

2013

Why Johnny can't regulate: the case of natural monopoly

Henry Ergas

University of Wollongong, hergas@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/smartpapers>



Part of the [Engineering Commons](#), and the [Physical Sciences and Mathematics Commons](#)

Recommended Citation

Ergas, Henry, "Why Johnny can't regulate: the case of natural monopoly" (2013). *SMART Infrastructure Facility - Papers*. 109.

<https://ro.uow.edu.au/smartpapers/109>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Why Johnny can't regulate: the case of natural monopoly

Abstract

This paper examines the difficulties inherent in regulation as a solution to market failure and, especially, to natural monopoly. It highlights the way regulation itself introduces new risks into the supply of natural monopoly services, including the risk of regulatory opportunism, and argues that delegating regulatory functions to 'independent' regulators does not in itself solve those risks.

Keywords

can, regulate, why, johnny, monopoly, case, natural, t

Disciplines

Engineering | Physical Sciences and Mathematics

Publication Details

Ergas, H. (2013). Why Johnny can't regulate: the case of natural monopoly. *Agenda: a journal of policy analysis and reform*, 20 (1), 47-54.

Why Johnny Can't Regulate: The Case of Natural Monopoly

Henry Ergas¹

This paper examines the difficulties inherent in regulation as a solution to market failure and, especially, to natural monopoly. It highlights the way regulation itself introduces new risks into the supply of natural monopoly services, including the risk of regulatory opportunism, and argues that delegating regulatory functions to 'independent' regulators does not in itself solve those risks.

Imagine that, in visiting a primary school, a politician were to ask a child what he or she wants to do as a grown-up. Now, if the child said 'I'm going to be a pilot', the politician would exude the milk of human kindness, pat the infant on the head, and move swiftly on to the next photo-opp. But what if little Johnny said 'When I grow up, I want to be a regulator'? Surely a moment of bewilderment would follow. After all, no one could impugn a child's aspiration to be a pilot, doctor, plumber or nurse; at a pinch, even neighbourhood dope peddler, pole-dancer or economist, all professions whose origins lie lost in the mists of time. But what would one make of the desire to be a regulator, other than a diagnosis of early-onset power-hunger? And could one in good conscience commend that career choice, as one likely to ultimately lead to personal fulfilment and worthwhile community service?

To raise that question is not to chastise the vocation of regulator. But is there a sense in which — to borrow Enoch Powell's famous dictum about politics (Powell 1977: 151) — all regulatory careers are doomed to end in failure? Because that is in the nature of regulation and human affairs? While any such assertion might go beyond the evidence, all those familiar with the regulation of natural monopoly would agree that it is inherently contentious in process, controversial in its results and all too often disappoints the expectations vested in it.

To a degree, that could be viewed as simply confirming the fact that there are no perfect solutions to market imperfections. But that is both trite and unrevealing — it says little about why the problems occur. Yet a better understanding of those problems is needed if issues such as the scope of regulation and its design are to be sensibly addressed.

In exploring those problems, this article seeks to go behind and beyond conventional explanations of the difficulties inherent in natural monopoly

¹ The University of Wollongong and Deloitte Access Economics; ergas.henry@gmail.com

regulation. Although diverse, those explanations often boil down to the view that regulation would work tolerably well if only pesky interest groups would not derail the process. Yet that obviously assumes the regulator — with its own interests and objectives — would merely pursue the public interest, presumably defined in terms of maximising social welfare. Why that cannot be simply assumed then has implications for the framing of mechanisms that hold regulators accountable.

The implication of this analysis is not that regulation is per se undesirable but that we need to pay more serious attention to the importance of providing effective safeguards against the abuse of regulatory discretion. In contrast, recent years have seen moves to water down those safeguards. The present course of regulatory policy in Australia therefore poses some serious issues, which deserve more attention than they have received.

The problem being addressed

Any introductory textbook in economics will explain that we regulate so as to correct market failures; that is, situations where voluntary exchange will not result in the best use of society's resources. It will cite a number of factors that may give rise to such failures, including public goods, natural monopolies and decision-relevant externalities.

