

Error and Design: Economics in (and some Economics of) the Australian Competition Tribunal

Henry Ergas¹

Abstract

The Australian Competition Tribunal plays a central role in the Australian regulatory system. Three of the Tribunal's recent decisions are examined and found to have significant errors. It is argued on the basis of the economics of the institutional design that these errors are inevitable in a system where regulatory decisions are based on vague standards. It is concluded that the quality of regulation would be improved by a shift to a regulatory system that relied less on such vague standards and more on prescriptive, explicit rules.

Introduction

How should regulatory institutions be structured? Contemporary models of the regulatory process view this issue primarily through the prism of principal–agent models, and focus on the trade-off between, on the one hand, securing scale economies in the exercise of regulatory functions, including through specialisation in the human capital they require, and, on the other, the risk of regulatory capture. Simply put, delegating functions to specialised bodies improves the effectiveness with which regulatory tasks are undertaken, but gives rise to agency costs, most notably the risk that the specialised regulator will pursue its own interests at the expense of those of the community on whose behalf it is intended to regulate (see, for example, Dal Bó 2006).

Agency costs can be viewed as reflecting differences in preferences over outcomes between a principal and an agent in the presence of monitoring costs. However, delegation also vests in the agent control over the choice of instrument, and hence relies on the agent's ability to map strategies into outcomes — to analyse the situation and choose the policy that stands the best prospects of securing the desired outcome. There is, in other words, not only a risk of bias (that the regulator selects on the basis of its own preferences, rather than the community's), but also a risk of error: that even in the absence of any misalignment of preferences over outcomes, the regulator will make mistakes in

¹ Concept Economics, HenryErgas@concepteconomics.com.au. I am grateful to the Editor, William Coleman, two referees, and to my colleagues Eric Ralph and Garth Crawford for very helpful comments. Responsibility for any remaining errors is, of course, my own.

the choice of policy. (A careful discussion of the difference between preferences over outcomes and preferences over policies can be found in Ganghof 2006).

This risk of error tends to be ignored in the formal models of regulation. This is not because the literature assumes the regulator is fully informed: on the contrary, models of optimal regulation typically start from the premise that the regulator does not have perfect access to the information it ideally requires. But it is then assumed that the regulator displays 'hyper-rationality' in responding to this situation.

It is, for example, usually assumed that the regulator knows, or can readily acquire, the 'technology' of regulation, a central component of which is the theory of optimal regulation itself. The regulator's problem, say in setting regulated prices, then becomes that of implementing that theory, given its own incentives and the constraints it faces in gathering the information required to determine the optimal prices (or the other parameters of the regulated firm's behaviour it is seeking to influence). The 'technology' recognises that the regulatory situation will be characterised by asymmetric information, in which key parameters of the model for determining optimal prices depend on information that is available to the regulated firm but not to the regulator. But while it accepts that the regulator does not know the exact values of those parameters, the regulator is assumed to know the underlying model and the joint distribution of those parameters, and to be capable of deriving, given that knowledge, the optimal policy response (see, for example, Sharkey *et al.* 2002).

In that sense, the standard theory deals with a recognition of ignorance by assuming even greater knowledge and computational capacity than would be required were the regulator fully informed.²

That assumption is, of course, plainly unrealistic. In practice, regulatory systems recognise that fact, and usually provide 'error correction' mechanisms, notably through independent review of regulatory decisions. Those review mechanisms serve a dual purpose: they can monitor against pure agency costs, in which the regulator's preferences replace those of the community; and they can also check for error by assessing the regulator's choice of policy in the light of community preferences over outcomes.

But how effective are the review mechanisms? To what extent can they avoid or reduce the risk of error? And what does that imply for the design of regulatory institutions?

² There are obvious similarities here to the Wilson problem in mechanism design theory, which points to the myriad common-knowledge assumptions made in deriving an optimal solution to the planner's problem of maximising a utility function subject to incentive constraints and highlights the restrictive (and implausible) nature of those assumptions (Wilson 1987). The common-knowledge assumptions made are often as strong as would be the assumption of complete information.

To examine these issues, I analyse three decisions of the Australian Competition Tribunal (ACT). The ACT plays an important role in the Australian regulatory system, as it hears applications for review of a wide range of decisions under the *Trade Practices Act, 1974* (TPA). Given its responsibilities, many of the Tribunal's decisions involve complex economic reasoning. So as to provide the Tribunal with the expertise needed to take these decisions, the Tribunal, when it hears a matter, is composed of three members: a presidential member, who must be a judge of the Federal Court, and two other members, who must have knowledge of or experience in industry, commerce, economics, law or public administration. Typically, one of the two non-presidential members would be an economist. The Tribunal's decisions should therefore provide a good benchmark for analysing the quality of the decisions taken by review mechanisms.

The conclusion I come to is that each of the decisions I analyse is flawed — indeed, involves what I believe could be fairly portrayed as errors in the underlying economics. The question then is what the implications are for regulatory design.

The second part of the article, in which I sketch a theory of the design of regulatory institutions, turns to that question. I suggest that errors such as those I have analysed are to be expected when generalist bodies must review decisions that span many areas of expertise, including in economics. But such review is inevitable and desirable if regulators (such as the Australian Competition and Consumer Commission (ACCC)) are given wide discretions, as those discretions increase both the risk of error by the regulatory bodies themselves and of rent-seeking. It is therefore sensible, under those circumstances, for the decisions taken by regulators to be subject to review by institutions that are less vulnerable to capture and other agency costs. Entities such as the Tribunal will be shown to have several advantages in this respect, but those advantages are bought at the cost of a lack of specialist expertise and a consequent risk of error.

In the third part of the article I consequently suggest an alternative strategy for reducing the risk both of rent-seeking and of error. That strategy would focus on reducing the scope of the discretion granted regulatory bodies. In particular, up-front rules, rather than vague standards, would be used to constrain regulatory discretion. This, it will be shown, is the approach that is now being adopted in regulating our energy industries, but not in other regulated sectors. While this approach has costs of its own, I conclude by recommending its extension to the regulated industries as a whole.

Review in action: three Tribunal decisions

In this section, I will examine three Tribunal decisions. I have selected these decisions on two bases.

