B The hold-up problem

As discussed in chapter 3, prominent Australian economists, most notably Biggar (2011b), suggest that the conventional rationales for regulating natural monopolies are not convincing. Instead, Biggar argues that the 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 of a natural monopoly. A sunk cost is an investment that, once made, cannot easily be recovered by selling it to another party. He argues that the hold-up problem justifies amendments to the National Electricity Rules and the National Electricity Law, including a new objects clause, the removal of merits review, and the greater role for regulatory discretion. He also argues that the hold-up problem explains why regulators behave as they do, such as *apparently* seeking price stability and limiting price discrimination.

Chapter 3 summarised the thrust of the Commission's concerns about this theory. This appendix explains its limitations in more detail.

B.1 Is the theory compelling as the primary rationale for regulation?

The underlying notion that the goal of regulation is to simulate an 'ideal' long-term contract between two parties is appealing. It captures the importance of cost minimisation, provision of information between parties, adequate infrastructure investment, and avoiding excessive prices, but also the desirability of avoiding strategic behaviour by either party (such as rent-seeking and the ex post appropriation of the kind described by Biggar). It also suggests that equity concerns described earlier should not be part of the competition regulatory contract because consumer-to-consumer transfers generally would not be addressed in a contract between consumers as a collective, and a supplier. Accordingly, the broad implications of the 'contract' approach appear to mirror 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 with the above justification for competition policy emerge when it gives primacy to the 'sunk investment' problem.

There does not need to be a single meta-theory

One immediate concern is that there can be more than one rationale for regulation of natural monopolies, even if these fall under a broad umbrella of economic efficiency. Among other factors, competition regulations can be justified on the grounds of allocative inefficiency, x-inefficiency, and foreclosure risks. Sometimes, dynamic inefficiencies, such as hold-up, may also be relevant. There is (some limited) evidence of strategic behaviour of the kind that might lead to investment hold-up. For example, many Californian oil refineries were designed to process a 'sweet' crude oil (an oil with low sulphur content) known as ANS. In the 1990s, ANS prices began to rise. The price charged by British Petroleum to supply ANS to a refinery was higher if the refinery was designed for ANS processing, but lower if the plant could refine other oils (McAfee 2002).[1] However, the fact that hold-up may occur sometimes does not mean that it is very important.

Do small users face hold-up problems?

Electricity network businesses transact with millions of businesses and households over long periods, with many millions of new household purchases in electricity-using appliances each year. The typical lumpy investments for households are appliances — such as computer equipment, televisions, air conditioning, heating and refrigerators. In 2011, there were more than one million purchases of air conditioners, with a stock of more than eight million (EnergyConsult 2010, p. 17). One in four households buy a new television each year (DEWHA 2008, p. 61). Moreover, technological change, innovation and lower prices will almost certainly lead to the greater diffusion of new electricity-using appliances in areas not currently anticipated by consumers. Since households are heterogeneous and new ones are constantly forming, different patterns of consumption can also be expected over time.

Given the above, the typical household will buy some electrical appliances each year and such a pattern can be expected to persist. While there is a second-hand market for these goods, it is likely that much of the value is 'sunk' because of the transactions costs of sales and uncertainty about the quality of second-hand goods. 'Sunkness' is even greater than that implied by individual purchases because people's investments in housing and lifestyle (for example, a willingness to live in a hot or humid environment) are partly based on expected future electricity prices.

1 Even this is weak evidence of actual hold-up, because hold-up is not about the actual expropriation of rents, but about the adverse investment effects of its anticipation by the non-monopoly.

It is helpful to provide a concrete example of what is purported to happen under the hold-up problem, and apply it to a household. The Smiths consider buying an air conditioner on a given date for $3000. However, in making any purchasing decision, the Smiths take into account that the immediate second hand value of a just purchased air conditioner is only $1500 because second-hand markets are imperfect and installed air conditioners are costly to remove. Accordingly, $1500 of the value of the air conditioner would be immediately sunk were the Smiths to purchase it.