If it is a good textbook, it will then go on to explain that none of those market failures in itself justifies regulation. After all, it will say, it may be that bargaining between individuals could address those market failures, so long as property rights are well-defined. Take, for instance, natural monopolies and local public goods. These could be procured or supplied by user cooperatives, an example of the Coase theorem at work (Hansmann 1969). But then it would go on to say that the assumptions underpinning the Coase theorem — well-defined property rights, negligible transactions costs — rarely if ever hold, so such an outcome is unlikely.

This would be followed by an explanation of the specific difficulties that would arise in trying to design and implement the Coase bargain involved in voluntary, unregulated, supply of a natural monopoly. It would list three:

- First, unless the number of participants is very small, so the natural monopoly is (say) a mini-local public good, no potential participant will be pivotal, creating incentives for free-riding and removing any incentive for truthful revelation of valuations (Dixit and Olson 2000). As those valuations are consumers' private information, there will be no easy mechanism for assuring that the marginal and total conditions for efficient supply (which

require that the sum of the marginal valuations is equal to marginal cost, while *ex ante*, the sum of the total valuations equals total cost) are met.

- Second, it would note that asymmetric information problems will also arise on the supply side. Information about costs is private to the supplier; unless an auction or tender (a la Demsetz 1968) can force revelation of that information — and such auctions can only be fully effective where sunk costs are low and learning economies insignificant — it will be impossible to avoid paying a rent to the provider (Williamson 1976; Spiller 2013; Gasmi *et al.* 2002).² In other words, the provider will extract a rent associated with its monopoly not merely over the supply of the good, but over information about the costs involved in its production.
- Third, it would say that even if those hurdles could be overcome, there would still be difficulties involved in designing, monitoring and enforcing an efficient contract between the consumers and the supplier (Goldberg 1976). Given that no contract can cover every contingency, consumers would be exposed to the risk of the chosen supplier increasing its returns at consumers' expense, for instance, by shading quality (assuming quality levels are difficult to contract for; that is, non-contractible). And equally, once its costs were sunk, the supplier would be exposed to the risk of *ex post* opportunism by the customers, who might seek to expropriate some or all of those costs, leaving the supplier with just enough revenue to cover variable costs. Given those risks, each side would invest in costly precautions so as to ward them off. But as these redistributions are zero-sum, those investments, seen from a societal perspective, are merely a waste, reducing welfare.

Given all of these impediments, our introductory textbook would say, government intervention is needed to implement the Coasian bargain; that is, to devise and enforce the bargain that, in the absence of transactions costs, would result from voluntary exchange.

Regulation as *deus ex machina*

All this is true, as far as it goes. But it is also obvious that the logic involves an element of sleight of hand. After all, merely saying government is needed to devise and enforce the Coase bargain tells you nothing about how government would know what that bargain was or, if it knew, why it would actually implement it. The bare fact that government is now a player in no way solves

² If those assumptions (low or no sunk costs and insignificant learning economies) are not met, incumbents will have advantages over new entrants in bidding for the concession contract. Those advantages will translate into a rent they secure; to the extent to which that rent is a rent on their superior information, it is referred to as an 'information rent'.

any of the problems that vitiated the Coase bargain in the first place. Rather, all of the obstacles to efficient outcomes listed above remain; but then there are some new ones as well.

Those additional obstacles to efficiency are inherent in the principal–agent problems interposing government between the consumers on the one hand and the supplier on the other creates.

To begin with, in the idealised Coase bargain, consumers, as the principals, bargain with the supplier by putting dollars on the table, and hence, subject to all the issues mentioned above, signal an intensity of preference. In contrast, when consumers as voters have to express their preferences to government as their agent, they use the ballot box, creating the preference aggregation paradoxes that have delighted economists and social scientists since the days of Condorcet.

Second, if the contract envisaged in the Coase bargain outlined above was imperfect, that between consumers as the principal and government as the agent is likely to be orders of magnitude worse (North 1990; Dixit 1998). In effect, few contracts are as fuzzy as the implied agreement between voters and their elected representatives in terms of defining what is being exchanged and how. It is therefore extraordinarily difficult to measure what is being traded in political markets and in consequence to monitor and enforce the implied agreements. The incentives for voters to shirk on monitoring — that is, to free-ride on the monitoring efforts of others — only compound the resulting inefficiencies.