The first is that they cover conceptual issues that are crucial to the ‘technology’ of optimal regulation: respectively, the determination of the cost of capital (obviously a factor of enormous significance in industries that are very capital intensive), the valuation of sunk assets (also of obvious significance, given that at any one time the assets of regulated firms are largely sunk) and the assessment of whether particular price structures are efficient. The second criterion is that the economic reasoning that will be discussed played an important role in the decision. I will suggest that in each decision that reasoning was flawed in crucial respects.

The three decisions relate respectively to the setting of the allowed charges for Telstra’s supply of the unbundled local loop service, the determination of the valuation of easements used by ElectraNet Pty Limited, and the extension of mandatory third-party access to certain facilities at Sydney Airport. I should state at the outset that I played a role in two of those decisions, namely that relating to Telstra’s unbundled local loop service and that relating to Sydney Airport. However, my part in the first was very limited, and did not touch on the aspect of the decision considered here. As regards the second, I provided evidence on behalf of Qantas, which was supporting the application for third-party access to be granted. This is what the Tribunal recommended, but my criticism goes to the reasoning it used, reasoning quite unrelated to the topic and substance of my evidence.

Rates of return and investment

The first decision regards an appeal by Telstra to the Tribunal in respect of the ACCC’s rejection of an Undertaking Telstra had offered covering the pricing of the Unconditioned Local Loop Service (ULLS).³ Simply put, under Part XIC of the Trade Practices Act access providers, such as Telstra, can propose Undertakings that set out the terms and conditions on which they will provide access to the services they are required to supply to access seekers. If accepted by the ACCC, those terms and conditions become binding on the ACCC should it be required to arbitrate an access dispute. In December 2005, Telstra submitted such an Undertaking for the ULLS, which is the service of providing the copper pair to an access seeker, who can then use that copper pair to provide services such as telephony and ADSL.

Telstra’s proposed Undertaking was rejected by the ACCC on the grounds that it was not satisfied the charges that Telstra proposed were reasonable. One element in the ACCC’s concerns involved the Weighted Average Cost of Capital (WACC) Telstra had relied on in calculating the cost of the ULLS. Telstra noted that determining the appropriate WACC was complex, but argued that in setting

³ Australian Competition Tribunal, *Telstra Corporation Ltd (No 3)* [2007] ACompT 3.

the WACC, it was important to take account of the fact that setting the allowed rate of return too low would deter investment.

The ACCC considered these issues within the framework of the conventional Sharpe-Lintner Capital Asset Pricing Model (CAPM) (Sharpe 1964; Linter 1965), as had Telstra, and disagreed with Telstra's contention. It was on this basis that the issue (along with a range of other points of disagreement) was appealed to the Tribunal. The Tribunal, relying on the CAPM, examined Telstra's arguments relevant to the WACC but found against Telstra, stating that:

Telstra assumed that setting a WACC that was too low would deter investors. However, different investors will inevitably have different attitudes to risk. Setting the WACC below the true value may deter some investors and therefore result in less investment taking place in the short run, but it will not be likely to cause all investors to cease providing funds.⁴

The obvious problem with this argument is that it is inconsistent with the Sharpe-Lintner CAPM on which the Tribunal relies. The Sharpe-Lintner CAPM allows investors to have different attitudes to risk. However, contrary to the Tribunal's suggestion, the Sharpe-Lintner CAPM shows that in spite of any heterogeneity in attitudes to risk, all investors will require the same risk premium in equilibrium — if expected returns fall below that level, the investment at issue will not be funded.

Of course, it is also true that simple versions of the CAPM assume investors have homogenous *expectations*. However, this is not the assumption the Tribunal refers to, as its concern is that 'investors will inevitably have different *attitudes* to risk' (my italics), and the Tribunal shows no sign of being at all aware of these subtleties in respect of the CAPM. Moreover, even had the Tribunal instead referred to heterogeneity of expectations, this would not have helped it reach the conclusion it reached. For example, under plausible assumptions, allowing for heterogeneous expectations leads to outcomes in which investors require a higher equilibrium return, as they need to be 'insured' against the risk of being over-optimistic (as in the 'winner's curse') (Harrison & Kreps 1978). As a result, a WACC estimated using the simple CAPM, when a heterogeneous-expectations CAPM was more apposite, would likely underestimate the 'true' value investors require. It seems difficult to believe that setting the allowed WACC even further below that would not adversely affect investment and efficiency.⁵

⁴ Australian Competition Tribunal, *Telstra Corporation Ltd (No 3)* [2007] ACompT 3 at paragraph 450.

⁵ Even with an upward-sloping supply curve for finance, setting the allowed return below the WACC will obviously cause investment to fall below its efficient level, with the entire surplus from the forgone capacity being lost. The welfare loss is clearly even greater if the initial estimate of the WACC is itself too low, and a further shortfall is then added to that initial underestimate.

In other instances, the Tribunal has criticised in trenchant terms those who claim to rely on the conventional CAPM, and do so for the substance of their analysis, but then make modifications or adjustments ‘on a pragmatic basis’, thereby ‘working out of a Rate of Return [that] is neither true to the formula nor a conventional use of the CAPM.’⁶ Yet the discussion in its ULLS decision suggests the Tribunal did not understand the assumptions on which the conventional CAPM rests nor the conclusions it reaches.

Asset valuation

The Tribunal recently had to consider asset-valuation issues in a decision under the National Electricity Law.⁷ The case involves the valuation of easements owned by ElectraNet Pty Limited, and is the result of an application for review of a decision taken by the Australian Energy Regulator.

Of primary concern here is the Tribunal’s discussion of alternative asset-valuation methodologies. In that discussion, the Tribunal correctly notes that Depreciated Optimised Replacement Cost (DORC) — an asset-valuation approach in which assets are valued on the basis of the current cost of replacing their service potential, taking account of any scope for efficiencies (hence the term ‘optimised’ in the methodology’s description) — ‘has become generally accepted as the most appropriate value to attach to assets when they are first brought into an RAB’⁸ (where RAB refers to the regulated asset base). But the Tribunal then goes on to say that DORC ‘provides a valuation consistent with the long-run marginal cost of service provision, supports the maintenance of the capital required to deliver the service looking forwards, and prices and investor returns which would be expected to occur in a competitive market and hence promotes the efficient allocation of resources.’⁹

It is questionable whether any of these statements is correct. For example, DORC, which reflects a physical capital maintenance approach to income determination,¹⁰ essentially measures the replacement cost of service potential,

⁶ Australian Competition Tribunal, Application by GasNet Australia (Operations) Pty Ltd [2003] ACompT 6, paragraph 47.

