# ­­3 The rationale for regulation of electricity networks

|  |
| --- |
| Key points |
| * In any given geographical area, electricity networks are natural monopolies as it would be uneconomic for another business to duplicate the infrastructure. * The theory and evidence about the behaviour of natural monopolies suggests that without strong regulation, network businesses could be expected to set excessively high prices and potentially provide too low a quality of services. * They would also face fewer incentives for internal efficiency and greater motivations for rent seeking to shore up their unregulated status. * These resulting static and dynamic economic inefficiencies provide a compelling rationale for regulation of network businesses. * Moreover, while not necessarily an economic inefficiency, high unregulated prices: * lead to potentially undesirable transfers from customers to businesses * would be regarded as unfair by many consumers. * These provide further grounds for regulation, though they do not greatly assist the actual determination of the right price. * Concerns about the impact of higher prices on particular customers are a less credible basis for competition regulation since their correction requires cross-subsidies between customers. Unless a special case can be made, such equity concerns should be addressed through other means. * Recently, the importance of the above rationales for competition regulation has been questioned. Instead, it has been argued that the main goal of regulation is to prevent the monopoly business from holding customers to ransom once they have made sunk investments that use the inputs of the monopoly business. Were this true, it would discourage customer investment in the first place (the ‘hold-up’ problem). * However, the market circumstances that would lead to the hold-up problem are mostly absent, markets have other solutions to the problem were it to occur, and regulators can create the same problem for the investments in the monopoly essential service (with even worse efficiency and distributional outcomes for customers). * The orthodox rationales for strong regulation of natural monopoly are sound. |
|  |
|  |

The appropriate regulation of electricity networks depends on the features of those networks (section 3.1) and the problems that would arise were they to be unregulated (section 3.2). The fact that electricity networks have market power is not controversial. However, the various rationales for regulating a natural monopoly — a core issue for regulatory remedies, including the role of benchmarking — remains complex and contested (section 3.3). Indeed, most recently, there have been challenges to the conventional wisdom on why regulation is necessary. Drawing on a new theory, different policy approaches have been proposed (section 3.4 and 3.5). The rationales implied by the considerations above help frame the appropriate policy goals of the regulator (section 3.6).

## 3.1 The characteristics of electricity networks

Electricity networks have several distinctive characteristics that affect the desirability and nature of regulation (most especially economic regulation).

In each geographic segment, supply of electricity costs are minimised through supply by a single distributor or transmission business — or a ‘natural monopoly’. This reflects some collective features of electricity networks, including the:

* very large fixed costs of the network (and low marginal operating costs)
* scarcity of easements required by distribution networks (and the opposition by householders and local governments to any potential duplication)
* safety concerns for consumers and workers were there to be multiple wires (overhead or underground) owned by separate businesses
* need for the system to act as a coherent network, with appropriate frequency and voltage control
* need for power on demand, and the incapacity to store power efficiently, which restricts the use of many distributed energy options and requires a very reliable network built to peak demand
* advantages of having generators located close to cheap energy sources that may nevertheless be some distance away from end users
* millions of customers, most with very limited countervailing bargaining power (unlike, for example, airports with airlines).

As such, there is no genuine capacity for new entry by a rival in any given area, unlike electricity generation or retailing. For example, an urban street would not sensibly have two sets of power lines owned by different suppliers.[[1]](#footnote-1)

In addition to its standard natural monopoly characteristics, there are special features of electricity networks that could lead to significant advantages to incumbents even where the networks expand to new geographical areas. For instance, to bring power to a new customer zone from generators already connected to the incumbent’s transmission system, a potential transmission rival must either:

* build a new line all the way to the existing generators (a high fixed cost) or
* need access to the incumbent’s network. As discussed in chapter 18, the high voltage inter-regional transmission lines (interconnectors) might feasibly provide contestability in some parts of transmission, since each new transmission line in an interconnected system can act as a substitute for other lines. However, there would need to be a regulated requirement for any-to-any connectivity, as in telecommunications, which without access requirements, the incumbent would not be obliged to provide. (In fact, in the absence of regulated technical standards, it might be justified in withholding access given the potential externalities posed by network effects.)

Even were the transmission businesses to negotiate a commercial agreement, non‑cooperatively determined charges for transmission rights could lead to higher prices than for an integrated monopolist. This reflects that each party acting independently could add a mark-up on marginal cost whose overall impact would be higher than the mark-up chosen by a horizontally integrated transmission business — the ‘double marginalisation problem’ (Laffont and Tirole 2000, pp. 184ff). In the United States, this has actually materialised, with customers and suppliers facing double margins on charges as their power crossed a utility’s corporate boundary (Massey 2007, p. 6).

The natural monopoly characteristics of networks could also allow a network business to foreclose supply upstream (generation) and downstream (retailing), so that vertical integration of these activities would occur. However, structural separation of most monopoly networks has occurred through privatisation, new entry and the creation of separate state-owned corporations in the contestable part of the market, so this particular risk has vanished. Nevertheless, there remain some concerns about:

* the integration of generators and retailers — ‘gentailers’. Whether this is genuinely a problem is not considered in this inquiry
* the market power that can be wielded by some incumbent generators (a matter discussed in chapter 19), which also has implications for the efficient use of interconnectors
* the potential for common share ownership of generation and network businesses, (albeit in separated entities) to influence decision making. Even though the Queensland and New South Wales Governments have structurally separated their network and other electricity businesses, they still own a significant share of generation capacity in their states (AER 2011b, p. 29). In a private situation, the Australian Competition and Consumer Commission would not approve (a transaction that led to) common shareholding of businesses that would be likely to ‘substantially lessen competition’ (increasing the ability to wield market power) in a market (s. 50 of the *Competition and Consumer Act 2010* (Cwlth)). It is unclear whether common government ownership of generation and network assets would necessarily have such an effect.

The above features of electricity networks need not eliminate the scope for some competition. The relevant market should not be defined as the technology for transporting power, but by the responsiveness of customers to increases in electricity prices, which in turn, will depend on the extent to which customers can choose alternatives. For example, gas already provides some competition as an energy source for cooking, and water and household heating.

