information — imposing an information collection, verification and analysis burden — and an authoritative process for resolving inevitably conflicting 'stories'. Of course, similar needs arise, by the nature of the situation, within the vertically integrated firm; but there is, then, a greater alignment of interests between the parties, and a greater subjection by those parties to common hierarchical control, than is the case in the relations between the access provider and the access seeker.

²⁰ This reflects the difficulty of defining the service in terms that are readily verifiable — the 'state verification' problem (Hillier 1997: 57ff), which then gives rise to wealth-reducing use (Barzel 1989: 13ff; Barzel 1982).

Problems evidenced in BHPIO/FMG

Both of these factors that can undermine the efficiency of the price mechanism (and hence of the liability rule) — the impact of non-convexities on the existence of a price-based, decentralised decision-making process, and the effects of contractual non-completeness — can be acute in the supply of intermediate input services. Indeed, BHPIO's evidence in the proceedings engaged with FMG is suggestive of these problems, though it needs to be remembered that this evidence has not yet been tested in court.

There is, for example, evidence on the public record suggesting that the Mt Newman line and BHPBIO's production system more generally are characterised by significant complementarities and indivisibilities:²¹

- Interdependencies between the wheels of the rolling stock and the track profile point to the existence of cost complementarities between BHPBIO's above- and below-rail inputs. This is consistent with the results of empirical studies of freight railways more generally, where similar cost complementarities have been identified (see, for example, Ivaldi and McCullough 2001).
- Other cost complementarities arise at an aggregate level between different activities that form part of BHPBIO's integrated operations, such as mining, loading/unloading, railing and port operations. Complementarities between activities give rise to 'system effects', so that the whole is more than the sum of the parts
- Supply indivisibilities arise at all levels of the production system, so that investment can only take place in substantial, discrete increments. For instance, a train comprises a minimum number of locomotives and ore cars, and loading and unloading processes require substantial infrastructure units.

Given these conditions, it is not obvious how any decentralised price system can be designed that will coordinate efficient use as between the access seeker and the access provider.

21 The fact that there are interdependencies at the wheel–rail interface is apparent from discussion in the Ridley Affidavit, 4 October 2004, ¶4.3, which discusses wheel and rail interaction, as well as ¶6, which notes the related costs that would be incurred with third-party access. These interdependences are also noted in CRA, FMG Application to Access Mt Newman Railway Line under Part IIIA, 3 June 2005, ¶259–260, and reflected in ¶269–271. Cost complementarities arising at the aggregate level are reflected in discussion in the Price Affidavit, 23 December 2005, ¶4–10, which describes the role of the Integrated Planning Group, as well as ¶11–14, which describes the integrated nature of BHPBIO's operations; the Hoare Affidavit, 22 December 2005, ¶12–47, which describes BHPBIO's mine, rail and port operations; the Monkhouse Affidavit, 23 December, 2005, ¶8–12, which discusses the strategic importance of integration of main, rail and port operations; and the Ridley Affidavit, 4 October 2004, ¶2.5–2.9. These complementarities are also noted in BHPBIO's Submission to the National Competition Council, 3 June 2005, ¶60–70.

At the same time, BHPBIO has pointed to problems that, seen from an economic perspective, involve contractual incompleteness and suggest a risk that access would give rise to 'unpriced margins' that could be exploited by the access seeker to inefficiently impose costs.

Thus, issues related to quality monitoring arise at the wheel and rail interface, where careful management is central to improved operational efficiency — for example, as a result of reduced incidences of track defects and derailments, and longer-term efficiencies from reduced repair, maintenance and investment costs. Thus, BHPBIO may not be in a position to monitor whether or not a third party's rolling stock conforms to BHPBIO's wheel profile, or to monitor ongoing compliance over time. The inability to monitor key performance aspects of the third party's train operations extends to other operational aspects of the Mt Newman line, such as maintenance standards, train-loading standards and train-driver conduct.²²

Equally, in a day-to-day context, hold-up risks could occur because of the considerable external and component variability that characterises all aspects of the operations of the Mt Newman line.²³ Irrespective of its cause, flexibility in the operations of the Mt Newman line, and in particular the flexible sequencing of trains, is said by BHPBIO to be central to managing system variability.²⁴ As a result, a requirement to secure agreement from an access seeker to vary the transportation schedules would offer that access seeker ongoing leverage to extract 'rents' in return for agreement to such deviations.