When there are many customers, individual contracts cannot be written to avoid the risks of hold-up, and vertical integration is clearly not an option. Accordingly, the Smiths are concerned that a canny network business could use its market power to increase its network charges after their air conditioner purchase to appropriate up to the $1500. Therefore, the sunk cost problem has effectively raised the price of the air conditioner by $1500. At that higher price, the Smiths decide not to purchase the air conditioner — an example of 'hold up'. This is a clearly inefficient outcome for both the Smiths (who want to have air conditioning) and the network business (which wants to transport power).

When put in such concrete terms, this outcome is implausible. Were it to occur, hold-up would rarely involve zero purchases of air conditioners across all households, but simply less than would be economically efficient. (And with rising incomes, the Smiths may well buy an air conditioner at a later time, even with the risk of appropriation of the sunk value of their investment, because their consumer surplus may then exceed the sunk cost.) The same would apply to countless other household investments that involve sunk costs. Consequently, the Smiths and other households would be able to observe whether, in fact, the risk of expropriation of sunk costs actually occurs. It is in the interest of a network business not to expropriate the sunk costs of early purchasers because this would signal that it would expropriate the sunk costs of later purchasers, with forgone revenue from transporting less power.[2] The hold-up problem would vanish.

Given this, the Smiths could buy appliances without concern about ex post appropriation. Much the same circumstances apply to hundreds of thousands of businesses. As shown in chapter 2, electricity networks are characterised by diffuse demand, accounting for a small share of the inputs of most industries.

---

2 It is worth emphasising that, to the extent that hold-up applies, it is a problem for the business with monopoly power as well as the consumer. The monopoly business would like to charge a consistently high price that maximise its revenue, but the hold-up problem results in a level of output lower than the profit maximising level.

However, households and small businesses could be sure that an unregulated monopoly would set inefficiently high average prices at all times, and may not guarantee sufficiently reliable supply (the usual textbook outcomes of monopoly). Such a perennially high average price would, at the margin affect the willingness of the Smiths (and indeed all others) from buying appliances that use electricity. This is one way in which the conventional deadweight costs of monopoly pricing emerge, but it does not rely on the hold-up problem.

Accordingly, simply by dint of their large numbers and the sequential nature of their electricity-intensive investments, the bulk of electricity customers would be unlikely to be subject to the risk of ex post appropriation by an unregulated electricity network (a point also made by Boffa and Kiesling 2006). Were hold-up the *primary* motivator (as claimed) for monopoly regulation, it suggests that the Australian Competition and Consumer Commission (ACCC)/Australian Energy Regulator (AER) should not regulate these parts of the market. That would be unfortunate given the myriad of costs that monopoly power could impose on such customers. While regulators should abstain from over-regulation, the Commission considers that the ACCC and the AER have strong reasons for regulating on behalf of consumers and small businesses, and should not relinquish that role despite the absence of credible hold-up problems.

However, perhaps the hold-up problem applies to business customers making very large investments, but that are unable to write sufficiently complete contracts that forestall this?

Overcoming credibility problems for a monopolist selling to large business customers

A monopolist may credibly signal that it will not expropriate the sunk value of a customer's large lumpy investments in several ways.

Writing long-term contracts

While it is impossible to write perfect contracts, long-term contracts are common between large customers and businesses with market power, as demonstrated in Australian airports (PC 2011b), and in gas and electricity markets in the United States and Europe. It is notable that while Biggar cites Laffont and Tirole (2000, pp. 74-5) approvingly when they discuss the hold-up problem, he neglects their observation that:

> Despite the usual difficulties involved in designing good long-term contracts, this is often a decent way to resolve the hold-up problem.

In fact, Biggar (2010) acknowledges that such long-term contracts appear to be viable in electricity:

> In practice, the best we can say is that in some limited situations, long-term private contracts might be able to achieve a reasonably efficient outcome. For example, [this might occur] where the users can negotiate with potential suppliers before one or other side has sunk any investment and when the industry is stable so that reasonably long-term contracts are possible. Possible examples include … a large electricity consumer negotiating with the electricity transmission network …[3] (slide 6)

Other solutions to hold-up may also exist

Even without contracts, hold-up may not occur where the monopolist:

- vertically integrates with those customers that need to make very large and relationship-specific investments in long-lived assets (Meade and O'Connor 2009 and Michaels 2006).[4] This is likely to be more feasible for this group of customers than for customers generally
- takes an equity position in the customer's business (Harbaugh 2001) or an option to do so (Nöldeke and Schmidt 1998)
- is cooperatively owned by customers (apparently occurring in the United States and New Zealand).[5]

Hold-up problems may also be overcome when there are many repeated transactions with a customer or where the investments occur in stages. For example, notwithstanding vague contracts, there is a stable long-run relationship between Japanese auto-component manufacturers and automakers because there are long-run gains for both parties from stable long-term trading (Hölmstrom and Roberts 1998). Similarly, electricity users will often expand or upgrade plants over time, with the additional power usage providing a revenue source for the networks. The monopolist can reveal that it is acting in good faith by not behaving opportunistically for each of the successive customer investments. As in the auto case, any single incident of hold-up would end that relationship. The authors note:

---

3 Another illustration is the long-term contracts forged in the 1980s (running to 2014) between Alcoa and the Victorian Government — effectively the supplier in this instance — for power tariffs that moved in line with global aluminium prices.

4 Notably, in this literature, the hold-up problems are more realistic, relating mainly to the relationships between generators and transmission businesses, not between transmission businesses and downstream customers.

5 Personal communication from Daryl Biggar. He also cites second-sourcing as a mechanism that a monopolist might use to credibly commit to the continuation of reliable supply (negating hold-up). However, while second sourcing may apply to some electronic products (such as proprietary computer chips), it is not clear how they would apply to electricity supply.

> The familiar logic of repeated games, that future rewards and punishments motivate current behaviour, supports the on-going dealings. An attempted hold-up would presumably bring severe future penalties. (Hölmstrom and Roberts 1998, p. 82)

Other (theoretically based) papers support the view that repeated transactions resolve hold-up.[6] But such is the fragility of results to alternative assumptions, that others find quite the opposite conclusion, or supply other reasons why hold-up might or might not occur. This reflects another fundamental problem in the practical policy application of hold-up ideas. The hold-up literature is a lively part of the burgeoning field of industrial organisation and game theory, driven by elegance and assumption. The overwhelming bulk of papers are theoretical and, to the extent that there are empirical papers, they often rely on experiments involving students (for example, Hoppe and Schmitz 2011). It is hard to be confident that a single interpretation of hold-up within this vast and chaotic literature provides a credible basis for price regulation by competition authorities.

It is also not clear why the problem of hold-up is unique to a natural monopoly's dealings with its customers. Most of the examples of potential hold-up and the way it is resolved occur in industries where (ex ante) no market power is obvious.

Would price regulation be the regulatory remedy for hold-up?

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. Some possible targeted alternatives would be:

- ex post action based on common law breaches of the duty of good faith and fair dealing (Shavell 2007)
- specific competition laws (resembling Article 82 in the European Union Competition Law[7])
- ex ante oversight of particular long-term contracts that risked anti-competitive outcomes.

In United States and European law it is unlikely that a business with monopoly power could confiscate the investment returns of purchasers through re-negotiation

6 As explored in various contexts by Pitchford and Synder (2004); De Fraja (1999); Che (2000); Che and Sákovics (2004); and Iyer and Schoar (2009).

7 Article 82 requires that any abuse by one or more undertakings of a dominant position shall be prohibited as incompatible with the common market in so far as it may affect trade between member states. Among other things, such abuse may consist of directly or indirectly imposing unfair purchase or selling prices or other unfair trading conditions.

of a contract, although it might reduce their profits (White 2012). Even if there are transactions costs in taking action of these kinds, the threat of litigation may deter opportunistic behaviour. So were hold-up the primary problem posed by monopoly, then that might suggest the redundancy of price regulation — a problematic policy position. A conventional model of monopoly would not reach that conclusion.