In the future, standalone distributed generation — such as rooftop photovoltaic cells or small gas-powered generators — may provide some competition to the network. However, the prospect of widespread competition of this form is not imminent. Even where they do exist, most distributed generators are still typically linked to the distribution network. This enables customers to feed their excess power into the grid, and access power at times when their generators have failed or cannot meet peak capacity requirements (chapter 13).[[2]](#footnote-2) Accordingly, their network savings will more likely apply to transmission. At least for inter-regional transmission lines, pipelines that transport gas to generators, which then feed power into regional transmission networks, is a more realistic source of competition.

For the immediate future, it is likely that without regulation, electricity networks could exercise substantial and enduring market power. (In this respect, they are not like innovating businesses that create momentary rents that competing innovative rivals then bid away — as in many electronic products or drugs.)

## 3.2 Evidence about the costs of market power

The *a priori* case for at least some regulation of businesses with enduring monopoly power in key goods and services is so strong that, at least in advanced economies, governments typically regulate them. Therefore, it is hard to find much detailed evidence about the actual behaviour of such businesses under a counterfactual when that regulation is absent. Some of the most notable exceptions are distantly historical — such as the US Standard Oil Trust, United States railroads in the 19th century, and salt monopolies in 17th century France — all of which engaged in the price gouging anticipated for monopolies.[[3]](#footnote-3) Gas distribution in the United States was initially unregulated, until prices rose to high levels after the formation of a monopoly gas trust (Troesken 2006). Private water companies appear to have set inefficiently high prices and delivered dangerously low-quality water to United States citizens in the late 19th century (Troesken 2006, pp. 274‑5).

The main, and hardly unsurprising, lessons from unregulated monopolies wielding market power are that they charged high average prices and that they spent significant resources trying to maintain that power.[[4]](#footnote-4) That provides little empirical guidance about the magnitude or source of welfare costs that might exist for an unregulated electricity network in contemporary Australia. It is likely that an unregulated essential service monopoly today would be far more sophisticated — for example, in its pricing strategies and in its bargaining with different groups of customers.

The problems in gathering evidence about the counterfactual have several implications.

Arguments about the nature and magnitude of the costs of the market power of electricity network businesses (and guidance on the best counteracting forms of regulation) will draw more strongly on theories and ‘reasonable’ assumptions than in many other areas of economics.

As a result, it is hard to rigorously test rival theories about the exact nature of the costs of natural monopoly, and consequently, the best type and benefits of regulation. The situation is akin to observing a patient’s health after receiving a treatment, but with less information about their health status before treatment. There is, however, the possibility of testing competing theories based on internal consistency, empirical evidence about behavioural relationships that hold more generally (for example, that businesses with high fixed costs often use two-part tariffs for efficient recovery of those costs) and based on the observation of partly regulated monopolies and unregulated businesses with transient market power. It is also possible to test rival theories about the impact of alternative regulatory regimes on efficiency and prices (a form of benchmarking).

Given the magnitude of the potential economic and social effects of different degrees of regulation, it is critical to separate the wheat from the chaff, else the community could become the victim of the ingenuity of particularly persuasive theorists. The last 150 years of competition theory demonstrates that alternative theories often reach radically different policy conclusions about the appropriate reach of regulation (none in some cases, to complete regulatory oversight or public ownership in others). No theory has proved entirely satisfactory, and some have appeared to excuse anti-competitive conduct (Crew and Kleindorfer 2004; Melody 2003). As Melody put it acerbically:

… the essential criteria around which public utility regulation has revolved are the reasonableness of prices and the universality of service coverage. It is on this skeleton regulatory framework that a little meat and an enormous amount of fat has been hung during the last century. (2003, p. 7)

Finally, there is enough uncertainty about the effects of (unregulated) natural monopolies that regulators also face uncertainty about the degree to which they have succeeded in their endeavours. Regulators have to be particularly mindful that they have less market and engineering expertise than the businesses they regulate.

## 3.3 The case for regulating monopolies

Putting aside foreclosure risks, the usual basis on which governments regulate natural monopolies are concerns about the impacts of higher prices and lower output. In the case of electricity networks, lower output may take the form of reduced quality and not just less network capacity.

There is no single adverse consequence from these outcomes. Over several hundred years, economists and legal experts have identified multiple channels by which these outcomes can affect people’s welfare. These multiple channels can represent a challenge to regulators because some types of regulatory responses may address one channel, but not another, in which case regulation may need to involve tradeoffs.

### (i) Consumer preferences are not fully and efficiently met

If a network business were to charge a price to consumers that exceeded the full cost to the business of transporting power, then customers forgo some consumption that was valuable to them, and that the business could have economically supplied. (A similar outcome might occur if the network business skimped on quality.) The immediate customers of a network business may often be other businesses. High electricity prices could have significant effects on their output and investment, especially if they compete on global markets. But the long-run effects are experienced by households.

### (ii) Production inefficiencies

Unless it has perfect price discriminating capacity, monopolists produce less output than is efficient. The inputs that the business would have used to produce that forgone output would be allocated elsewhere in the economy, but to lower-value uses.

The inefficiencies resulting from (i) and (ii) are the most commonly cited conventional sources of welfare losses from monopoly power. These inefficiencies can be static or dynamic. For example, the latter would include adverse effects on investment or on the drive for innovation or cost minimisation by a customer or upstream supplier of an unregulated monopolist.

Since they are well understood, this chapter devotes little space to (i) and (ii), but brevity should not imply that they are not significant (especially once their long-run effects are considered).

### (iii) ‘X-inefficiency’ or business underperformance

Regardless of the degree of competition in a market, it is in the interests of shareholders to minimise costs to maximise profits.

However, there are several theoretical reasons, backed by empirical evidence, that suggest that persistently low levels of competition may weaken the internal *managerial* incentives for cost minimisation and innovation — creating so-called x‑inefficiency.[[5]](#footnote-5) This does not contradict the shareholder goal of profit maximisation.

For example, employees and managers may (rationally) work less hard, knowing that shareholders find it hard to observe their lack of effort (so-called ‘principal-agent’ problems) and that the business will still survive without strenuous effort because of the absence of competitors or, in the case of state-owned corporations, the risk of takeovers. Some of the signals available to shareholders about the performance of enterprises in workably competitive markets are weaker in the case of natural monopolies:

* it is harder to establish yardsticks against which to compare performance, whereas in workably competitive industries, comparative prices and product innovation of rival businesses are more easily observable
* persistently poor managerial performance does not result in lost market share (preserving 100 per cent of the market in a perfect natural monopoly) nor in losses, since a monopolist can apply a margin to its inflated costs and still earn a positive return.