Finally, further hold-up and hold-out risks could arise in areas where third parties are in a position to hinder investment initiatives that would deliver greater operational efficiency, reduce failures and delays and improve rail safety. Many research projects (such as changes to the wheel–rail interface, signalling or control systems, train management and scheduling) require the full cooperation of all third-party users in order to be effective.²⁵ To the extent that some such advances involve disruptions to, or reorganisation of, the rail

22 The careful monitoring and management required at the wheel–rail interface, as well as in other aspects of the BHPBIO's rail operations, can be inferred from discussion in the Ridley Affidavit, 4 October 2004.

23 Variability in mining operations arises from operational differences at the mines, such as limited stockpile space or longer load times. Factors that cause variability at port are ship arrivals, ship-loading constraints, stockpiling constraints, interactions between reclaiming and stockpiling equipment, or high demand for a specific product. The variability inherent in the BHPBIO system, and the implications for BHPBIO's operations, are noted in the Ridley Affidavit, 22 December 2005, ¶24 and ¶43–53. The impact of variability on BHPBIO's integrated mine, rail and port system are discussed in the Hoare Affidavit, 22 December 2005, ¶20–37.

24 The Price Affidavit, 6 September 2004, ¶5.1–5.40, describes the system used by BHPBIO to minimise and control variability. Section 6 of this affidavit notes the importance of maintaining flexibility in the rail component of the overall system. The importance of system flexibility is also noted in CRA, FMG Application to Access Mt Newman Railway Line under Part IIIA, 3 June 2005, ¶261.

25 That these types of research projects require the full cooperation of all third-party users in order to be effective can be inferred from discussion in the Ridley Affidavit, 4 October 2004.

system, they would require negotiations with, or even the consent of, third parties. Additionally, innovations such as the introduction of driverless trains could only be implemented if all users of the rail system incorporated the new technology.

The continuing need for central administrative control

Combined, these features limit the extent to which a pure liability rule, which centres on determining a Pigouvian charge that can support decentralised decision-making, could be made to efficiently 'fill the gap' (if there is one) relative to private bargaining. Rather, where supply-side non-convexities and contractual incompleteness limit the effectiveness of the price system, resource allocation will need to rely on some type of administrative mechanism that allocates capacity, authoritatively controls use of that capacity and determines investment decisions (given that those decisions have consequences for both parties). Whatever pricing mechanism there is, in other words, would have to be supplemented by extensive non-price direction and administrative control, much as happens within the vertically integrated firm.

However, the design of such a parallel administrative mechanism involves obvious problems of its own, including those of truthfully eliciting the information required and then enforcing whatever outcomes the mechanism results in. It also raises the difficult question of who should run this mechanism on a day-to-day basis. Clearly, the access provider seems by far the best-placed party to do so, both in terms of information and incentives (given that it has most at stake in the efficient management of the assets, which, after all, it owns). However, if the right to administer this process *is* vested in the access provider, then there must be the risk that the 'system administrator' will have incentives to discriminate against the access seeker, giving rise to further transactions costs as and when any conflicts occur and are dealt with, and to potential inefficiencies.

Dealing with these incentives for discrimination then brings us to another important finding of the modern literature on efficient regulation.

The standard analysis of the problem of vertical discrimination in access situations shows that whether vertical discrimination is profitable for firms depends on the specifics of the case (see Mandy and Sappington 2007, section 1, which provides citations to the literature):

- When the effect of the discrimination is to increase the costs of the (otherwise efficient) rivals to the vertically integrated firm, there is, at any given price, a reduction in output and part of the retail demand shifts toward the vertically

integrated firm (some of the retail demand may shift to a third party and some may be lost, as is the case if there is some product differentiation).

- However, this reduction in output also reduces demand for the vertically integrated firm's wholesale services, but in general by more than the gain in retail sales.

The net effect on the vertically integrated firm's profits, assuming wholesale service prices exceed short-run marginal costs (that is, make some contribution to total costs), is ambiguous, and so specific analysis is required for each given situation (though, Mandy and Sappington (2007) show that in telecommunications at least the likelihood that demand-reducing vertical discrimination is profitable is small).