Third, political markets do a poor job of disclosing the information needed to monitor performance. For instance, if governments shift benefits to the present, while postponing costs to the future, there is no capitalisation mechanism, similar to that which operates (however imperfectly) in securities markets, that will instantly crystallise the resulting change in wealth in a way readily visible to consumers and taxpayers. This limits the ability of the principals to assess the agent's performance and, by the same token, weakens the disciplines the agent faces.

Fourth, the agent involved in this transaction is highly unusual: it is an agent with many powers of compulsion over the principal. Indeed, that is its main merit, as it is those powers of compulsion that allow it to overcome the obstacles voluntary exchange faces to attaining the Coase bargain. But those same powers of compulsion mean it can refuse to disclose information, force the supposed principal to make financial contributions regardless of that principal's valuations, and redistribute income directly and through cross-subsidisation. In bargaining with this agent, the supposed principals are mere lambs.

Fifth, those coercive powers magnify the risks of *ex post* opportunism. Consider a supplier investing substantial funds that, once committed, will be sunk. The likelihood of those sunk costs being expropriated is surely all the greater when the decision to do so rests with government on a three- or four-year electoral cycle, given the political gains even short-term price reductions could secure. The fact that the government has many instruments it can use to compel ongoing supply or punish quality degradation can only make the danger of expropriation greater. There is, in other words, a risk of time-inconsistency in government behaviour, where time-inconsistency refers to situations in which conduct of a policymaker that is rational *ex ante* is not rational *ex post* (and is known to be not), so that rational actors will discount the probability of a commitment to that conduct being maintained (Ergas 2008; 2009a).

In short, there is nothing in the mere fact of introducing government that solves the problems that impede the Coase bargain. Rather, the very coercive powers that are required to resolve the difficulties of voluntary exchange create new problems and inefficiencies.

Delegation and independence as the solutions

At this point, the conventional solution is to hand-wave about these difficulties being overcome through the delegation of power to an independent regulator. The story goes something like this (Komesar 1997; Vibert 2007; Helm 1994; Helm 2004):

First, a specialised regulator can reap economies of scale and scope in securing information, diminishing the information asymmetry between consumers and the natural monopolist. This can both reduce the information rent that needs to be paid to that monopolist and allow the provision of stronger incentives for productive efficiency.

Second, it may be easier for consumers to monitor a specialised regulator than a more generic bureaucracy, such as a ministerial department. Specialisation, in other words, imposes a degree of specificity in the disclosure of information and, in turn, can facilitate the development of specialised monitors, improving the efficiency with which the regulatory contract is implemented and enforced.

Third, and perhaps most important, distancing the regulator from central government can help overcome commitment problems; that is, can reduce the risk of time-inconsistency, lowering the cost of capital to the regulated industry and improving both producer and consumer welfare.

Now, the first two of these claims are possibly correct, but it is by no means clear that they are especially strong. It seems odd, for instance, to say that the mere fact of shifting functions from a ministerial department to an independent agency increases the ability to acquire information: perhaps it does, but one would need more to explain why. As for disclosure of specialised information, perhaps, but surely that too depends. Rather, the real work in this argument is done by the third claim: the claim that formal independence allows distancing, which in turn enhances policy credibility.

The (erroneous) analogy to central banks

In theory, there are indeed situations where delegating authority to an independent regulatory agency could increase the credibility of commitments in ways that reduce the risk of time-inconsistency. Such an enhancement of credibility most likely occurs where the regulator has a greater *ability* than general government, and/or greater *incentives* than general government, to act consistently over time — for example, because it does not benefit from time-inconsistent behaviour and/or has a unique preference for time-consistent behaviour.

These conditions, it is often claimed, underpin the case for the delegation of powers over monetary policy to central banks, providing a useful analogy for analysing the issues of regulatory design (Singleton 2011). Whether this is a sensible characterisation of central bank independence is a matter of intense academic debate, both as regards the solidity of its theoretical foundations and its empirical relevance. So too is the question of whether, as a factual matter, central bank independence, however defined, actually reduces the sacrifice ratio (that is, the costs of disinflation) or more generally shields central banks from political pressures, all the more so given the at times controversial (and seemingly politically responsive) decisions supposedly independent central banks have taken in the wake of the global financial crisis.