It is notable that where international competition authorities query contracts, it is not because of their potential to create hold-up but because of risks of foreclosure and concerns about average pricing levels, such as cases:

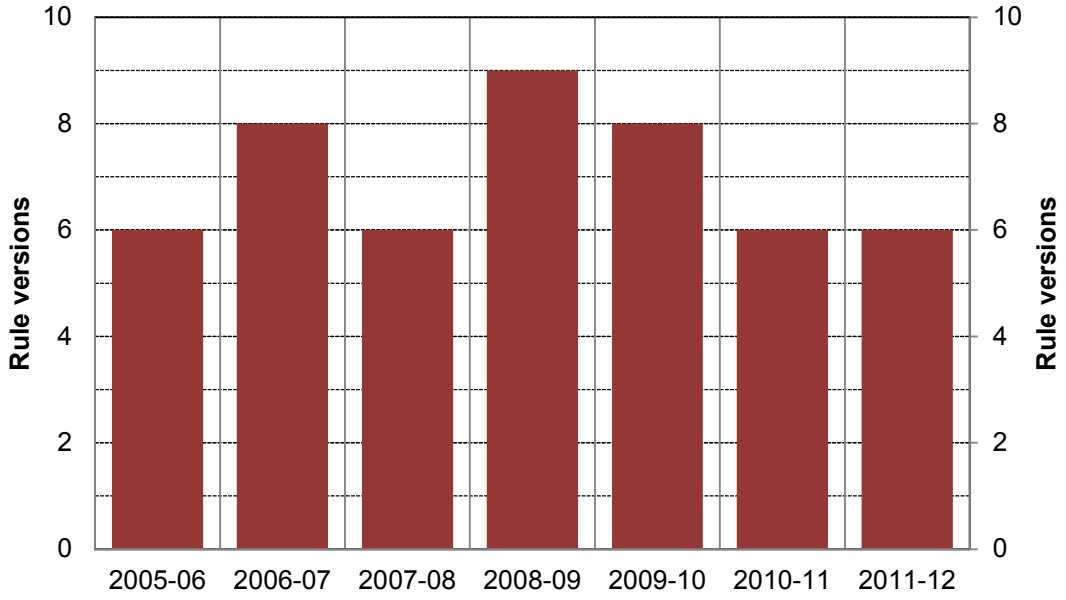
- where a party with market power seeks to tie up its purchasers in long-term supply contracts, so that parties competing with the dominant firm cannot supply its customers (Bellantuono 2008, Bessot et al. 2011). For example, a major gas supplier in Europe attempted to tie up many generators in long-term contracts
- where, ex ante, parties sign long-term contracts with 'eyes wide open', but that the purchaser *without* market power wishes to void at a later time advantageous to itself. For instance, utilities made long-term energy contracts with generators during the Californian electricity crisis to protect themselves from high spot prices, and then subsequently sought to have these contracts voided. This is a complex area in law, with considerable debate about the circumstances where a court may (or should) void a contract, but does not resemble the hold-up of Biggar's kind.

Are regulations a superior form of implicit contract?

Regardless, the notion that regulations (and their underpinning laws) create credible long-term contracts belies how regulations change and how regulators behave.

The asset life of electricity transmission equipment is often 40 to 50 years, and some Australian networks have operating equipment up to 80 years old (Crisp 2003, p. 22). Similarly, customers' investments can be in long-lived assets whose returns are dependent on efficiently supplied and priced network services (such as a smelter, glass works or manufacturing plant). Yet from 2005-06 to 2011-12, there have been 49 versions of the National Electricity Rules (and already six more versions up to 7th March 2013 in the 2012-13 period). There is no sign that the numbers of new versions per year are trending down (figure B.1).

Figure B.1 **There have been 49 versions of the 'Rules' from 2005-06 to 2011-12**



Data source: Information obtained from the AEMC website.

Were this to continue for an asset with a life of 40 years (whether owned by the customer or the monopolist), the economic value of the asset would be exposed to nearly 280 sets of possible new regulatory influences during its life. While most of these influences would probably be benign, there is no contract to ensure that. In effect, governments (and regulators through their discretion) can re-negotiate the implicit contract between network businesses and their customers. Neither party would know what the terms would look like five years later. Accordingly, both parties are exposed to regulatory hold-up.

The risks are likely to be greater for the network businesses. While some customers of electricity networks make large lumpy investments that is not true for most of them. In comparison, *all* network businesses are highly capital intensive and have very long-lived, relationship specific assets (chapter 2). Indeed, some prominent competition economists consider that one of the major risks in competition regulation arise because regulators are unable to commit that they will not expropriate the returns from the monopolist's long-lived sunk assets (Vogelsang 2010, Panzar 2012, Knittel 2006).