It is also the case that the vertically integrated firm's revenue losses from engaging in vertical discrimination increase with the difference between the wholesale price and short-run incremental cost. Thus, the larger this difference, the less likely it is that a discriminatory action that causes a given increase in rivals' costs (which reduces wholesale demand) is profitable, and, hence, the less attractive is discriminatory behaviour (as the effect of that behaviour is to increase rivals' costs or reduce their revenues). If, however, the wholesale margin is negative, then there is no ambiguity — raising rivals' costs, unless that is a costly exercise, is profitable. In short, the higher access prices are set, the less likely it is that vertical discrimination will be profitable (as that discrimination causes the vertically integrated firm to forgo some profitable wholesale sales).

This, however, is only part of the story. In the present context, the regulator wishes to achieve two objectives:

1. set a cap over access prices that provides appropriate investment incentives without granting the firm unnecessary profits; and
2. prevent (to the extent that this is an issue) non-price vertical discrimination.

It is likely that efforts spent on preventing excess profits and non-price vertical discrimination both have positive, diminishing marginal benefits and at least constant, if not increasing, marginal costs. At the same time, by the proceeding analysis, there is a natural effort substitutability between the actions the regulator can take. For example, a larger gap between access price and cost reduces the profit gained by discriminating against downstream rivals, and so reduces discrimination. In this circumstance, the important result derived by Holmstrom and Milgrom (1991) shows that an efficient regulatory response is to reduce the harshness of the controls over monopoly pricing so as to obtain optimal incentives to avoid vertical discrimination. In other words, in the presence of costly information about conduct, the risk of discrimination can be reduced by reducing the severity of controls over monopoly pricing, as higher access charges themselves are an instrument for preventing discrimination.

Two further factors reinforce this effect. First, the need to ensure appropriate investment incentives complements the regulator's desire to reduce vertical discrimination, and this favours accepting more generous access-price regulation (such as a more lenient price cap). Second, many efficient actions may well appear discriminatory (for example, it is difficult to determine whether the vertically integrated firm faces lower costs because vertical integration is efficient, or because it discriminates against other firms). Distinguishing 'discriminatory' from 'efficient' responses to situations involving disputed use of shared facilities is likely to be even more difficult than securing accurate information about costs. As a result, there are likely to be high costs (including as a result of error) associated with enforcing non-discriminatory regulations, even in vertically separated firms (which may still discriminate through contracts and understandings).

It follows that the optimal stance of policy in these circumstances is to err on the side of allowing relatively high access prices (or equivalently, accept relatively lax access-price regulation), as such a stance:

- (a) In itself, is a cost-effective approach to reducing the risk of discrimination while preserving efficiencies of vertical integration; and
- (b) Is consistent with the need to provide credible incentives for investment and avoid or mitigate the risk of *ex post* expropriation of investors' sunk costs.

The other way of putting this is that because a perfect 'price only' system for allocating the resource cannot be devised, some form of administrative allocation must also be used. However, that administrative allocation inevitably involves some risk of non-price discrimination. That non-price discrimination has an unavoidable efficiency cost, some part of which is most efficiently taken in the form of a higher access price than would prevail were there no possibility of discrimination. Combined, the higher access price and the remaining risk of distortions due to discrimination then whittle away at the net gains, if any, from the liability rule.

Summary on 'regulatory failure'

In short, it is simply a mistake, from an economic perspective, to assume that the types of decisions that are taken administratively within the vertically integrated firm — in determining investment, managing asset availability and selecting efficient levels of output — can readily be solved in the liability rule by the device of a decentralised pricing process based on the access charge. While there may be instances in which that is possible, it will not be if the

production processes at issue are characterised by the non-convexities and imperfect contracting problems highlighted above. It seems reasonable to think that these features will be common in the sorts of situations that give rise to access disputes.

As a result, a liability rule, rather than allowing potential Pareto gains to be achieved, may well struggle to do no worse than private bargaining, even accepting that private bargaining will be imperfect. It may, in particular prove difficult to:

- Devise a price system for allocating the shared resource that is consistent with cost optimisation and efficient investment signals in the presence of imperfect information on the one hand, and indivisibilities and complementarities in asset structures on the other;
- Control the costs and distortions associated with ‘unpriced margins’ and opportunistic behaviour by the access seeker;
- Design an administrative system of resource allocation that both vests control of that system in the party best placed to control it, and avoids the high costs of possible discrimination over the non-price terms of access.