⁷ Australian Competition Tribunal, Application by ElectraNet Pty Limited (No 3) [2008] ACompT 3. The National Electricity Law prescribes the approach to be adopted in respect of asset valuation. However, the valuation of the easements at issue in these proceedings was excluded from the scope of the relevant provisions, as it was the subject of ongoing dispute.

⁸ *Ibid.*, at para. 192.

⁹ *Ibid.*

¹⁰ Accounting theory distinguishes two concepts of income maintenance. The first, generally referred to as financial capital maintenance, defines the activity cycle as ‘cash to cash’, and broadly accounts for changes in the value of the funds owners have made available to the entity. This corresponds to an investor view of capital maintenance, in which income is defined with respect to the command the investor’s financial resources provide over goods and services. The second, referred to as ‘physical’ or ‘operating’ capital maintenance, views the activity cycle as ‘physical unit to physical unit’, and accounts for changes in the cost of providing a specified level of service potential. This rests on an enterprise view of capital maintenance, in which income is broadly defined as the amount left over for distribution

so that unit charges set on that basis reflect *average* costs. In an industry with decreasing costs, average costs will not reflect marginal costs in the short run or the long.¹¹ Nor is it clear that in a competitive market prices would be set to average DORC, and even less so that an equilibrium outcome could occur in which (in the absence of public subsidies) efficient prices would recover marginal, but not total, cost.

But be that as it may, the Tribunal then goes on to say (at paragraph 194) that despite its conclusions in respect of DORC ‘there is merit in the ACCC and the AER’s approach’ of using ‘indexed historic cost’ as the basis for valuing the easements, seemingly for two reasons. The first is that ‘indexed historic cost ... maintains investors financial capital intact, it covers the opportunity cost of their financial capital’. Again, it is by no means obvious that this is correct. For example, if the WACC has been set on a nominal forward-looking basis (so that it reflects anticipated inflation), it is not clear why indexation is needed to secure *ex ante* financial capital maintenance;¹² moreover, even were the Tribunal’s claim correct, it seems difficult to reconcile that claim with its earlier conclusion that setting charges on the basis of physical (not financial) capital maintenance ‘promotes the efficient allocation of resources’.

But the Tribunal’s second reason for endorsing historic costs in this context is even more striking: it is that the replacement cost of the relevant assets is rising over time. Specifically, the Tribunal says (at paragraph 194) that using DORC where replacement costs are rising would have the consequence of ‘providing windfall gains to investors and price shocks for consumers. The Tribunal has been alert to avoid such an outcome.’ But how can it be the case that one approach to capital maintenance is generally correct and efficient, *except* where asset values were rising? This implies a ‘heads I win, tails you lose’ approach to asset valuation, in which an access provider’s assets are valued on

once the firm has put aside enough to replace its current service potential. The choice between these has important implications for the treatment of changes in asset prices. An increase in replacement cost, for example, is treated as a holding gain in most valuation systems based on financial capital maintenance, but would be treated as requiring an increase in provisions under physical capital maintenance. These differences in the approach to changes in asset prices are reflected in differences in whether those changes flow into the income statement or are treated as solely affecting the balance sheet. An extended discussion can be found in Sterling & Lemke 1982.

¹¹ Of course, an industry can be a natural monopoly even with a U-shaped cost curve, if demand is only large enough to sustain one firm operating at minimum efficient scale. In that event, average and marginal costs could be equal at the regulated output level (assuming that level is set to the minimum point on the envelope curve defined by long-run average cost). However, there is no reason to think this describes cost conditions in electricity transmission and reticulation, which are more likely to display everywhere decreasing costs, especially in a relatively small system such as ElectraNet’s.

¹² Even if the WACC has been set on a forward-looking nominal basis, issues of *ex ante* financial capital maintenance may still arise if depreciation is set on an accounting, rather than an economic, basis. However, the Tribunal does not seem to have considered this issue, and even were it the problem, using indexed historical cost as the valuation base would not resolve it, as it would require recalculating the depreciation provisions and correcting any shortfall or surplus. The corrections required are discussed in Whittington 1983 and Franks and Hodges 1996.

the basis of the *lower* of physical (DORC) or financial (indexed historical cost) maintenance, which is necessarily inconsistent with either concept of capital maintenance ever being achieved (except in the uninteresting and unrealistic case in which there are no changes over time in relative asset prices).

Put slightly differently, what the Tribunal is saying is that the regulated firm should be:

- Fully exposed to 'windfall' losses (to use the Tribunal's terms), as would occur under a DORC methodology when the efficient replacement cost of service potential has fallen over time; but
- Not allowed to retain any 'windfall' gains, as would occur under DORC when the efficient replacement cost of assets has risen.

It is obvious that the effect of such an approach, considered *ex ante*, would be an expectation of less than full cost recovery, so long as there was some probability of some asset prices falling. Put in terms of the standard theory of regulation, this breaches the participation constraint, which requires that the investor have an *ex ante* expectation of full cost recovery, subject to the investment being efficient, and hence will deter otherwise socially desirable investment.¹³

The Tribunal seeks to deal with this by saying that unlike other assets, easements 'do not wear out' (paragraph 194). Putting aside whether easements are truly perpetual (and they are not if technological changes or population movement reduces or eliminates the benefits they provide), this overlooks the obvious point that any difference in asset life as between this asset and other assets is already picked up in the depreciation charge (which in the case of a perpetual asset would be set to zero). Why then would the fact that an asset was perpetual require or justify a change in the asset valuation methodology? And if the change is made for perpetual assets, why not also for assets that are not perpetual but have extremely long lives, and if so, for those that have lives only marginally shorter than extremely long, and so on? What is the relevant basis of principle?

