
The ‘baby and the bath water’: avoiding efficiency mishaps in regulating monopoly infrastructure.*

Gary Banks

Chairman, Productivity Commission

Over the past decade and a half, Governments have radically transformed regulatory, governance and ownership arrangements for essential infrastructure — such as telecommunications, water, energy and transport. Statutory monopolies have largely been swept away to be replaced by competition in infrastructure services. This has facilitated greater innovation and brought more efficient pricing of essential infrastructure. The gains to the community are big, as recent Productivity Commission reports have shown. This in turn reflects the inefficiency of previous arrangements and the fact that such infrastructure accounts for about one fifth of Australia’s total capital stock and plays such a pivotal role in producing services to business and directly to consumers.

This competitive transformation has involved the construction of an elaborate regulatory apparatus. Left unchecked, incumbents operating in previously legislated monopolies, or firms in markets characterised by natural monopoly, may wield enduring market power as a result of the large and (usually) irreversible investments — sunk costs — required by entrants. In seeking to maximise their profits, they may strive not only to be more efficient in their operations, but also to set prices above costs.

National Competition Policy sought to address this infrastructure problem by:

- enhancing competitive disciplines on government business enterprises through the application of competitive neutrality principles and structural reform of public monopolies;

* Presentation to the IPART Conference, *Incentive Regulation at the Crossroads*, Sydney, 5 July 2002. The presentation draws on the policy discussion in the Commission’s latest annual report and inquiry reports into the national access regime, the Prices Surveillance Act, airports, telecommunications and harbour towage services (visit www.pc.gov.au).

-
- establishing in each jurisdiction arrangements to oversee prices charged by utilities and other corporations with substantial monopoly power; and
 - establishing rules to enable potential competitors to gain access to the services of significant monopoly infrastructure.

There are also specific arrangements for sectors such as telecommunications, where the initial lack of competition and the fast pace of change meant that the Government was reluctant to rely on general trade practices law.

Given the ‘experimental’ nature of some of these reforms, and the potential for them to generate costs as well as benefits, it was envisaged at the outset that there would need to be independent reviews of the new regulatory arrangements after a few years. A recent suite of Productivity Commission inquiries — encompassing reviews of the National Access Regime, telecommunications competition regulation, airport services pricing, rail reform, harbour towage services and the Prices Surveillance Act — has provided the opportunity for an independent and public stocktake of key elements of the pro-competition regulation governing Australia’s economic infrastructure.

When I accepted IPART’s invitation to discuss the Commission’s views in these areas, I was confident that our reports on the National Access Regime and the Prices Surveillance Act would have been released. Unfortunately that is not the case, which means that I am obliged to stay at a higher level of generality than otherwise.

The Commission has generally found that there are legitimate grounds for maintaining regulatory oversight of some form in the areas it analysed. However the current regulatory framework is not free from flaws, risks and some important unresolved issues.

These issues, captured by the familiar ‘baby’s bathwater’ analogy in the title, constitute the main focus of my talk today. Those who have read some of the Commission’s reports, including our most recent Annual Report, will recognise the themes. However, they bear repeating, especially at a conference with ‘crossroads’ in the title, because in the Commission’s view how regulatory structures develop from this point on is of considerable importance to Australia’s future.

Competition is not an end in itself

National Competition Policy and related policies were based on an understanding by all governments that, by and large, competition leads to greater productivity, stronger incentives for innovation, lower costs and improved service, and so eventually to higher incomes. It is broadly understood by regulators — as

emphasised by Hilmer back in 1993 — that competition itself is not the goal of competition regulation, but a means of achieving higher standards of living for Australians.

Even so, the notion of exposing former monopolies to competitive disciplines is so alluring, the political rhetoric so compelling and the statutes themselves so structured, that at times regulators may understandably give undue emphasis to achieving competition and push the fundamental objective of efficiency into the background. In the context of infrastructure regulation, this poses several risks for regulators and to adjudicators of competition regimes.

‘Perfect’ competition would be costly

First, and most importantly, the trigger for regulatory action must not just be a departure from some competitive ideal. Entry barriers and market power lie on a continuum, with some market structures and outcomes closer to those of perfect competition than others, but none attaining that theoretical abstraction. The costs of a particular unconstrained market outcome depend on what alternative is realistically achievable. Given the compliance, administrative and other more significant costs of regulation, there is limited scope for beneficial policy interventions in markets that remain ‘workably’ competitive.