These difficulties, individually and together, give rise to real costs, which must be taken into account in considering whether social efficiency would be enhanced by replacing a property rule by a liability rule. There must, in other words, be a material likelihood that the ‘regulatory failure’ associated with mandated access will be more costly than the ‘bargaining’ or ‘market’ failure it is seeking to correct.

Conclusions

Seen in international perspective, Part IIIA is unusual. No other OECD country has a generic, economy-wide, provision of this kind, though many impose third-party access requirements in particular regulated industries, and a few (essentially the US and to a limited extent the EU) have some scope for the courts to mandate third-party access as a remedy under the competition laws.

Moreover, mandatory third-party access is an unusual, and in many respects especially invasive, form of remedy. In particular, while most liability rules provide for occasional and largely incidental incursions on a property right (Ellickson 1993), third-party access effectively converts the access seeker into what might be regarded as a joint tenant with the access provider. However, it does so in circumstances where, almost by definition, the parties have sharply

divergent interests, differing views of the world, and (given the scope for gaming associated with ongoing regulatory intervention) limited ability to make credible long-term commitments to one another.

As a result, where efficient use and investment requires many joint decisions to be taken, a substantial — and potentially impossible — burden of conflict resolution is likely to be placed on the regulatory apparatus, which must seek to discharge that burden in a situation where each party can have strong incentives to distort the information it provides.

Overall, the profit-maximising monopolist that engages in anti-competitive and inefficient ‘leveraging’ of an essential facility is likely to be a rare bird indeed, should such a bird exist. And high risks and costs may be incurred in seeking to catch such birds, especially if in the process one mistakenly labels as anti-competitive conduct that actually protects and promotes efficiency, as refusals to allow access may. Consequently, before imposing mandatory third party access, one would want to be convinced that:

- The potential access provider must have control over an essential input, and use that control to extract monopoly rents in the downstream market;
- The refusal to supply access must be related to the protection of those rents by means of the exclusion of no-less-efficient competitors, rather than by the seemingly more plausible desire to preserve the efficiencies of vertical integration; and
- It must be more efficient to regulate the monopolist by mandating the supply of access than by directly regulating the supply of the final goods that it monopolises or simply not regulating at all.

In the companion piece, I argue that Part IIIA fails to properly test for these factors and creates a substantial risk of ‘false positives’ — that is, of mandating access when it should be denied. To that extent, our current access arrangements seem more likely to reduce efficiency than to increase it.

References

- Alchian, A. 1977, ‘Some Implications of Recognition of Property Right Transactions Costs’, in K. Brunner (ed.), *Economics and Social Institutions: Insights from the Conferences on Analysis and Ideology*, Boston, MA: Martinus Nijhoff.
- Ayres, I. 2005, *Optional Law: the Structure of Legal Entitlements*, Chicago: University of Chicago Press.

- Barzel, Y. 1989, *Economic Analysis of Property Rights*, Cambridge: Cambridge University Press.
- Barzel, Y. 1982, 'Measurement Cost and the Organization of Markets', *Journal of Law and Economics* XXV(1): 27–48.
- Calabresi, G. and Melamed, A. Douglas 1972, 'Property Rules, Liability Rules, and Inalienability: One View of the Cathedral', *Harvard Law Review* 85(6): 1089–1128. Reprinted in R. C. Ellickson, C. M. Rose and B. Ackerman 1995, *Perspectives on Property Law*, 2nd ed., Boston, MA: Little Brown.
- Coase, R. H. 1937, 'The Nature of the Firm', *Economica* 4(16): 386–405.
- Coase, R. H. 1960, 'The Problem of Social Cost', *The Journal of Law & Economics* 3(1): 1–44.
- Coase, R. H. 1988a, *The Firm, the Market, and the Law*, Chicago, IL: University of Chicago Press.
- Dahlman, C. J. 1979, 'The Problem of Externality', *Journal of Law and Economics* 22(1): 141–162.
- Demsetz, H. 1967, 'Toward a Theory of Property Rights', *American Economic Review Papers and Proceedings* 57(2): 347–59.
- Demsetz, H. 1997, 'Agency and Nonagency Explanations of the Firm's Organization' in *The Economics of the Business Firm: Seven Critical Commentaries*, Cambridge: Cambridge University Press: 15–39.
- Ellickson, R. C. 1993, 'Property in Land', *Yale Law Journal* 102(6): 1315–1401.
- Ergas, H. 2008, *Wrong Number: Resolving Australia's Telecommunications Impasse*, Sydney: Allen & Unwin.
- Holmstrom, B. and Milgrom, P. R. 1991, 'Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design', *Journal of Law, Economics, and Organization* 7 (special issue): 24–52.
- Gasmi, F., Kennet, M., Laffont, J.-J., and Sharkey, W. 2002, *Cost Proxy Models and Telecommunications Policy: A New Empirical Approach to Regulation*, Cambridge, Mass: The MIT Press.
- Heal, G. M. 1973, *The Theory of Economic Planning*, Amsterdam: North-Holland.
- Heller, M. 2008, *The Gridlocked Economy*, New York: Basic Books.
- Hillier, B. 1997, *The Economics of Asymmetric Information*, New York: St Martin's Press.