The Tribunal additionally cites as a relevant consideration that the easements do not have alternative uses (again, at paragraph 195). However, even were this correct, it is merely a way of saying that investments made in easements are entirely sunk. But again, why an asset that is entirely sunk should be treated differently from one that is sunk to 99.9, 90, 80 or 50 per cent of its original value is unclear. Additionally, if some assets, once committed, are entirely sunk,

¹³ Analytically, of course, the correct approach in the event of the asset being periodically redetermined is to shape the depreciation profile so as to match in expectation the path of future price changes. This requires credible commitment by the regulator to the implied repayment profile. See Mandy 2002, Mandy and Sharkey 2003 and Ergas 2009.

one would have thought that this only made it more important that investors credibly believe that efficient cost recovery will be allowed in expectation; that is, taking account of the risk of future changes in valuation. Asset valuation on a heads I win, tails you lose' basis surely erodes that confidence.

In reality, the Tribunal seems concerned to set a fair' charge, rather than a charge determined by the consistent application of an economic model of asset valuation. Yet in doing so, the Tribunal again seems inconsistent, as it has previously (in a decision subsequently upheld by the High Court¹⁴) criticised regulators for 'departing [in the setting of access charges] from a quest for value and entering upon a quest for some form of justice or equity'.¹⁵

Airport charging

The third and final decision I wish to consider regards access to aeronautical facilities at Sydney Airport.¹⁶ Specifically, the matter before the Tribunal was an application for review sought by Virgin Blue Airlines Pty Limited (Virgin) of the decision of the Parliamentary Secretary to the Commonwealth Treasurer not to declare (that is, regulate) a service, described as the 'Airside Service', provided by Sydney Airports Corporation Limited (SACL) at Sydney (Kingsford-Smith) International Airport. The 'Airside Service' comprises the use of runways and the service of taking off and landing at the airport, as well as the charges the airport levies for those uses of its facilities.

The Tribunal set aside the decision by the Parliamentary Secretary to the Commonwealth Treasurer and decided that the Airside Service should be declared; that is, brought within the scope of the mandatory third-party access regime. Many contentious elements are involved in that decision and, as noted above, I do not seek to review them all. However, there is an economic argument that the Tribunal advances that seems to lie at the heart of its reasoning. That argument can be summarised as follows:

- SACL, having previously charged airlines on the basis of the weight of the aircraft at take-off or landing, adopted a charging arrangement that charged domestic airlines for use of its aeronautical facilities (such as runways and tarmacs) on the basis of an amount per passenger;¹⁷
- It did so in a context where there was competition between a 'Full Service' Airline (FSA), Qantas, and a 'Low Cost' or 'Value Based' Airline (VBA), Virgin;¹⁸

¹⁴ East Australian Pipeline Pty Limited v Australian Competition and Consumer Commission [2007] HCA 44.

¹⁵ Application by East Australian Pipeline Limited (2004) ATPR 42-006 at 48,804 [19].

¹⁶ Virgin Blue Airlines Pty Limited [2005] ACompT 5.

¹⁷ Ibid at para. 167.

¹⁸ Ibid at paras. 46-60.

- Per-passenger charges were likely to bear more heavily on the VBA than on the FSA, because the VBA would have higher levels of occupancy (that is, capacity utilization);¹⁹
- This would distort competition between the FSA and the VBA relative to alternative charging arrangements, and notably one where charges were based on the weight of the aircraft at take-off or landing;²⁰
- Moreover, charges based on weight were likely to be less distorting, in an economic sense, overall, as they related better to cost causation. As a result, an airport that was simply seeking to operate efficiently would retain charges based on weight.²¹

Serious issues could be raised about whether the Tribunal's arguments with respect to impacts of the different charging bases on competition are correct. What is clear is that any such impact requires consideration of a model of oligopolistic competition, and the relevance of the standard models to the situation in question is open to considerable doubt.²² But to the extent to which such models are relevant, it seems peculiar to think of a change in input prices as reducing competition. True, an increase in the relative price of labour will presumably 'advantage' those competitors that have chosen capital-intensive technologies, and 'disadvantage' those that have chosen labour-intensive technologies. But it will leave unchanged the coefficient of conjectural variation (which measures each firm's perception of its rivals' response to any attempted price change), and it is that coefficient that captures the intensity of competition, at least in its conventional sense (see Ergas, Ralph and Robson 2009). Otherwise, every material change in input prices — of which there must be many, in any period — could potentially 'lessen' competition (if it affected competing suppliers differently), and on the reasoning set out in the decision, expose input suppliers to regulation.

Perhaps so as to deal with this concern, the Tribunal suggests that the input price change at issue (that is, the shift from charging on the basis of weight to a per-passenger charge) was inefficient, and hence would not have been adopted

¹⁹ Ibid at paras. 204–6.

²⁰ Ibid at paras. 222, 524, 526, and 534.

²¹ Ibid at paras. 236, 240–1, 245, 253–4.

²² For example, the Tribunal appears to have relied on a Cournot model of competition between Qantas and Virgin. In effect, there is no reason to believe that competition between a VBA and an FSA can be properly reduced to a Cournot model: rather, it is more sensible to think that such a model will understate the intensity of competition between these airlines. This supposition — that a Cournot model will understate the intensity of rivalry between a VBA and an FSA — is consistent with empirical analyses of airline competition. In effect, the general result of the relevant literature is that while it may be the case that competition between FSAs can resemble the Cournot model of competition, competition between an FSA and a VBA is fiercer than the Cournot model would suggest. See (Weisman 1990; Oum, Zhang & Zhang 1993; Oum 1995; Reiss & Spiller 1989; Brander & Zhang 1990; Armantier & Richard 2003). Of course, as a referee has pointed out to me, even with this stronger form of competition, the Tribunal's finding that an increase in Virgin's marginal costs will increase output prices may still be correct.

in the absence of some intention to harm competition. Specifically, the Tribunal claims that the weight-based measure is cost-reflective, while the per-passenger basis is not. As a general matter, it is correct that mass at landing and take-off has some effect on the extent of the damage aircraft movements cause to runways. However, these marginal costs account for a very small share of the average costs of providing aeronautical services (Oum & Zhang 1990; Turvey 2000). As a result, any charge that recovers average costs will have to be very significantly marked up relative to the marginal damage cost if total costs are to be recovered. There is no particular reason to believe that imposing this marked-up charge (in which marginal damage costs will account for a small share) on a per-passenger basis will lead to outcomes that are materially worse, in an efficiency sense, than would arise from a charge based on weight. The Tribunal itself provides no such reason and seems to believe — wrongly — that charges based on weight would be cost-based.