While the term x‑inefficiency originated with Leibenstein (1966), the nexus between the absence of competition and managerial underperformance has been persistently observed since economics began as a discipline. Nearly 250 years ago, Adam Smith (1776) observed, ‘monopoly … is a great enemy to good management’, while 80 years ago Hicks (1935) argued that the ‘best of all monopoly profits is a quiet life’. Australia’s principal competition regulator, the Australian Competition and Consumer Commission (ACCC), has recently echoed this sentiment, observing that market power can be manifested as a ‘lazy monopolist with a quiet life’ (Pearson 2011, p. 4).

There is substantial empirical evidence from 30 years of international studies that x‑inefficiency is a major source of economic waste in industries protected from fierce competition — particularly essential services. These inefficiencies are potentially larger than the allocative efficiency losses described in (i) and (ii).[[6]](#footnote-6) There is a parallel literature in trade policy providing further empirical support for the debilitating efficiency effects of weaker competition (in this case arising from barriers to import competition).[[7]](#footnote-7)

As an illustration of its potential importance, empirical research found x‑inefficiency of around 20 per cent in the Australian electricity industry prior to microeconomic reform, though this may partly reflect state ownership of vertically integrated electricity businesses, as well as statutorily guaranteed market power (Whiteman 1998). Mountain and Littlechild (2010) and Mountain (2011) claim even more profound inefficiencies for both state-owned and private network businesses.

One of the dilemmas of regulation of natural monopolies is to ensure that it does not create x‑inefficiency in its own right.[[8]](#footnote-8) Conventional rate-of-return regulations — which limit the rate of return of a regulated monopoly — reduce rents, but create incentives for inefficiency by encouraging cost padding. Indeed, a strong motivation for benchmarking of regulated monopolies is to identify and penalise x‑inefficiency (and not just the targeting of monopoly rents). The regulator may use benchmarking results as one source of evidence about the acceptably efficient operation of business to help determine regulated prices or revenues. Incentive regulation is based on rewarding businesses using the performance of a leading practice firm close to (but not usually at) the efficient frontier as the benchmark. In that sense, while the avoidance of x‑inefficiency may (partly) motivate regulation in the first place, it also partly motivates the regulator’s choice of regulation.

### (iv) Rent seeking

Given the monopoly ‘rents’ achieved through transfers from consumers to businesses, well-run monopolies have the potential to earn high rates of return on assets. This can prompt wasteful lobbying and other ‘rent seeking’ — a term coined by the World Bank Chief economist, Anne Kreuger — to preserve monopoly power.

In theory, rent seeking (and its protection) could dissipate all or, indeed some multiple, of the rents as unproductive activity, though in practice that is unlikely (Congleton et al. 2008). Certainly, academics and others have identified rent seeking in almost all areas of economic activity where governments, through statute or regulation, have the discretion to restrain or increase market power. This is one reason why excessive regulatory discretion can be inefficient. Concerns about such rents apply to areas as disparate as preferential trade agreements (Cutbush 2010), water (Goesch and Hanna 2002), taxi regulations (Moore and Balaker 2006), union market power (Connolly et al. 1986), import protection (Olson 2005; Ville 2007) and pharmacy (Philipson 2003).

The desire to preserve rents can also lead to excessive political power to the monopoly business or other parties, such as unions, that may share the rents. For example, this was apparent in some unregulated railroad monopolies in the United States in the 19th century. The ‘Camden and Amboy’ monopoly railroad acquired sufficient political power that the State of New Jersey became known across the United States as the ‘State of Camden and Amboy’ (Stein 2012). One executive of the railroad famously said ‘he carried the State in his breeches pocket, and meant to keep it there’ (Rutgers University Libraries 2011).

However, the regulation of natural monopolies does not eliminate rent seeking since the monopoly business and its customers have strong incentives to influence the regulator to re-cast the regulatory contract to their own advantage. Indeed, sometimes incumbent utilities have lobbied government for the creation of a statutory monopoly, arguing that the quid pro quo of economic regulation is the guarantee of insulation from competitive entry (a strategy used by the giant United States telco, AT&T, in the early 20th century — Melody 2003). Nonetheless, to the extent that a regulator significantly reduces rents, the resources that competing parties are willing to use to secure the remaining prize are also diminished.

The relevance of rent seeking for competition policy is the need to have a credible commitment to long-run constraints on monopoly power (without excessive discretion). As for other major natural monopolies, in electricity networks this is best achieved through clear regulatory objectives and statutes, a politically independent regulator (such as the Australian Energy Regulator or ‘AER’) and institutional arrangements that give customers sufficient visibility and negotiating power to counter the monopoly businesses (chapter 21). Biggar (2011b, pp. 18‑19) and Littlechild (2009) have given particular emphasis to the latter. The Commission agrees with this emphasis.

### (v) The distributional effects of consumer to business transfers

Consumers and regulators often raise distributional matters. However, they are much less straightforward than the above issues. Monopoly rents are transfers from consumers to those with a stake in the monopoly business (shareholders and those employees who are able to achieve the ‘quiet’ life or can negotiate higher wages). Even where the initial impacts of monopoly are higher prices for downstream business customers or lower prices for upstream suppliers, the costs are ultimately likely to be borne by households.

Such ‘consumer-to-business’ impacts *may* matter in several ways. Society may value consumer welfare more than the welfare of shareholders or the business’s employees (Armstrong and Sappington 2006; Sims 2012a). That might occur because the shareholders that have benefited from the rents are foreign so that their welfare is not relevant to maximising the welfare of Australians. Or, it may be because, as a group, the relevant consumers facing higher prices have lower income than shareholders. That presupposes that the marginal benefit to lower-income consumers of an additional dollar is higher than for shareholders. While this appears to be a reasonable claim, there is significant debate about when that should be considered in policy analysis (PC 2011a, pp. 955‑956).

Distributional issues raise questions of fact and careful interpretation.