- Independent Committee of Inquiry into Competition Policy in Australia 1993, National Competition Policy ('Hilmer Report'), Australian Government Publishing Service, Canberra.
- Ivaldi, M. and McCullough, G. J. 2001, 'Density and Integration Effects on Class 1 U.S. Freight Railroads', *Journal of Regulatory Economics* 19(2): 161–82.
- Lafontaine, F. and Slade, M. 2007, 'Vertical Integration and Firm Boundaries: the Evidence', *Journal of Economics Literature* 45(3): 629–85.
- Mandy, D. and Sappington, D. E. M. 2007, 'Incentives for Sabotage in Vertically Related Industries', *Journal of Regulatory Economics* 31(3): 235–60.
- Parish, R. 1980 'Methods of Non-Price Rationing', mimeo, Monash University, Melbourne.
- Perry, M. 1980, 'Forward Integration by Alcoa: 1888–1930', *The Journal of Industrial Economics* 29(1): 37–53.
- Pirrong, S. C. 1993, 'Contracting Practices in Bulk Shipping Markets: A Transactions Cost Explanation', *Journal of Law and Economics* 36(2): 937–76.
- Quinet, E. 2003, 'Short Term Adjustments in Rail Activity: The Limited Role of Infrastructure Charges', *Transport Policy* 10(1): 73–9.
- Roberts, J. 2004, *The Modern Firm Organizational Design for Performance and Growth*, New York: Oxford University Press.
- Roth A. E. 1985, *Game-Theoretic Models of Bargaining*, Cambridge: Cambridge University Press.
- Schelling, T. C. 1983, 'Prices as Regulatory Instruments', in Thomas C. Schelling (ed.), *Incentives for Environmental Protection*, Cambridge, MA: MIT Press.
- Stigler, G. J. 1966, *The Theory of Price*, New York: Macmillan.
- Sunder, S. 1997, *Theory of Accounting and Control*, Cincinnati, Ohio: South-Western College Publishing.
- Werin, L. 2003, *Economic Behavior & Legal Institutions: An Introductory Survey*, River Edge, NJ: World Scientific Publishing.
- Wilson, R. B. 1998, 'Negotiation with Private Information: Litigation and Strikes', in Donald P. Jacobs, Ehud Kalai and Morton I. Kamien (eds), *Frontiers of Research in Economic Theory: The Nancy L. Schwartz Memorial Lectures, 1983–1997*, Cambridge: Cambridge University Press.

Applications and submissions

CRA, FMG Application to Access Mt Newman Railway Line under Part IIIA, 3 June 2005.

BHPBIO, Submission to the National Competition Council, 3 June 2005.

BHPBIO, Modelling and Related Submissions to the National Competition Council in Response on Draft Recommendation, 4 November 2005, submitted 6 January 2006. *Affidavits filed with the NCC and available at www.ncc.gov.au*

Price, Phillip Anthony, Affidavit, 23 December 2005.

Hoare, Rodney Tytler, Affidavit, 22 December 2005.

Monkhouse, Peter Herbert Lloyd, Affidavit, 23 December, 2005.

Ridley, Robert Ian, Affidavit, 4 October 2004.