In fact, and this is perhaps the central issue in respect of the Tribunal's consideration of efficiency, the efficiency impact of which basis was used for airport charging (per-plane, or per-passenger) would depend on where output was relative to the capacity margin (that is, whether planes are typically close to full or not). A per-passenger charge would be distorting when planes were relatively empty (so that capacity utilisation was the relevant margin), and a per-plane charge when they were relatively full (so that the relevant decision was whether to increase the number of planes). When this is uncertain, the least distortion is likely to result from some average of the two; that is, from a charge that combined a per-plane and a per-passenger element. This was not considered by the Tribunal, which simply concluded that per-passenger charges were undesirable.

Moreover, a per-passenger charge would seem closer to a tax on final demand than a weight-based charge (which might give rise to substitution between different types of aircraft). The modern theory of public finance suggests that the distortionary taxes required to fund any shortfall between average and marginal costs should fall on final consumers rather than on intermediate inputs. This follows from the Diamond-Mirrlees theorem, which states that so long as production efficiency is feasible, and private profits are negligible or can be taxed away, then public policies should be designed to achieve production efficiency, which entails shifting the burden of distortionary taxation away from intermediate inputs. To that extent, a per-passenger charge (or a weighted charge that placed greater weight on the per-passenger element) would be more efficient.

The more general point is that securing cost recovery in industries operating under declining costs requires consideration of how joint and common costs should be recovered. That consideration should reflect the impact that alternative approaches will have on decisions at the margin of use and of investment. It is

here that the Tribunal's decision appears least satisfactory, as it draws conclusions about efficient pricing without consideration of what the relevant margins of investment and use are, how they will be affected by alternative charging bases and what that implies for the design of an efficient charge.

Why do errors occur?

It would be unwise to infer, simply on the basis of the three decisions reviewed, that all, most or even many Tribunal decisions involve errors of this kind. Doing so would require a more comprehensive assessment than the one I have undertaken. Yet the errors I have pointed to are not minor, and should give rise to some consideration of the mechanisms that might be at work. My contention is that it is not surprising that such errors should occur; indeed, strong grounds can be found for thinking that an added risk of error is inherent in any mechanism that can effectively correct for the kinds of agency costs that arise in the regulatory regimes at issue. In this section, I sketch a theory of regulatory design that explains why that is the case.

Thus, agency costs arise because regulators, in these regimes, are given wide delegations. This is done so as to allow regulators to achieve economies of scale and scope in the regulatory task, including in the use of specialised expertise, and so as to allow them to structure and conduct regulatory operations in an efficient manner. This, in turn, involves granting regulators substantial institutional capabilities, in the form of access to expertise and to financial resources.

There is, however, an unavoidable tension between institutional capability on the one hand and effective accountability on the other (see Komesar 1997). Thus, granting an institution greater capabilities — in terms of its capacity to initiate and determine action, its access to resources (including of expertise), and its control over how it interacts with outside actors — increases both the information asymmetry between it and its principal (making it more difficult for its behaviour to be monitored and controlled), and its permeability to external influence.

One way of offsetting these risks is to subject that institution's decisions to review by a body that is less permeable to those influences, and generally less vulnerable to capture by its own interests or those of favoured constituents. Courts and tribunals are a natural choice for this role. First, these institutions are broadly outside the bureaucratic process of jockeying for budgets and resources, and individual adjudicators typically have little interest in bureaucratic 'empire building', reducing the risk of goal displacement. Second, legal proceedings impose relatively tight constraints on the way in which the parties and the Court interact (see Aubert 1967), and there is no equivalent in the Court setting to the close and informal interaction between the parties and

the decision-maker that characterises the regulatory process. This makes influence-seeking more difficult. Third, legal proceedings, though they may involve several stages of determination (first instance, appeal, and so on) are, at each stage, a one-shot game. There is consequently less scope for an interested party to invest in developing and maintaining expertise and influence specific to a particular decision-maker in legal proceedings, as well as less scope for collusion between the decision-maker and an interested party. Fourth, the rules pertaining to the disclosure and use of information in legal proceedings are generally highly formalised. The repeated 'give and take' game around access to information which prevails in most regulatory processes therefore does not arise, or arises only to a confined extent, in legal proceedings. This too makes influence-seeking by the party with an informational advantage more difficult.

However, the negative counterpart of these positive factors is substantially more limited access by the review body than by the primary or first-instance regulator to specialised expertise and less scope to seek and test information autonomously. After all, a central reason courts and tribunals are less permeable to influence activities than other institutions is precisely the fact that despite their expansion in recent years, they are not Weberian bureaus, in which bureau chiefs seek to secure, maintain and expand a pool of resources which give specialist expertise and the ability to implement complex decisions (Rosenberg, 2008). This inevitably constrains their scope for dealing with complex problems, with the resulting limitations likely to be particularly acute in areas that rely heavily on specialised expertise. As a result, while reliance on review by formal adjudicative structures reduces the risk of rent-seeking and may help correct errors made by the regulatory body, it can introduce errors, and consequent error costs,²³ of its own.

Implications for regulatory design: Rules and standards

Are there alternative approaches that could allow both agency costs and error costs to be reduced? The most obvious alternative would be to limit regulatory discretion. In particular, there could be scope to alter the balance between reliance on rules and reliance on standards.

Rules can, for this purpose, be defined as 'if-X-then-Y' statements crafted with sufficient specificity to determine the outcome once the qualifying condition (the 'X') is met.²⁴ A simple example of such a rule is the setting of a speed limit.

²³ Error costs are the efficiency costs to society of incorrect decisions. On the analysis of error costs and their relation to imprecise, vague or contentious tests see Kaplow 1994.