That said, whatever the merits may be of the claim that delegation enhances policy credibility in the context of central banks, it is by no means obvious that the same benefits will flow in the context of industry regulation (Ergas 2010). Indeed, there are a number of obvious difficulties with the argument that the delegation of powers to utility regulators strengthens and enhances the credibility of socially beneficial commitments.

In effect, virtually by definition, delegation entails vesting a degree of discretion. Now, while the notion of regulatory discretion is inherently complex, the extent of the discretion given to a regulator is typically characterised in the economic literature in terms of the extent of the constraint the regulator faces in two

respects: 1) the outcomes it seeks to achieve ('goal' or 'outcome' independence); and 2) the instruments it can use to achieve those outcomes ('input' or 'instrument' independence).

That characterisation helps refine the comparison between central bank independence and the independence of utility regulators.

Thus, the Reserve Bank of Australia is independent in many respects, but the outcomes it is to pursue are narrowly defined in terms of a target inflation rate over the course of the economic cycle. Measuring whether it is achieving that target is not simple, but it is clear that gross departures from the target range would be quickly detected and widely commented on. At the same time, the range of instruments it can use to pursue that target is also narrowly defined, with the cash rate being the primary lever under its control. As a result, the substance of the Reserve Bank's independence consists of the freedom it has in gauging how best to use a well-defined set of instruments to achieve a narrowly set goal.

In contrast, an economic regulator, such as the Australian Competition and Consumer Commission (ACCC), may have objectives that are quite vaguely defined, as is certainly the case for Parts IIIA and XIC of the Competition and Consumer Act (which set out, respectively, a framework for economy-wide access regulation and an industry-specific regulatory regime for telecommunications). Moreover, a regulator can have substantial control over the range of instruments it uses in pursuit of those objectives, both through the ability to alter the range of services to which regulation applies and through its control over what it includes or excludes from the scope of regulatory determinations (for instance, in terms of the degree to which it prescribes particular levels of service or other aspects of the non-price terms and conditions of access).

In other words, in contrast to the typical central bank, utility regulators have a high degree of both outcome and instrument independence. Four additional points strengthen this contrast.

First, the scope for industry regulators to commit to acting or not acting in particular ways in future is limited as a matter of statute, practicality or both. For example, with very long-lived assets, the regulator may not be able to commit its conduct for the life of those assets, if nothing else because changes in circumstances may require the re-opening of regulatory commitments (or make the costs of failing to do so especially high). For example, shifts in supply and demand conditions are likely to create differences between existing prices and efficient prices; the longer the adjustment lag, the greater may be the resulting

social costs. This constrains the degree to which industry regulators can commit to not revisit previous decisions and so reduces the credibility of commitments against time-inconsistency.

Second, the gains to a utility regulator from time-inconsistency are likely to be substantially greater than those accruing to a central bank. Thus, industry regulators may benefit from delivering immediate benefits to well-defined, favoured constituencies in a way that is extremely unlikely to occur for a central bank. For example, an industry regulator concerned about its resources and wishing to expand its political support may well face very strong incentives to cut regulated prices (be it to access seekers, particular end-consumers, or both), regardless of the impacts on future investments (and on regulators many years hence). While the whole notion of an independent central bank is that it does not 'internalise' the gains from time-inconsistency in the way politicians do, an independent utility regulator may well 'internalise' those gains (say, in the form of a greater budget) while the costs are shifted into future periods.

Third, the costs of time-inconsistency are likely to be less readily and quickly apparent to a utility regulator than for a central bank. In monetary policy, both output and inflation expectations respond relatively quickly to changes in the stance of the central bank. Moreover, those inflation expectations are capitalised into interest rates and quickly affect the slope of the yield curve. As a result, deviations from time-consistency are relatively rapidly translated into visible signals that rebound on the reputation and standing of the central bank. In contrast, in utility networks, especially those with some buffer of excess capacity, prices can fall far below long-run marginal costs before much changes. Moreover, in such networks, capacity augmentation is usually quite lumpy, and capable of some degree of postponement. As a result, there are usually few visible signs of time-inconsistency, other than the share prices of those regulated entities that are privately owned, which are inevitably subject to many and varying interpretations. While the costs of time-inconsistency in monetary policy come home to roost reasonably fast, in utility regulation it can be many years before the crunch occurs.