* If foreign shareholders purchased the monopoly business at a price that already fully capitalised the rents, then the preceding owners acquired the rents, not foreigners (a point made by Posner 2001). In some cases — and this is true for all Australian electricity networks and many other businesses with market power in Australia — state and territory governments were the original owners. Accordingly, it is possible that citizens acquired the stream of any future rents as a lump sum from the asset sale at privatisation. (Whether that is true depends on whether the sale process effectively captured all rents.)
* Monopolies are often state-owned. If the monopolies wield market power, the transfers are from consumers to citizens generally (either through reduced taxes or additional services). Whether there are any adverse distributional outcomes depends on judgments about the capacity of governments, through the political process, to make wise distributional decisions.
* While there may still be distributional effects, they may be mitigated by diversified share ownership. For example, superannuation funds, which hold many of the retirement savings of Australian workers, sometimes have significant holdings in businesses with market power. In theory, the implicit taxes that monopoly imposes on consumers may be less for lower income consumers (depending on how the monopoly business sets its tariffs). In that instance, it is possible that transfers from consumers to shareholders have distributional benefits. Whether that occurs is a question of empirics, not principle. Currently, Australian private electricity networks are largely foreign-owned.
* There is little question that society values some re-distribution of income from higher-income to low-income people — as revealed by the tax and transfer system, and by people’s private voluntary re-distribution of resources. However, the overall level of re-distribution in a society depends on the cumulative effects of the myriad ways in which re-distribution occurs. There is a point after which people see further re-distribution as undesirable. Accordingly, without knowing this bigger picture, it is unclear whether any single policy with re-distributive impacts is optimal. This aggregate perspective is overlooked in populist views about the desirability of achieving re-distribution in every instance where a policy has income or wealth effects.

The policy implication of the latter point is that the desirability of different re-distributive policies depends on their relative efficiency costs. It is possible to be strongly in favour of re-distribution of income in a society, but to be doggedly opposed on efficiency grounds to a particular policy that attempts to achieve that.

A possible argument for curbing any re-distributive effects of unregulated monopoly is that passing monopoly rents back to consumers resembles the outcomes of an efficient rent tax (Armstrong and Sappington 2006; and more broadly, Laffont 2005).[[9]](#footnote-9) *If* correctly calibrated so as not to take too much from the monopolist, rent taxes should not raise relative prices or reduce production and investment incentives.

In achieving a desirable level of re-distribution in an economy, rent taxes could displace other inefficient taxes. Another way of looking at this is to imagine the political calls for greater family transfer payments, higher pensions and allowances, and more utility assistance to relieve cost-of-living pressures were essential services to be priced at monopoly levels. Any such fiscal measures would involve new taxes to fund these outlays, with their associated inefficiencies. The excess marginal burden of taxes vary considerably,[[10]](#footnote-10) but the most commonly applied taxes for gathering new revenue are income taxes. The accompanying technical paper to the Henry Tax Review found that labour income taxes, payroll taxes and corporate income taxes had marginal excess burdens of 24, 41 and 40 per cent respectively, while municipal rates (a rent tax) had a burden of 2 per cent (KPMG Econtech 2010, p. 44).

Consequently, there may be *efficiency* grounds for removing genuine rents through the regulation of natural monopolies and passing these back to customers as lower prices. The main drawback to this observation is that an implicit monopoly rent tax distributed in proportion to consumers’ use of electricity is not equivalent to the outcome were the tax to be collected by government and then used to reduce debt or distortionary taxes. Nevertheless, there may be some efficiency benefits.

Regardless, addressing the conventional efficiency costs of monopoly will limit transfers from consumers to a monopoly business. As noted by Banks (2012, p. 12), the focus of National Competition Policy ‘on promoting efficiency should not be seen as contrary to distributional goals.’

This is consistent with the National Electricity Objective (NEO) (as described in chapter 1), and particularly its emphasis on promoting ‘efficient investment in, and efficient operation and use of, electricity services for the long-term interests of consumers of electricity’. The NEO captures the central rationales for the regulation of monopoly businesses:

* efficient production by network businesses (static and dynamic)
* the efficient use of network services by customers (both households directly purchasing electricity and downstream business users that ultimately supply goods to customers)
* the necessity of ensuring efficient investment and other required expenditure needed to ensure long-term reliable supply.

While it would also imply a preference for transfers to favour consumers where efficiency was not at stake, the NEO can be seen as fundamentally an efficiency objective.[[11]](#footnote-11) As noted in the second reading speech for the Bill containing the NEO:

The market objective is an economic concept and should be interpreted as such. For example, investment in and use of electricity services will be efficient when services are supplied in the long run at least cost, resources including infrastructure are used to deliver the greatest possible benefit and there is innovation and investment in response to changes in consumer needs and productive opportunities. The long term interest of consumers of electricity requires the economic welfare of consumers, over the long term, to be maximised. If the National Electricity Market is efficient in an economic sense the long term economic interests of consumers in respect of price, quality, reliability, safety and security of electricity services will be maximised. (Second reading speech, *National Electricity (South Australia) (New National Electricity Law) Amendment Bill 2005*)

In their stage one interim report on the limited merits review regime, the panel (Yarrow et al. 2012b, p. 40) noted that the NEO has the advantage that it is both clear and emulates the outcomes of effectively competitive markets:

The primacy of the long term interests of customers as an evaluation criterion, set out in the NEO and the NGO, gives the conduct of regulation the same focus as that of the supply-side of an effectively competitive market (how can we improve the consumer offering?). This is admirably clear, and avoids the confusions of multiple, conflicting objectives that have had adverse effects in jurisdictions such as Great Britain.

#### However, ‘equity’ is a problematic goal

A quite distinct distributional concern relates to the desirability or otherwise of reducing the costs of essential services for *specific* groups of consumers (regional dwellers, the old, the disadvantaged) financed through higher regulated charges for other customers, rather than financed by taxpayers generally. This is generally not a persuasive rationale for *competition* regulation. However, social concerns provide a rationale for other measures, such as hardship measures implemented by the businesses,[[12]](#footnote-12) and sometimes financial assistance to particular groups.

The reason such policies often accompany competition regulation is that:

* once a government regulates to address the market power of an essential service, it is easy to add other regulations with social (and environmental) goals
* a major reason that essential services have market power is because people cannot take or leave the services they offer — these are services critical to the everyday lives of all people. This characteristic also means that governments see it is as highly desirable that essential services are universally available and affordable for all people, regardless of their circumstances.