²⁴ The qualifying condition for a rule is usually referred to as its protasis, while the state to which meeting the protasis gives rise is the apodosis. In a rule which says 'If a process meets condition X, then that process is exempt from consequent Y', the criterion of meeting condition X is protasis, while the prescriptive consequence (being exempt from Y) is the apodosis. There is nothing that says that the test involved in applying the protasis needs to be same as that involved in implementing the apodosis.

In contrast, a standard, rather than determining the outcome consequent on the qualifying condition being met, defines the criteria that a decision-maker will use in determining, through an individualised assessment of each case, what that outcome should be. For example, subjecting the assessment of a driver's behaviour to a consideration of whether it was safe under the circumstances would be a standard rather than a rule.²⁵

Broadly speaking, rules provide greater up-front certainty and specificity in the decision situation, and economise on the costs of individualised assessment of particular cases. However, these advantages are bought at the cost of both likely greater required outlays in the up-front formulation of the rule, and of potential under- or over-inclusiveness in particular instances. For example, formulating the rule 'Driving at more than 80 kph is prohibited' involves an investment in determining *ex ante* what the appropriate speed limit should be; and then imposes costs because there are circumstances in which it would be perfectly safe to drive at 100 kph (so that the rule is over-inclusive), and others in which driving at 20 kph is reckless (the corresponding under-inclusiveness). In contrast, standards may require less investment in their initial formulation, but significant costs — in *ex ante* uncertainty and *ex post* enforcement — can arise from a standard's more open-ended formulation (see Shapiro 1965; Diver 1983; Kaplow 1992).

Put slightly differently, a central difficulty in governing through rules lies in their potential lack of flexibility; this makes it important to invest, possibly heavily, in getting the rules right. In contrast, standards avoid that initial investment and allow greater adaptation of decisions to individual cases, but involve higher costs in taking individual decisions, greater uncertainty as to regulatory outcomes and greater agency costs. The result is a greater need for substantive review of decisions based on standards, which increases the costs of relying on a standards-based approach.

Given those trade-offs, reliance on up-front formulation of a rule would seem most clearly justified where three conditions are met: first, the rule covers an issue of wide applicability, so that the initial investment can be spread over many individual instances; second, the costs of relying on such a rule, in terms of lost opportunities for particularisation, are low because there are few gains from adapting the treatment to the circumstances of specific cases; and third, individualised assessment of consequences will involve inherently contentious judgements that ultimately may need to be taken by decision-makers with limited expertise, will induce (essentially unproductive) expenditure in seeking to

²⁵ In thus characterising 'rules' and 'standards' I am following Hart & Sacks 1958–1994, and Kennedy 1976. Others, notably in the English legal tradition, view these as merely different points on a continuum, defined by dimensions of specificity: see Atiyah & Summers 1987; Twining & Miers 1991.

establish a case one way or another, and can result in high error costs and unnecessary uncertainty.

Issues such as the methodology to be used in determining allowed rates of return, asset valuation and efficient prices all seem like strong candidates for such a rule-based approach. After all, determining these on a different basis in different industries would merely distort investment flows and resource allocation, while creating opportunities both for error and for rent-seeking.

Relation to Australian regulatory regimes

How then does the prescription to place greater reliance on up-front rules relate to current Australian regulatory regimes?

Overall, one can distinguish two broad approaches within the Australian regulatory context.

The first of these approaches relies almost entirely on loosely specified standards. Its origins lie in the reforms enacted subsequent to an Inquiry chaired by Professor Fred Hilmer (Independent Committee of Inquiry into Competition Policy in Australia 1993). Among its many recommendations, that Inquiry recommended that there be a comprehensive, unified approach to the regulation of infrastructure industries, with that regulation to be based on principles set out in a new Part of the TPA. That new Part was enacted in 1995 as Part IIIA of the Trade Practices Act, which provides for mandated access to the facilities of a firm, when that access would meet a number of conditions and, in particular, would 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 on the facility owner a right of third-party access to that facility by means of 'declaration' of the service provided by that facility. Two features of the Part are particularly relevant here.

First, a National Competition Council (NCC) is established, which reviews applications for facilities to be declared and makes a recommendation to a Minister, typically the Commonwealth Treasurer. Decisions by the Minister about whether a facility is to be declared are subject to review by the Australian Competition Tribunal (the Virgin case, discussed above, being such an instance). While the Part sets down a number of criteria that must be met for a facility to be declared, in practice those criteria now involve a very wide discretion (see Ergas 2009).

Second, once a facility is 'declared', there is a process whereby, in the event of dispute, the terms and conditions of access by third parties to that facility are set by a regulator, generally the Australian Competition and Consumer Commission (ACCC). Some general principles are set out in the legislation to guide the ACCC in the exercise of its powers, as are a number of factors it must take into account, but it would be fair to say that even with recent amendments,

they leave a wide discretion. There is provision for review of the ACCC's decisions, again by the Australian Competition Tribunal.

As a result, the Part IIIA process can be seen as defining a broad canvas which it is broadly up to the regulator to fill, subject to some general standards — such as the requirement that pricing decisions take account of the legitimate interests of the facility owner — and to review of those decisions by the Australian Competition Tribunal.

Although the Hilmer recommendation was for a single, comprehensive approach to infrastructure regulation, the Part IIIA structure has come to co-exist with industry-specific regulatory arrangements covering a range of other sectors. Two of those regimes are especially relevant here.

The first, governing telecommunications, is primarily set out in Part XIC of the TPA, which came into effect on 1 July 1997. While I have described the details of those provisions elsewhere (Ergas 2008a; Ergas 2008b), suffice it to say they vest in the ACCC even greater discretion — both as to what is regulated, and as to the terms and conditions of regulated supply — than it enjoys under Part IIIA, again subject in some areas to review by the Australian Competition Tribunal. (The ULLS proceedings, discussed above, are such a case). Here, too, the regulators (the ACCC at first instance, and the Australian Competition Tribunal on review) make decisions subject only to very general standards.

The second regulatory regime, which contrasts with the pattern of wide-ranging delegation described above, now covers the regulated energy sector. Here a structure has been put in place that separates the functions of setting regulatory policy, translating that policy into detailed rules and then implementing those rules as regulatory decisions.