Fourth and last, there is a striking difference between central banks and utility regulators in terms of behaviour. In most models of the gains from delegating monetary policy to an independent central bank, a newly independent bank must invest in creating a reputation for time-consistency — that is, for placing a greater weight on inflation-reduction than on avoiding falls in output. To credibly signal that attachment to an anti-inflation stance, it must force on the community a higher cost, in terms of forgone output, than it would be rational for a decision-maker to impose were that decision-maker any less averse to inflation. Typically, this is achieved by appointing as the decision-maker an

'inflation hawk'. To some extent, the behaviour of the European Central Bank in its first years and of the Reserve Bank of New Zealand in the early 1990s is consistent with this characterisation.

Transposed to the utility regulation context, this would suggest that the initial period following the transition to regulatory independence would be associated with regulators credibly signalling a strong aversion to expropriating sunk investments. But in no major country has that been the case, and it certainly did not occur in Australia. If anything, utility regulators seemed to preference delivering price reductions to consumers, as evidenced by tough price caps and steep falls in regulated revenues. While many contrasting interpretations can be placed on those events, they do not seem to conform to the pattern typically ascribed to independent central banks and which underpins the claim that delegation (and investment in a reputation for 'toughness') enhances time-consistency.

Dealing with the risks

In short, there are significant differences in the substance of the discretion vested in independent central banks on the one hand and utility regulators on the other. As a result of those differences, the risks of opportunism and time-inconsistency in utility regulation are likely to be high, despite or indeed because of statutory independence.

In practice, these risks are compounded by the fact that statutory independence from the political process may not translate into independence as a matter of substance. After all, the statutory framework in which an agency operates is devised by the very politicians from whom it is intended to be independent. There is an extensive literature in the economic theory of politics that discusses what is called the 'mirroring principle', by which a regulatory agency's rules, budget processes and membership are designed so as to mirror and entrench the preferences and bargaining environment that best suit its political masters (McCubbins, Noll and Weingast 1987).

At the same time, that design and its operation are likely to be shaped by what the political scientists call 'the law of anticipated reactions', which simply says that actors are forward-looking and strategic in choosing among alternative courses of conduct (Weingast 1997). For instance, behaviour which is likely to severely compromise an agency's political support is surely less attractive to rational actors than a more moderate alternative. This is all the more the case as sagas such as that of the Murray Darling Basin Commission (which has repeatedly

changed its views and even leadership in the light of political reaction to its recommendations) highlight just how transient the claimed insulation of expert processes from political intervention may prove to be.

As a result, it is not unreasonable to expect regulatory outcomes to be considerably affected by the wider political context. Indeed, there are many episodes in recent Australian regulatory history which are inexplicable in other terms. Given that, how much can be done to discipline regulatory decision-making in situations where the risks of opportunism are high?

This is a huge issue, which I have covered to some extent elsewhere (see Ergas 2009b and 2012). But my own sense is that in terms of the design of regulatory institutions, at least at a Commonwealth level, there are many instances where we are going backwards.

The scope, for example, for merits review of regulatory decisions has been almost entirely eliminated in telecommunications, though (having done that to Telstra) new, substantial, restrictions on the regulator are being introduced so as to protect the National Broadband Network (NBN). However, those restrictions are not really safeguards protecting the integrity and balance of the regulatory process. Rather, they mainly involve enhanced ministerial powers of direction, and hence, if anything, make decision-making more rather than less political.

The legislative changes also provide far-reaching exemptions for the NBN from the competition laws, removing whatever safeguards the scope to turn to those laws would otherwise have provided. These are the first such exemptions of any substance since the Hilmer reforms. It is surprising, and not readily reconciled with conventional notions of regulatory independence, that our competition authority appears to have publicly acquiesced in these changes, despite its avowed commitment to a public-advocacy role in defending and promoting the integrity of the competition laws.

Moreover, a similar watering down of the safeguards against regulatory error and bias seems under way in electricity, following public outcry over rising electricity prices. While the precise changes to the electricity regime are still largely to be legislated, there is little doubt their effect will be to materially increase regulatory discretion.