Whether such re-distributional policies should attach themselves to economic regulation depends on the desirability of such subsidies and on the relative efficiencies of the alternatives for financing them. Either way, they should not be an objective of the *competition* regulator. Budget-funded measures — such as utility allowances — have the advantage of transparency, clear parliamentary accountability and can use the same (targeted) eligibility criteria for other distributional policies to create a more coherent framework. In contrast, it can be hard to target price regulation for social welfare purposes (ERAA 2012a, p. 4).

However, where social goals are achieved through regulatory means, they should be:

* targeted as best as possible and be as consistent as possible with budget-funded social welfare measures
* imposed as an explicit government directive so that it is clear who is responsible for the policy, but leaving some flexibility in how they should be met. The danger is that prescribing particular tariff arrangements to meet equity goals might undermine the main objectives of the regulator to set prices as efficiently as possible, with unintended long-run impacts for consumers generally
* accountable and transparent, with full disclosure of their revenue equivalents, to avert another source of rent seeking.

This issue is relevant to this inquiry in that it may affect the appropriate ‘efficiency’ benchmarks in incentive regulations for businesses subject to extraneous social goals, and if poorly undertaken, could undermine some approaches for achieving network cost-savings through demand management (chapters 9–12).

### (vi) Unfairness and ‘injustice’

‘Unfairness’ has several dimensions. One is the distributional issue described above. The other concerns justice. For example, a common example of this aspect of fairness is the desirability of procedural fairness.

While little empirical work on this issue has been undertaken in Australia, the history of anti-trust regulation in the United States demonstrates that addressing unfairness is a key objective of utility regulation, with requirements for prices to be ‘just and reasonable’ (Beecher 2010). A survey of the 207 state and federal utility regulators in 2001 found that the majority of commissioners considered fairness to be a central and often more important goal of utility regulation than efficiency (Jones and Mann 2001). Their view of unfairness extended beyond unreasonable prices, suggesting that the distribution of transfers were not the only consideration.

One specific case is the Michigan Public Service Commission, which lists, among other factors in determining utility prices, the desirability of fairly apportioning costs among consumers, fairness to both the regulated utility and ratepayers, and avoidance of unjust or undue discrimination between rate classes or consumers.[[13]](#footnote-13)

Likewise, competition law in the European Union gives strong weight to fairness. Abuse of a dominant position may consist in ‘imposing unfair purchase or selling prices’ (Akman and Garrod 2010).

In Australia, the *Competition and Consumer Act 2010* (Cwlth) also places significant weight on prohibiting unfair practices, but does not refer explicitly to unfair prices. The competition and consumer regulator, the ACCC, has some (limited) capacity to oversight prices under Part VIIA of the Act. However, the ACCC has on occasions characterised its price monitoring role in terms of fairness:

A price inquiry is conducted to determine whether buyers are getting a fair deal in the supply of goods and services. (2005, p. 8)

While unfairness is hard to define, people clearly identify it as an important feature of transactions. As scholars in this field have noted:

While the idea of fairness is elusive and perceptions differ, it clearly is a potent force in regulation, as indicated by the vehemence with which participants complain when they feel they have been treated unfairly. (Jones and Mann 2001, p. 153)

… few persons give much thought to what is fair. But they know when they have been treated *unfairly*; perceived unfair treatment is what makes people shout, “I have been had (screwed, taken to the cleaners, etc.)! … the sense of unfair treatment typically comes from a perception that a contract, explicit or implicit, has been broken. (Zajac 1996, p. 117)

People regard transfers made under circumstances of duress or with significant unequal bargaining power as unjust, even if they do not entail significant distributional consequences. For example:

* even if the perpetrator is disadvantaged and the victim not, most people would regard the theft of money as an undesirable and unfair form of re-distribution
* survey research has shown that most people regard the use of market power to raise prices as unfair, and that the degree to which prices are raised does not affect their judgment of unfairness by much (Kahneman et al. 1986, p. 735). The present chair of the ACCC (Sims 2012a) characterises public attitudes to monopoly pricing in fairness terms (although he suggests that the best rationale for regulation lies elsewhere)
* in cases of costly litigation, people are sometimes willing to use significant resources to pursue matters where the economic losses are small, revealing the value they place on correcting what they perceive to be injustice and unfair treatment
* in its inquiry into consumer policy, the Commission found people had an aversion to ‘unfair’ contracts that related to far more than their transfer effects.

The desire to enforce social norms against the unfairness or injustice that sometimes accompanies unequal bargaining provides an additional rationale for the regulation of natural monopolies wielding market power.

However, in practice, unfairness and injustice present difficult challenges for the price regulation of utilities, in contrast with the clear relevance of these concepts to misleading and deceptive conduct, anticompetitive actions, fraud and other forms of misconduct generally prohibited by, and redressed through, general trade practices and the common law. It is notable that the test of the misuse of market power in s. 46 of the Competition and Consumer Act requires that a business uses its power for an illegal purpose — not that it sets prices too high. Indeed, one of the targets of s. 46 is exactly the opposite — the strategic use by a business to strangle competitors by initially setting prices too low (‘predatory’ pricing).

The fact that some form of overpricing is widely seen as unfair indicates that some regulation may be warranted, but does not provide any analytical tools for calculating a ‘fair’ price. As with the issue of transfers to consumers, pursuing efficiency objectives will typically achieve the goal of fairness without any need for supplementary policies.

Moreover, judgments of fairness are highly subjective and depend on how a situation is framed (as shown in the many careful survey questions posed by Kahneman et al. 1986). Even where a monopoly business’s prices are heavily regulated, many will still claim that the prices remain unfairly high (a point apparent in submissions made as part of regulatory determinations for electricity revenue and price caps).