The more sophisticated regulators (such as those here present!) no doubt accept this. However, tolerance for imperfection in competition seems to be rather lower for infrastructure than for other sectors of the economy, despite the fact that the costs of regulatory error may well be higher — a point I’ll get back to.

In its telecommunications inquiry, the Commission recommended that one way of reducing the risk of interventions in workably competitive markets was to set a stricter threshold for regulatory action. We proposed that a necessary hurdle for declaring a service was an expectation that this would promote a *substantial* increase in competition, not just any increase.

One example of where this would have made a difference is the recent declaration of CDMA mobile services by the ACCC. It appears that a major rationale for declaring CDMA was to give equal regulatory treatment to two functionally close technologies. However, the likelihood is that any increment to competition from declaring CDMA will be relatively slight at best.

Another dimension of ‘workable’ competition is how quickly any market power is likely to be eroded. Market power may be high, but short-lived, as new technologies compete with the old or as services converge. In telecommunications services, wireless local loops, new fibre optic networks and additional satellite services are

increasingly threatening the dominance of incumbents reliant on copper lines. Competition can spring from surprising sources. When Alexander Bell patented the telephone in 1876 it was dismissed as a fleeting novelty. Western Union, the largest telegraph service of the time, decided not to buy the patent, and the rest, as they say, is history. In some countries, internet and telecommunications services are now being provided through the electricity system — a new and unexpected source of competition to the conventional conduits for such services.

The point to emphasise is that the prospect of market power is what motivates firms to innovate and new firms to enter markets. Such transitory market power is not inimical to competition. Rather, it invites it. The patent system recognises this by statutorily protecting intellectual property from competitors — and the TPA exempts access to these forms of property. However, the Act does not cover other aspects of innovation. A danger is that the pursuit of static competitive outcomes might choke the incentives for innovation. This is particularly relevant to in those areas of infrastructure where technologies are evolving quickly, such as telecommunications.

The collective impact of substitution possibilities

Second, the availability of substitutes limits the exercise of market power. For example, rail track providers may be sole suppliers but, for most freight or other services, they face intense competition from other transport modes, notably road.

A key issue in determining whether competition is adequate is not to get pre-occupied with technological descriptions of markets or indeed in mechanically defining markets at all, but on testing the extent of competitive pressure on firms arising from the collective impact of a whole set of substitution possibilities.

An example that illustrates the drawbacks of a technologically-oriented approach to market definition is given by the declaration of analogue subscription pay TV services back in 1999. Among other reasons given for this decision, it was judged that videos were not a substitute for pay TV because they are less convenient to acquire, and that free-to-air TV was also not a substitute because it is paid for by advertisers rather than by subscription. But the fact that the form in which these services are provided differs from pay TV need not disqualify them as substitutes. Even if, individually, services are not *close* substitutes, collectively they may exert enough discipline to remove any significant scope for excess profits.

Incomplete incentives to exploit market power

Third, there will not always be an incentive or the scope to exploit market power, even where it appears to exist. Taking airports as an example, with more than four times as much operating profit earned from non-aeronautical activities such as retailing and car parking, the owners of the major airports have clear incentives to moderate their pricing of aeronautical services so as not to reduce passenger throughput and undermine total airport profitability.

This has been recognised by the Government in recently accepting the Productivity Commission's recommendations to adopt a more light handed approach to airport regulation.

Countervailing power

Fourth, any countervailing power of major users can also be a constraint on monopoly behaviour. Taking the example of airports again, attempts to exercise market power can expect to be resisted by airlines, who have some commercial clout of their own (especially in dealing with smaller airports reliant on holiday markets). The concentration of the industry since Ansett's demise will have strengthened the scope for exercising such countervailing power.

Avoiding a numbers game

Undue emphasis on competition may encourage a 'numbers game' in which competition is not measured primarily by entry barriers but by how many actual competitors there are. This risks:

- potentially inefficient entry in certain technologies; and
- the 'double marginalisation' problem, whereby multiple firms, each with monopoly power in a local market, charge even more inefficient prices to each other than would the arms of an integrated monopolist.

In short, regulators have a tough task in gauging and responding to potential market power. The complexities facing them cautions against assessing competition and market power in a static context, or one which doesn't account for market reactions to that power being exercised.