Specifically, a reform of the energy regulatory arrangements began in 2004 under the auspices of the Ministerial Council on Energy (MCE). The MCE comprises Ministers with responsibility for energy from the Federal Government and all States and Territories, and is the national policy and governance body for the Australian energy market, including for electricity and gas. It holds, in other words, primary responsibility for setting policy for energy markets, and hence clearly removes that policy responsibility from the remit of regulators.

Having thus secured and clarified the location of the policy role, the MCE has moved to separate from the regulators the responsibility for setting the rules under which they regulate. Specifically, the Council of Australian Governments (COAG), operating through the MCE, established the Australian Energy Market Commission (AEMC) in July 2005 to be the Rule-maker for national energy markets under the Australian Energy Market Commission Establishment Act 2004 (South Australia).

The MCE also established the Australian Energy Regulator (AER). The AER, though a constituent part of the ACCC, operates as a separate legal entity and has separate membership from the ACCC. The AER implements and enforces the rules that are made by the AEMC, and its powers are intended to be limited to that 'rule implementation' and enforcement role.

At the same time, given these more prescriptive up-front rules, the scope for substantive review of regulatory decisions in the energy industries has been somewhat reduced. There has, in other words, been a trade-off, in which greater specificity at the outset is offset by less scope for reconsideration of issues of principle once initial decisions are taken.²⁶

In short, Australian infrastructure regulation has two broad kinds of regimes: one that delegates to a decision-maker a broad discretion, subject to substantive review of its decisions by a review body; and another that places the primary weight on clearly specifying the rules up-front and doing so separately from the taking of particular decisions, while confining both the scope of the regulatory task and the potential for review of that task.

However, it is difficult to identify any compelling reasons why issues such as the setting of the weighted average cost of capital, the determination of the value of the asset base, or the determination of depreciation charges, should differ as between the industries covered by these regimes.²⁷ Mirroring in other regulated industries the rules-based approach that has been adopted in energy would help avoid the distortions that could be caused by dealing with these issues differently in different industries, improve certainty and reduce error and agency costs.²⁸

Conclusions

In summary, I have argued that:

1. Important areas of infrastructure in Australia are regulated on the basis of relatively vague standards, with the task of interpreting and applying those

²⁶ While the prescriptive nature of the rules narrows the scope of reviews of regulatory decisions, it does not necessarily leave the regulated businesses without redress. They can lodge rule-change applications with the AEMC, and have done so over a wide variety of rule requirements (with some successes and some failures).

²⁷ It is sometimes argued that the rapid pace of technological change in telecommunications justifies a more discretionary approach. I examine this argument in detail in Ergas 2008a and conclude that is poorly thought through. In fact, the rapid pace of technological advance in telecommunications provides compelling grounds for limiting regulatory discretion, both because market forces are more likely to be self-correcting and because regulatory interventions that distort the pace of technological advance have especially high social costs.

²⁸ There is some evidence of some of these benefits already being achieved by informal means, through State and Territory regulators covering other infrastructure sectors giving significant weight to the outcomes of the comprehensive review of cost of capital which the National Electricity Rules prescribes.

- standards being delegated to a regulator (generally the ACCC). That delegation creates a substantial risk of agency costs.
2. Controlling those agency costs requires vesting review rights in a body less permeable to outside influences and less vulnerable to goal displacement. Courts and tribunals — including the Australian Competition Tribunal — have natural advantages in this respect. However, a major element in making these structures less vulnerable to agency costs is precisely that they are not Weberian bureaus, in which bureau chiefs seek to secure, maintain and expand a pool of resources which give specialist expertise and the ability to implement complex decisions. This inevitably constrains their scope for dealing with complex problems, with the resulting limitations likely to be particularly acute in areas that rely heavily on specialised expertise.
 3. The result is that relying on review of regulatory decisions by these bodies can reduce agency costs, but does so at the risk of introducing error costs, as evidenced in my three case studies.
 4. An alternative would be to place greater reliance on regulation through relatively prescriptive up-front rules. This would require greater investment in the initial formulation of these rules, but would constrain regulatory discretion, and hence reduce agency costs. Additionally, devising these rules through a considered and focused deliberative process could allow errors to be detected and corrected in advance, more effectively avoiding error costs.
 5. Such an approach has now been adopted in the regulated parts of the Australian energy sector, and its extension to other regulated industries could improve the quality of regulation and help avoid distortions to resource allocation from differential treatment of issues such as the setting of the WACC.

Three further points are worth making by way of conclusion.

First, to the extent that such a rules-based approach were to be adopted to those areas which are currently regulated on the basis of broad standards, it would be desirable to also adopt the separation-of-functions model that is used in the energy industries.

Thus, a system that claimed to rely on rules would have little credibility if the rules were, or could be, merely a sham that allowed the regulator to achieve whatever outcome it believed desirable. Relative to those risks, vertical separation between rule-making and rule-implementation can play the same role that structural separation is often claimed to play in the design of the structure of regulated industries. More specifically, it can reduce the ability and incentive of the rule-maker to favour the rule-implementer, or more generally to facilitate rent-seeking behaviour by the rule-implementer. At the same time, separation means that the output of each stage in the vertically separated process — policy

formulation, rule definition, rule implementation — needs to be sufficiently transparent to allow the other stages to operate, with the added transparency facilitating accountability. Finally, the temporal sequencing of the tasks, the inevitable lags between them, and the resulting stickiness of the rules, reduces the risk of rules being ‘cooked’ so as to secure desired outcomes in individual cases; that is, the risk of collusion between the rule-maker and the rule-implementer. This adds to the credibility of the mechanism, and hence can make it more effective in securing increased investor confidence.

Second, given a move in this direction, it would be important to ensure consistent treatment across all regulated industries of generic issues, such as the setting of the cost of capital. This could be achieved either through a single rule-maker, or by statutory requirements on the separate bodies to ensure consistency of treatment.