### Deadweight costs as an umbrella concept

The costs represented by (i) to (iv) are the standard ‘textbook’ economic efficiency losses of monopoly, with a parentage stretching back at least 50 years. Sometimes these costs are categorised collectively as ‘deadweight’ costs.[[14]](#footnote-14) Even (v) can be categorised as a deadweight cost, since its relief though regulation may also yield potentially substantial economic efficiencies.[[15]](#footnote-15)

There is little evidence about the magnitude of the above costs were electricity networks to be unregulated. However, some of the important aspects of the story are:

* electricity is an ‘essential’ service (a basic service that few households can do without)
* electricity is used by every business and government (chapter 2). Any monopoly price effects cascade throughout the economy, before being ultimately borne by households (domestic and foreign). Accordingly, the price of a soft drink to the consumer will, among other things, reflect the impacts of excess electricity prices on the glass used to manufacture the bottle, the manufacturing process used to produce the drink and bottle it, the cost of power to make the bitumen and concrete for the roads that transport the drink, and for the power costs of wholesalers and retailers
* the distortionary impacts of monopolies depends on the responsiveness of customers to higher prices (‘demand elasticities’), which may vary by customer type. The existing evidence suggests relatively low elasticities of electricity demand for households (Fan and Hyndman 2010 and AEMO 2012a, pp. A.1‑A.5). In part, this is because it is an essential service. Since conventional deadweight losses are a function of such demand elasticities, this suggests relatively small *static* consumption-side deadweight losses. However, the existing demand studies relate to an industry that is already highly regulated. Demand elasticities may well be greater for an unregulated monopoly, and so too, the degree of inefficiency.[[16]](#footnote-16) Moreover, long-run price elasticities are typically higher than short-run elasticities. This reflects many factors. For example, with high prices, technological developments would favour energy-saving equipment and dwellings. While people might not immediately replace energy-intensive technologies, they would eventually do so. Distributed generation would become more attractive. Accordingly, the long-run distortionary impacts of monopoly are likely to be significantly larger than the static costs inferred from short-run price elasticities applying to already regulated monopolies
* even small increases in electricity prices can entail significant losses for some commercial customers that have electricity as a major input into their outputs and that sell in competitive global markets. That could have large long-run effects on investment by these customers, even if does not substantially affect current output. (The long-run effect can be much larger because over the short run, the assets in such businesses are sunk.) This was the primary reason for offsetting the effects of carbon pricing on trade-exposed emissions intensive industries
* to the extent that demand is not responsive for any given group, transfers from customers to businesses are large for that group. This has several implications. First, it means that the distributional issues raised in (v) above may become relevant. Second, since demand is least responsive when the monopoly business faces few risks of entry by others, it suggests that x‑inefficiency may be high. Third, the prize from preserving monopoly rents through lobbying and the political process is large, encouraging wasteful resource use.

## 3.4 Are deadweight losses passé? New theories of why monopolies should be regulated

The effectiveness and role of benchmarking as a regulatory tool depend critically on the underlying problems that regulations are intended to resolve. Recently, some prominent Australian economists have disputed the relevance of deadweight costs to competition policy, indicating that most economists (including the Productivity Commission) poorly understand the issue:

The fundamental rationale for public utility regulation has not been well understood, particularly by economists. Mainstream neoclassical economists have argued that the primary rationale for regulation is the minimisation of [the] so-called ‘deadweight loss’. But, on close inspection, this hypothesis does not fit the observed facts. (Biggar 2011b, p. 6)[[17]](#footnote-17)

The alternative perspective is that the rationale for price regulation is to create an implicit contract that protects customers’ investments from expropriation by the monopolist (a hypothesis explained further below). If this alternative is the correct rationale, this could constitute a major break from the primary expressed basis for regulation of essential services by many international competition authorities.[[18]](#footnote-18) New theories abound in economics and most are not influential. However, when they originate from Australia’s premier competition regulator, and have implications for the way it might regulate or advocate policy change, these ideas should be listened to carefully.

In that context, Biggar claims that the contractual approach leads to the imperative for a quite different[[19]](#footnote-19) regulatory framework from the current version. He proposes that:

* consumers should play a bigger role because they are a key party to the ‘contract’ and more informed about their preferences than the regulator
* the regulator should be an independent arbitrator, not a consumer advocate as its role is to construct the right contract between a monopoly supplier and customers. Accordingly, there would be no grounds for appeals on merit because, as an arbitrator, the regulator would have already fairly addressed both parties’ concerns. Appeals should be limited to matters of law, which would also allow the regulator to exercise significant discretion and judgment, rather than to use significant resources ‘covering its back’ on every issue, given the risk of merits review
* the Rules should not require that the regulator apply the highest cost that just passes a reasonableness criterion, rather than using its best cost estimate, since this favours one party to the implicit contract. Instead, the regulated entity should prove that the regulator’s price/revenue proposal is unreasonable
* the object clause in the National Electricity Law should be changed to include long-term price stability, non-discrimination, and a ‘no disadvantage’ test, so that no customer could be made worse off following a regulatory change (such as those that would arise from a tariff change — for instance, peak load pricing).

These issues are central to aspects of this report, affecting the potential role of benchmarking in incentive regulation and the scope for increasing network efficiency by removing some of the regulatory constraints network businesses face. Given the above policy conclusions partly depend on the credibility of the new theory, this chapter briefly examines the validity and insights of the theory that led to these conclusions (with a detailed discussion in appendix B).

### The new theory: protecting the sunk assets of customers

Biggar’s alternative view of the rationale for regulation of monopolies is based on the problem that customers can be held to ransom once they have made sunk investments that use the inputs — such as electricity — of a natural monopoly. A sunk cost is an investment that, once made, cannot easily be recovered by selling it to another party. Some clear examples include the special dies made by a supplier of automotive panels,[[20]](#footnote-20) a person’s investment in a qualification, or a specialised product for a particular customer. Not all fixed assets are sunk. For example, as they are usually leased, aircraft are called ‘capital with wings’. Nevertheless, many saleable assets would still recover only a proportion of their value when sold — and so are partly sunk. ‘Sunkness’ is therefore ubiquitous.

Biggar’s claim is that:

… the ***primary*** economic rationale for public utility regulation is the protection of the sunk (relationship-specific) investment of customers of the regulated firm. (2011b, p. 30) (emphasis added)

The logic underpinning this is that:

* customers of the monopolist make sunk investments whose value is dependent on continued supply of the monopolist’s inputs (a ‘relationship-specific’ investment in the language of this literature)
* customers know that after they have made the investment, the monopolist can raise prices significantly because, once the investment is sunk, the customer only needs revenue sufficient to cover the marginal costs of supply. The monopolist can therefore expropriate all of the returns that would otherwise have made the investment profitable. Prices would be unstable, varying before and after the investment was made
* customers are unable to write *binding* long-term contracts with the monopoly supplier (but regulators can do so on their behalf)
* vertical integration to internalise the long-term contract within a single firm is often not possible
* As customers know that the monopolist can behave opportunistically after the investment is made, they do not make the investment at all (the ‘hold-up’ problem).