Monopoly power may be used efficiently

Even where monopoly power is exercised, it may not have significant negative impacts on efficiency. In particular, to the extent that monopolists can structure

their price menus efficiently, so that prices are high for the inelastic segment of demand and low for the elastic segment, there may be little distortion in supply or consumption patterns.

For example, in the case of airports, there are numerous examples of airport price structures designed to promote or retain marginal users, including direct incentives designed to encourage additional flights and new entrant airlines.

Of course, there may be *distributional* consequences, but whether these warrant concern is not always straightforward. For example, the losers from higher than necessary airport charges would potentially be passengers paying higher fares and airline shareholders earning lower returns. But the diversity of share ownership in airlines and airports — directly or indirectly through large superannuation funds — and the mix of foreigners and Australian residents amongst shareholders and passengers, mean that any distributional effects may be largely ‘neutralised’.

Ironically, a possible victim of the regulatory response to market power has been to limit the scope for the very feature that reduced the adverse efficiency effects of that power in an unregulated setting — multi-part pricing at the access and retail level. Regulated access prices have generally been uniform and cost-based. Where services use common fixed costs — which is a ubiquitous feature of infrastructure services — the regulator is forced to use arbitrary cost allocation rules, instead of seeking to recover a greater portion of common fixed costs from inelastic demand.

Investment matters too

Access and price regulation have the potential to improve efficiency where natural monopoly is a problem and/or markets are in transition. However, the regulatory challenge is to ensure that prices are set neither too high nor too low. There are dangers both ways. Given the legacy of government ownership and control of vertically integrated monopolies, it is not surprising that much of the initial regulatory focus has been on reducing prices. This has been to the direct benefit of consumers and using industries and has led to market innovations and expanded choice.

However, the major risk associated with the regulation of essential infrastructure is that setting prices too low could deter new investment in the facilities themselves. At a conceptual level it is clear that access and price regulation involve a significant intrusion into the property rights of facility owners and can distort their investment behaviour. While available evidence of adverse impacts on past investment is largely anecdotal and difficult to verify, the potential risks of adverse consequences

from regulatory action appear to be looming larger. Some of these are documented in the Commission's final report on the National Access Regime.

There is a potential tension between the efficient use of existing facilities and incentives to build new ones. Once investments have been made, the actual costs of running transmission and distribution networks are relatively low. With capital effectively having no alternative uses, there is a theoretical case — at least from a short-term perspective — in setting prices to recover only marginal or operating costs. However, this would deny the firm the opportunity to recover its fixed costs. While the service might continue to be provided for the asset's economic life (though possibly with inadequate maintenance), such a pricing policy would destroy incentives for any replacement investment.

The 'truncation problem'

No firm, including existing facility owners, will commit to major new capital outlays without the expectation of profits commensurate with the commercial risks involved. Realised returns can be affected by unforeseen delays and costs during the construction phase, unanticipated changes in market demand, uncertainty about how an untried technology will perform or the possible emergence of a superior competing technology. So even without regulatory risk, profitability cannot be assured. For investments which are particularly risky, or that have the expectation of only normal returns allowing for such risk, the potential for regulatory action to deter or even stop new investment is very real.

Regulators may sometimes unwittingly appropriate what appear to be excess returns, but which are in fact the necessary upside of a risky investment. By contrast, regulators cannot compensate firms for any downsides. With an investment that is already in place, a regulator can engage in regulatory taking without threatening the existing service. But future investment incentives may be undermined. Indeed, an important function of a regulator is to provide signals to guide future investment.

An example from Melbourne University's Stephen King illustrates the problem. A cable TV provider is considering investing in a regional town. The investment costs \$51 million and, in the absence of access provisions, returns \$100 million if pay TV is very successful, \$60 million if it is moderately successful and only \$20 million if unsuccessful. Say the likelihood of each of these outcomes is 25 per cent, 50 per cent and 25 per cent respectively. In that case, the expected net return is \$9 million. Now consider the situation whereby the facility is not immune from access and that the regulator would reign in any apparent excess profits by granting access to the facilities to rivals. The maximum potential return drops from \$100 million to \$60

million. The regulator sees this as a benign outcome because the provider would still get \$9 million if the venture were very successful. But because the regulator does not subsidise the less successful outcomes, the expected *ex ante* return from investing is now negative. By ‘truncating’ the potential returns from risky investments, an apparently benign regulatory policy can actually kill the incentive to invest in the first place.