In short, regulation is no magic wand that can, at no or low cost, replicate the outcomes of the idealised Coasian bargain. Moreover, conventional nostrums for the conundrums it raises — most obviously, the ‘solution’ of delegating powers to an independent regulator — create many problems of their own. To those problems there are no simple solutions, and most of the attempted solutions have yielded very mixed results.

As a result, Johnny may be on to something — there is certainly no shortage of career opportunities for regulators and regulation is, as they say, steady work. But if he expects regulation to meet the lofty goals so often set for it, that is less simple and far less assured.

References

- Demsetz, H. 1968, 'Why Regulate Utilities?' *Journal of Law and Economics* 11(1): 55–66.
- Dixit, A. 1998, *The Making of Economic Policy: A Transaction-Cost Politics Perspective*, MIT Press, Cambridge MA, ISBN 0262540983, 9780262540988.
- Dixit, A. and Olson, M. 2000, 'Does voluntary participation undermine the Coase Theorem?' *Journal of Public Economics* 76(3), June: 309–35, ISSN 0047-2727, [http://dx.doi.org/10.1016/S0047-2727\(99\)00089-4](http://dx.doi.org/10.1016/S0047-2727(99)00089-4)
- Ergas, H. 2008, *Wrong Number*, Allen & Unwin, Sydney, Australia.
- , 2009a, 'Time Consistency in Regulatory Price Setting: An Australian Case Study', *The Review of Network Economics* 8(2): 153–63, ISSN (Online) 1446-9022, DOI: 10.2202/1446-9022.1174.
- , 2009b, 'Error and Design: Economics in (and some economics of) the Australian Competition Tribunal', *Agenda* 16(3), ISSN 1447-4735 (Online).
- , 2010, 'New policies create a new politics: issues of institutional design in climate change policy', *The Australian Journal of Agricultural and Resource Economics* 54(2): 143–64, DOI: 10.1111/j.1467-8489.2010.00484.x.
- , 2012, 'Review of the Limited Merits Review Regime: Submission on Economic Aspects of Merits Review', Standing Committee on Energy Regulation.
- Gasmi, R., Kennet, D. M., Laffont, J.-J. and Sharkey, W. W. 2002, *Cost Proxy Models and Telecommunications Policy*, Cambridge, MA: MIT Press.
- Goldberg, V. P. 1976, 'Regulation and Administered Contracts', *Bell Journal of Economics* 7(2): 426–48.
- Hansmann, H. 1996, *The Ownership of Enterprise*, Harvard University Press.
- Helm, D. 1994, 'British Utility Regulation: Theory, Practice, and Reform', *Oxford Review of Economic Policy* 10(3), Oxford University Press: 17–39.

- , 2004, 'Regulatory Reform, Capture, and the Regulatory Burden', *Oxford Review of Economic Policy* 2(22): 169–85, DOI: 10.1093/oxrep/grj011.
- Komesar, N. K. 1997, *Imperfect Alternatives: Choosing Institutions in Law, Economics, and Public Policy*, University of Chicago Press, ISBN: 0226450899, 9780226450896.
- McCubbins, M. D., Noll, R. G. and Weingast, B. R. 1987, 'Administrative Procedures as an Instrument of Political Control', *Journal of Law, Economics, and Organization* 3(2): 243–77.
- North, D. C. 1990, *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, Cambridge UK, ISBN: 0521397340, 9780521397346.
- Powell, J. P. 1977, *Joseph Chamberlain*, Thames and Hudson, London, ISBN: 0500011850, 978-0500011850.
- Singleton, J. 2011, *Central Banking in the Twentieth Century*, Cambridge University Press, ISBN: 1139495208, 9781139495202.
- Spiller, P. T. 2013, 'Transaction cost regulation', *Journal of Economic Behavior & Organization* 89: 232–42, ISSN 0167-2681.
- Vibert, F. 2007, *The Rise of the Unelected: Democracy and the New Separation of Powers*, Cambridge University Press, ISBN:1139464728, 9781139464727.
- Weingast, B. R. 1997, 'The Political Foundations of Democracy and the Rule of Law', *The American Political Science Review* 91(2): 245–63, <http://www.jstor.org/stable/2952354>.
- Williamson, O. E. 1976, 'Franchise Bidding for Natural Monopolies — In General and With Respect to CATV', *Bell Journal of Economics* 7: 73–104.