Under the hold-up problem, the customer knows that the unregulated monopolist has incentives to expropriate the value of a customer’s relationship-specific sunk investments, and the monopolist knows that the customer knows this. Accordingly, the customer will be unwilling to invest (or if so, to a less than optimal level). A customer’s decision not to invest reduces the revenue of the monopolist. So, the key hold-up problem is not that the monopolist *will* expropriate the sunk investments of a supplier, but that it must credibly commit *not* to do so to preserve its long-run monopoly profits (which it earns as a mark-up on long-run prices). Biggar suggests that regulation solves the problem because it re-creates the long-term contract that parties would have negotiated to protect each party from later opportunistic behaviour.

Moreover, he argues that the theory’s validity is strengthened because only it explains why regulators behave as they do. He claims that it explains, among other things, why they apparently:

* encourage price stability to ensure ex ante and ex post prices are the same
* dislike price discrimination because ex post a customer with a large sunk asset has a very low elasticity of demand, and so would face a high price. This would appear to be efficient as a means of recovering the fixed costs of the monopolist, but would do so at the cost of undermining the investment incentives of customers
* do not generally accept peak load pricing, despite its benefits for reducing deadweight losses
* want monopolies only to cover their full costs and not additional rents (because the rents would most likely arise from expropriated sunk investments).

### Is the theory compelling as the primary rationale for regulation?

As discussed in appendix B, there are attractions to conceptualising the regulation of natural monopolies in contractual terms. The broad implications of the ‘contract’ approach appear to echo those resulting from maximising community welfare. That is not surprising given that the only contract that two parties with equal bargaining power would mutually agree to would be one that involved no removable inefficiencies. The real difficulties arise when the contract approach gives primacy to the ‘sunk investment’ problem.

Without repeating the detail in appendix B, the main flaws in the ‘hold-up’ argument are that:

* there does not need to be a single meta theory of regulation of natural monopolies. Hold-up may sometimes be relevant, but in other instances, the concern will simply be that, absent regulation, the consumer would be exposed to excessive prices and the monopolist might live an inefficiently ‘quiet’ life
* most users of electricity do not make the large irreversible investments that underpin the hold-up problem. Moreover, from a pragmatic perspective, it would be very costly for an unregulated monopolist to behave opportunistically in the way supposed by the hold-up problem for millions of customers
* even those business customers making very large investments would often be able to write long-term contracts that would preclude the hold-up problem. Moreover, other non-regulatory solutions like vertical integration could be used where contractual difficulties appear insurmountable
* it is also not clear that even were a dominant business to ex post exploit a customer making sunk investments that the solution would be price regulation. Competition laws (as in Article 82 the European Union Competition Law) or oversight of particular long-term contracts that may be subject to hold-up might address any issue in a more targeted way
* the notion that regulations (and their underpinning laws) create credible long‑term contracts belies how regulations change and how regulators behave. By early 2013, there had been 55 versions of the National Electricity Rules. Over an asset life of 40 years (easily possible for network assets), the economic value of the asset would be exposed to nearly 300 sets of possible new regulatory influences during its life. The hold-up problem may be larger under regulation than under unregulated monopoly, which is why the National Electricity Law should be designed to reduce that risk
* regulators do not, in fact, say that their actions are based on the hold-up problem — something that would be expected were it to be their true underlying motive for the regulation of natural monopolies. Indeed, they typically refer to conventional arguments about the problems of natural monopoly. Regardless, the idea that a good rationale is revealed by the actions of regulators presumes those actions are always appropriate. The ultimate test is not what regulators do, but whether, given market behaviour and its outcomes, regulation is justified on welfare grounds.

## 3.5 The alternative policy implications of different theories of monopoly regulation

As noted in section 3.4, the proponents of hold-up reach strong policy conclusions, with significant implications for this inquiry. However, those conclusions are not convincing.

The conclusions could arise from alternative theories. For example, the importance of empowering consumers and, where possible, involving them in regulatory decisions is consistent with conventional theories of competition regulation. This approach can sometimes overcome asymmetries of information for regulators, achieve ‘buy in’ by consumer groups that recognise the need for compromise in regulation, and may allow less heavy-handed regulation (where supported by the potential for intervention by the regulator if negotiations are imbalanced). The Commission is in strong agreement with Biggar’s (and the ACCC’s) views on the importance of a consumer role. The difference is only in the reasoning to reach that conclusion.

Often the policy conclusions are non-sequiturs. It is not clear why the existence of hold-up would justify a no-disadvantage test. The implicit contract struck by the regulator on behalf of a large heterogeneous group of customers and a monopoly business would usually not advantage all consumers. For instance, a prohibition on peak pricing would disadvantage those consumers who use less power in peaky periods (who will often be poorer customers)

Similarly, the fact that the AER should be an impartial arbitrator rather than a consumer advocate — a position the Commission supports — does not preclude significant regulatory errors. Customers or businesses should be able to contest any material errors. The proposed reversal of the onus of proof so that businesses would be obliged to prove the unreasonableness of a determination by the AER would need to be justified on other grounds. This inquiry finds that benchmarking is a useful but inexact tool (chapters 4 and 8). The existence of merits review creates strong incentives for the regulator to be prudent in its application.

The conclusions also fail to consider the long-run interests of customers as a group by taking a narrow approach to the implications of regulatory error on investments. Where a monopoly business provides inputs to many diverse customers — as in electricity networks — opportunistic behaviour by a monopolist can only affect a small share of customers (as shown in appendix B), whereas regulatory error by the regulator can affect the investment adequacy of the monopoly business to the disadvantage of all of its customers. Moreover, the intention of incentive regulation is to offer some rents (‘headroom’) as the carrot for a regulated business to minimise its costs and to innovate. Setting the ex ante regulated price at a point equal to that of some hypothetically most efficient business does not allow such headroom, and would be likely to reduce dynamic efficiency — with losses to consumers.

## 3.6 In summary

This chapter has devoted considerable space to the issue of hold-up because it:

* appears to have gained purchase in some regulatory circles, albeit not uniformly in either its rhetoric or practices
* is a useful thought experiment, whose flaws highlight many of the practical necessities of competition regulation (such as merits review and compromise).