In the case of access regulation, there may also be perverse incentives for investors to build smaller facilities than would be socially desirable, so as to ensure that there is little spare capacity beyond their immediate requirements, thereby removing any threat that they would be required to grant access at prices they considered too low.

Regulatory-induced service failure

Another disadvantage of unduly low regulated prices is that the investment required to maintain, extend or replace existing infrastructure may be delayed. This can result in a deterioration in service through breakdowns, increasing congestion and, depending where price restraints are imposed, profit squeezes on intermediary suppliers. The effects of this may go unnoticed for some years, until a crisis point is reached.

The California energy crisis in 2000 provides an apt illustration. For some time, retail prices were low, reflecting retail price caps imposed by the regulator. However, in the summer of 2000, wholesale energy prices rose steeply as electricity generating capacity failed to keep up with soaring demand and higher energy prices. The supply problems mainly reflected the fact that no new generating capacity was built in the 1990s and few transmission lines were constructed. This stemmed from environmental requirements and an uncertain regulatory environment. The retail price caps meant that the wholesalers were unable to pass on their higher costs to customers, so that they made huge losses. This undermined the confidence of electricity generators to supply wholesalers when they might not get paid, which was a major factor leading to rolling black outs. Ultimately, the retail caps were not sustainable, and prices for consumers rose by 40 per cent in April 2001 and some by 80 per cent. The knock on effects to the broader Californian economy have been severe.

The Californian electricity crisis is sometimes attributed to deregulation. In fact, what it demonstrates is that the regulations that persist after breaking up monopolies have to be carefully designed and adapted as markets change. Otherwise consumers can be made worse off in the long run.

To motivate adequate investment, prices need to be *at least* sufficient to cover the long-run costs of facility operators, including an adequate return for the risk involved.

But prices should not be set so far above costs as to detract from the efficient use of services or to inhibit investment and innovation in related markets. This is a particular risk in markets such as new value-adding telecommunication services, where investors are already grappling with rapid and unpredictable technological change and demand for the platform service is sensitive to price. There is also the possibility that high regulated prices could lead to the inefficient duplication of facilities where users have no option but to build their own.

It follows that there needs to be a balance between the short-term gains for users and consumers in having low prices, and the long-term interests of those same users and consumers, which requires the efficient timing and scale of investment. The Commission considers that regulatory frameworks need to provide clearer signals about how this balance is best achieved.

Regulatory measures must be ‘workable’

Notwithstanding the risks and complexities, the rationale for regulation *is* strong in many infrastructure areas. However, the choice of regulatory instrument and the processes used for declaring, arbitrating and otherwise managing a regulated regime then come to the fore. These can make or break the regulatory regime.

For example, in its recent draft report on harbour towage services, the Commission found that the instrument used was ineffective. The requirement for price notification under the Prices Surveillance Act had no impact on the (modest) market power present in those services. Probably the most effective constraint on market power is the potential for entry — if, as they say, planes are capital with wings, then tugs are mobile floating capital. From a policy perspective, allowing port authorities the discretion to license towage operators through competitive tender ‘for the market’ was found to be a superior option for lowering prices while maintaining quality than price regulation.

For a given approach, the *details* of regulatory provisions can matter a great deal. For example, lack of clarity about price capping arrangements for airports promoted strategic behaviour by all parties. This led to increased compliance burdens and discouraged commercial negotiations.

And of course, the speed and ease with which the regulatory regime works can be critical. Access regulations have often proven to be cumbersome and slow. Even the

Telecommunications Access regime, which was designed to be speedy to suit the pace of technological change in that sector, has taken years in some cases to reach preliminary determinations. For example, for the fundamental telecommunications access service (originating and terminating PSTN access services), the delay between the first notification of a dispute and the finalisation of the appeal is expected to be around five years. Changes have since been made to speed up the process.

A factor shaping these delays is the breadth of the regulatory regime. The longest lasting disputes in the telecommunications access regime have tended to be for services that are workably competitive and involve parties without much market power. The median delay to first resolution of a dispute involving Telstra as an access provider was around 300 days; it was 550 days for a dispute that did not involve Telstra at all. This is testimony to the wise decision by the ACCC to prioritise the cases that really matter, but it is also symptomatic of the willingness of parties to use the regulatory regime to press for commercial advantage in negotiations that should really be outside the scope of that regime. (Some of the disputes have been trivial, with one party claiming that in one dispute the ACCC arbitrated over a total sum in dispute between two carriers of \$12 per month.)