Third, it is not my intention to suggest that implementing the approach I recommend would be a panacea for the myriad weaknesses of our regulatory arrangements. Rather, my contention is that to the extent to which activities have to be regulated, it is better if that regulation is predictable, transparent and free of bias. A rules-based approach has numerous advantages in that regard. But all forms of regulation are seriously imperfect, and the ‘first best’ solution is to be extremely cautious — far more cautious than our current regimes are — in what is brought within the regulatory net.

References

- Armantier, O. and Richard, O. 2003, ‘Exchanges of cost information in the airline industry’, *Rand Journal of Economics* 34(3): 461–7.
- Atiyah, P. S. and Summers, R. S. 1987, *Form and Substance in Anglo-American Law*, Clarendon Law Series, Oxford.
- Aubert, V. 1967, ‘Courts and conflict resolution’, *The Journal of Conflict Resolution* 11: 40–51.
- Brander, J. A. and Zhang, A. 1990, ‘Market conduct in the airline industry: an empirical investigation’, *Rand Journal of Economics* 21(4): 567–83.
- Dal Bó, E. 2006, ‘Regulatory capture: a review’, *Oxford Review of Economic Policy* 22(2): 203–25.
- Diver, C. S. 1983, ‘The optimal precision of administrative rules’, *Yale Law Journal* 93(1): 65–109.
- Ergas, H. 2008a, *Wrong number: resolving Australia's telecommunications impasse*, Allen & Unwin, Crows Nest, NSW.
- Ergas, H. 2008b, ‘Setting access prices: a critique of the ACCC’s approach in telecommunications’, *Agenda* 15(4): 35–58.

- Ergas, H., Ralph, E. and Robson, A. 2009, 'The ACCC Merger Guidelines: A Review', *Competition and Consumer Law Journal*, forthcoming.
- Ergas, H. 2009 (forthcoming) 'An excess of access? The economics of declaration under Part IIIA of the Australian Trade Practices Act'.
- Ergas, H. 2009, 'Time Consistency in Regulatory Price Setting: An Australian Case Study', *The Review of Network Economics* 8(2): 153–63.
- Franks, J. R. and Hodges, S. D. 1996, 'The Meaning of Accounting Numbers in Target Setting and Performance Measurement: Implications for Managers and Regulators' in Brief, R. P. and Peasnell, K. V. (eds), *Clean Surplus: A Link Between Accounting and Finance*, Garland Publishing, Inc, New York.
- Ganghof, S. 2006, *The Politics of Income Taxation: A Comparative Analysis*, ECPR Press, Colchester.
- Harrison, J. M. and Kreps, D. 1978, 'Speculative investor behaviour in a stock market with heterogeneous expectations', *Quarterly Journal of Economics* 92(2): 323–36.
- Hart, H. M. and Sacks, A. M. 1958–1994, *The Legal Process*, Eckridge and Frickey (eds) Foundation Press, Westbury, N.Y.
- Independent Committee of Inquiry into Competition Policy in Australia, 1993, *National Competition Policy ('Hilmer Report')*, Australian Government Publishing Service, Canberra.
- Kaplow, L. 1994, 'The value of accuracy in adjudication: An economic analysis', *The Journal of Legal Studies* 23: 307–402.
- Kaplow, L. 1992, 'Rules versus standards: an economic analysis', *Duke Law Journal* 42: 557–629.
- Kennedy, D. 1976, 'Form and substance in private law adjudication', *Harvard Law Review* 89: 1685–1778.
- Komesar, N. K. 1997, *Imperfect Alternatives: Choosing Institutions in Law, Economics and Public Policy*, University of Chicago Press, Chicago.
- Lintner, J. 1965, 'The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets', *Review of Economics and Statistics* 47(1): 13–37.
- Mandy, D. M. 2002, 'TELRIE Pricing with Vintage Capital', *Journal of Regulatory Economics* 22(3): 215–49.
- Mandy, D. M. and Sharkey, W. W 2003, 'Dynamic Pricing and Investment from Static Proxy Models', *Review of Network Economics* 2(4): 403–39.

- Oum, T. H. 1995, 'A productivity performance of the world's major airlines', *Journal of Air Transport Management* 2(3-4): 181-95.
- Oum, T. H., Zhang, A. and Zhang, Y. 1993, 'Inter-firm rivalry and firm-specific price elasticity in airline markets', *Journal of Transport Economics and Policy* 27(2): 171-92.
- Oum, T. H. and Zhang, Y. 1990, 'Airport pricing: congestion tolls, lumpy investment, and cost recovery', *Journal of Public Economics* 43(3): 353-74.
- Reiss, P. and Spiller, P. 1989, 'Competition and entry in small airline markets', *Journal of Law and Economics* 32(2): S179-S202.
- Rosenberg, G. N. 2008, *The Hollow Hope: Can Courts Bring About Social Change?* The University of Chicago Press, Chicago, Ill.
- Shapiro, D. L. 1965, 'The choice of rulemaking or adjudication in the development of administrative policy', *Harvard Law Review* 78: 921-72.
- Sharkey, W., Laffont, J.-J., Gasmi, F. and Kennet, M. 2002, *Cost Proxy Models and Telecommunications Policy: A New Empirical Approach to Regulation*, MIT Press, Cambridge, Mass.
- Sharpe, W. F. 1964, 'Capital asset prices: a theory of market equilibrium under conditions of risk', *Journal of Finance* 19(3): 425-42.
- Sterling, R. R. and Lemke, K. W. 1982, *Maintenance of Capital: Physical and Financial*, Scholars Book Co., Houston, Texas.
- Turvey, R. 2000, 'Infrastructure access pricing and lumpy investments', *Utilities Policy* 9(4): 207-18.
- Twining, W. and Miers, D. 1991, *How To Do Things With Rules*, Butterworths, London.
- Weisman, E. 1990, *Trade in Services and Imperfect Competition: Application to International Aviation*, Kluwer Academic Publishers, Holland.
- Whittington, G. 1983, *Inflation Accounting: An Introduction to the Debate*, Cambridge University Press, Cambridge.
- Wilson, R. 1987, 'Game-Theoretic Analyses of Trading Processes' in T. Bewley (ed.), *Advances in Economic Theory: Fifth World Congress*, Cambridge University Press, Cambridge: 33-70.