The hold-up problem is a hypothetical and theoretically elegant example of a particular deadweight loss that might sometimes arise were monopolies to be unregulated. However, it is not persuasive as ‘the’ reason for economic regulation of electricity network businesses, and nor have energy users, the AEMC, AEMO and other major stakeholders perceived market power in this way.

The hold-up idea does not displace the usual efficiency concerns of regulators, governments and economists, and is not required as a basis for the sensible regulation of network businesses. Ironically, were it to actually inform regulatory policy to any great extent, the hold-up construct has the potential to undermine the long-term interests of end users, as it could threaten investment by the monopolist in long-term assets, innovation and cost minimisation. It does not appear to be compatible with incentive regulation (and one possible use of benchmarking). It would be unfortunate were it to assume primacy as the conceptual basis for Australia’s economic regulation of electricity networks (or indeed other infrastructure).

#### There are points of substantial agreement

The differences of view on this issue do not change one essential point. Electricity customers, the ACCC, the AER and the Commission all strongly argue for the economic regulation of electricity networks, broadly along the lines set out in the recent Rule changes (AEMC 2012r). Unregulated network businesses would have significant market power, with the potential for adverse effects on customers. Correcting that market power requires nuanced but powerful regulation. Subsequent chapters specify the desirable characteristics of such regulations.

1. Though there may be ‘underbuild’ when distribution wires are strung under higher-voltage wires on a sub-transmission line. [↑](#footnote-ref-1)
2. Of course, in some circumstances, distributed generators are not linked to the grid at all — especially in non-metropolitan areas. [↑](#footnote-ref-2)
3. There have been many ostensibly unregulated state-owned utilities, but they were often charged with social goals that meant they were, in effect, partly regulated. Moreover, they did not usually face the same incentives as private firms due to their affiliation with government, and were often less exposed to even the risk of competition due to statutory barriers to entry. In that sense, their conduct may not be representative of an unregulated private monopolist. [↑](#footnote-ref-3)
4. Monopoly per se does not *necessarily* confer significant market power. For example, there may be countervailing power from customers (Oxera 2012). Airlines and airports are one illustration (PC 2011b, pp. 74‑5, 80‑2, 193, 331‑33). Moreover, a monopolist that exploits its market power for most customers may do so to a much lesser extent for a commercial customer that uses the monopolist’s output as an input into the production of a good exposed to very competitive markets. [↑](#footnote-ref-4)
5. However, x‑inefficiency is not all waste. A quiet life is valuable to those enjoying it, and to that extent, x‑inefficiency is a transfer from shareholders to employees (Alchian and Kessell 1962). However, there may still be waste because many employees would prefer to take the rent as wages, but cannot do so because wages are observable to shareholders. [↑](#footnote-ref-5)
6. As discussed by Perelman (2011) and documented by Frantz (2007) for many industries, De Witte and Marques (2010) for water, Brons et al. (2005) for urban transport, and a large number of studies of the electricity sector (described throughout this report). [↑](#footnote-ref-6)
7. See Panagariya (2002) for a survey. [↑](#footnote-ref-7)
8. Indeed, one of the empirical difficulties in assessing the degree to which monopolies create x‑inefficiency is that most natural monopolies are heavily regulated, so that it is hard to disentangle the effects of monopoly and the effects of the regulation of monopolies. [↑](#footnote-ref-8)
9. This is subject to the proviso that the rents are not the outcomes of innovation or risky investments. Taxing such ‘rents’ would reduce incentives for risk-taking. That proviso is not likely to hold for standard investments in electricity networks. [↑](#footnote-ref-9)
10. This measures the deadweight inefficiency costs as a share of each dollar of revenue that government raises. These costs arise because taxes affect decisions to invest and to work. [↑](#footnote-ref-10)
11. The NEO suffers from some ambiguity — a point made in the Commission’s analysis of the similar object clause in telecommunications competition regulation (Part XIC of the then Trade Practices Act); by the Energy Networks Association at the time of the Exposure Draft of the National Electricity Law (ENA 2004); and by Kerin (2012). However, these concerns are largely allayed given the guidance on the interpretation of the NEO in the second reading speech. [↑](#footnote-ref-11)
12. These are measures adopted by the businesses that provide advice on energy use, and allow people to make instalment-based payments and deferred payments. They are not usually subsidies. [↑](#footnote-ref-12)
13. Information provided by Darryl Biggar (pers.comm.). [↑](#footnote-ref-13)
14. Any basic textbook would categorise (i) and (ii) as deadweight costs. Moreover, many economists would argue that x‑inefficiency and rent seeking also result in deadweight losses (for example, Crew and Kleindorfer 2004; Dobson 1992; Johnston and Trembath 2005, p. 24). Some argue that x‑inefficiencies should not be *labelled* as deadweight costs because they arise from non-profit maximising behaviour by the business. However, x‑inefficiency is consistent with profit maximising behaviour subject to principal-agent costs. [↑](#footnote-ref-14)
15. Some also argue that failure to address unfairness in (vi) is economically inefficient because fairness is a ‘good’ included in people’s utility functions, with people willing to give up the consumption of other goods to increase it (such as in altruistic behaviour). This may be technically correct, but enumerating the inefficiency effects of unfairness is difficult, and as such, the point is conceptual rather than practical. [↑](#footnote-ref-15)
16. For example, if a demand curve is linear, an unrestrained (profit-motivated) monopolist will price up to the elastic portion of the demand curve. [↑](#footnote-ref-16)
17. Biggar has also expressed these concerns in a range of other papers (2009, 2010, 2011a). The chair of the ACCC, Sims (2012a) has also raised similar issues. [↑](#footnote-ref-17)
18. The Commission examined the criteria for regulating among various prominent regulators — which all emphasised traditional economic efficiency as a major (if not always the only) motivator for regulation. For example, these included the New Zealand Commerce Commission and major UK regulators (such as Office of the Gas and Electricity Markets, the Water Services Regulation Authority, the Utilities Regulator (of Northern Ireland), and the Financial Services Authority). [↑](#footnote-ref-18)
19. In fact, some elements are already features of the current framework — a point raised later. [↑](#footnote-ref-19)
20. This is the classic example given by Klein et al. (1978). [↑](#footnote-ref-20)