If nothing else, these points suggest that the more heavy-handed price or access regulations can involve some steep transaction costs, while not always being effective. Regulatory forbearance might avoid some of these. Or, where regulation is required, lighter-handed alternatives may have fewer downsides. For example, the prices monitoring regime proposed by the Commission for major airports, and effectively accepted by Government, provides some deterrence for abuse of market power, while allowing a lot of commercial latitude and sustained incentives for investment. Similarly, the ACCC has appropriately not proposed stipulating prices for mobile telephone services.

Regulators, too, are only human

The expectations placed on the competition regulator are high. The regulator is required to find a reasonably implementable system for encouraging efficient competition; to avoid appropriating the returns that motivate investment, and prevent tangling everyone in a mire of complex directives and procedures in the process. Regulators must simultaneously hold at bay game-playing participants and meet public requirements for fairness and transparency.

The tasks regulators must perform require considerable information if they are to improve consistently on market outcomes. That information must largely come from incumbents, who are naturally reluctant to lay all the relevant information in

the hands of regulators. When technology and market demands are changing rapidly — which is increasingly the case — the informational difficulties for regulators multiply.

Persistent methodological quandaries over such delphic issues as tilted or non-tilted depreciation, provisioning, common trenching costs, and fierce battles of the accounting acronyms — the DACs versus the DORCs — suggest we still face big uncertainties in some key areas that define the relevant asset bases and valuation.

Ultimately, a large element of judgment is unavoidable in deciding whether and how to intervene, and this raises the prospect of regulatory error.

The inevitability of error has special risks where long-lived investments in essential infrastructure are involved. The Californian energy crisis sent a chill through regulators around the world. Here was a situation in which a modern sophisticated regulatory authority had presided over, at least in hindsight, what appeared to be grave regulatory mistakes. This is sobering stuff and hopefully has done some good for the future.

Statutory guidance is fundamental

Regulatory discretion cannot be eliminated, and indeed, some discretion is desirable. However, to reduce the risk of regulatory error, statutes need to be clear about three things:

- the objectives of regulation;
- the behaviour at which intervention should be targeted, and
- the principles governing the type of intervention.

These basic requirements are not often met. For example, the Productivity Commission found that the National Access Regime was deficient in all three respects. It proposed the inclusion of an objects clause and pricing principles in the national regime and a change to the regime's 'declaration' criteria to reduce the possibility that services will be subjected to access arrangements without the prospect of a significant economic payoff.

Given the manifold uncertainties and information difficulties, there are limits to what regulators can be expected to achieve. Rather than aiming for an ideal but unattainable outcome, the public policy goal should *be a set of regulatory arrangements that will improve efficiency through time, while minimising the scope for regulatory errors*. A framework is needed in which regulators are encouraged to

intervene only when significant improvements in efficiency are in prospect and not to be too ambitious in fine tuning the prices they regulate.

A number of the Productivity Commission's general regulatory proposals, while they mainly apply to existing assets, would also help facilitate efficient investment in essential infrastructure services. However, the Commission has come to the view that special additional provisions will also be needed if new investment is not to suffer from an inherent regulatory tendency to truncate the up-side potential of a proposed investment, while allowing investors to bear all the downside risks — illustrated by recent references by regulators to average rates of return earned by companies on the stock exchange, as if these were an appropriate benchmark for prospective investment.

The problem of regulatory truncation is an important policy issue, but determining the best approach to dealing with it is not at all straightforward. It should nevertheless become a priority for government consideration and the Commission's final report on the National Access Regime provides guidance on how governments might go about it. These and other detailed proposals in that report are also highly relevant to the current COAG Energy Review, and the foreshadowed review of the Gas Access Code.

The need for balance

In conclusion, the task of regulating monopoly infrastructure is complex and faces major informational obstacles. Much has been achieved over the past decade, however, and the community has realised substantial gains. On the basis of the Commission's detailed reviews of particular areas of infrastructure regulation, further gains now hinge on making adjustments to the regulatory apparatus to minimise the risks for investment and to ensure that competition can discharge its rightful function of achieving more efficient outcomes for the long-term benefit of the community.