Unlike commercial arrangements, a regulator has an unassailable power to expropriate the monopolist's sunk investment unless moderated by some kind of merits appeal process. In effect, a regulator may possess monopoly powers well in excess of the business they regulate. As noted above, Biggar proposes considerable regulatory discretion, no merits review process and that (reversing the current onus

of proof in the existing Rules) the business should prove that the regulator is wrong in any determination (2011b, p. 8). These elements undermine the credibility of the regulator to create long-term implicit contracts immune from regulatory hold-up.

Do regulators ignore deadweight costs in practice?

Biggar (2011b, p. 23) suggests that:

> … regulators do not, in practice, behave as though minimising deadweight loss is their primary concern (or virtually any concern at all).

There are several problems with this contention.

Regulatory practice is not a guide to good regulation

The massive literature on regulatory capture contradicts the notion that regulatory behaviour necessarily reveals the legitimate rationale for regulation.[8] As an illustration, in the early 20th century United States, corruption among municipal energy regulators was apparently rife (Troeskin 2006). More recently, the United Kingdom competition economist, George Yarrow (2011a) quipped:

> On the basis of observed behaviours, the likely conclusion of a visiting social scien[tist] could well be that 'the principal objective of regulators is to convey a good impression of themselves'.

Nevertheless, it is reasonable to observe regulators' behaviour and ask them about their motivations because, through experimentation and learning, they might have discovered a different or additional basis for regulation. However, the ultimate test is not what regulators do, but whether, given market behaviour and its outcomes, regulation is justified on welfare grounds.

Mostly Australian regulators do not say that hold-up is their principal concern and they do not behave as if it is

In any case, Australian regulators often raise the relevance of deadweight costs. For example, the ACCC Deputy CEO of Regulatory Affairs, Mark Pearson (2011) has recently affirmed the latter 'text book' view, noting that the fundamental problem of market power is a 'contraction in supply resulting in higher prices or poorer quality of service'.

---

8 For example, as described by Levi-Faur (2012), Baldwin et al. (2011); Boehm (2007); Levine and Forrence (1990) and Stigler (1971).

A recent ACCC/AER working paper has also highlighted the primacy of the standard efficiency criteria for regulation of infrastructure:

> Across the OECD countries, it is common for governments to operate regulatory regimes aimed at producing more efficient outcomes than the unrestrained market. In Australia, as in most OECD countries, efficiency is interpreted broadly to include … cost efficiency… allocative efficiency… and dynamic efficiency …. These efficiency criteria underlie both the National Electricity Law (NEL) and the National Gas Law (NGL). They date back to the Hilmer reforms of the early nineties… and pervade all of Australia's infrastructure regulation. (ACCC/AER 2012a, p. 9)

What customers say is also revealing. It is notable that large users do not cite contractual or hold-up problems as their prime concern. Rather, they are concerned about excessive costs and the deadweight losses these entail:

> The NEM design is based on providing strong incentives for the supply side to provide a vibrant and responsive electricity supply. If incentives are inappropriate and over-incentivised investments are made in transmission (and distribution) networks – as have been the case under the existing Rules — users of energy will face significantly higher but arguably unnecessary costs (and hence adversely affect downstream investments). Even more importantly, the Australian economy will be incurring large dead weight losses. (Major Energy Users, sub. 11, p. 6)

A counterargument might be that it is not what people say, but what they do that sheds light on the theory underlying their actions. That may be true for some consumers and business managers, but it would be worrying if it were to hold for experienced competition economists representing regulators and large businesses. It would also undermine the value of articulating sound rationales for regulation if regulators somehow instinctively get it right anyway.

Moreover, it is not clear that regulators *do* behave as if hold-up was their unstated, but actual motivation for regulation. For example:

- while sometimes regulators may question particular tariff structures, it is not clear that they always oppose price discrimination. Two-part tariff schemes (second-degree price discrimination) are common in infrastructure access arrangements in Australia and internationally, and are seen as economically efficient, a point made by Biggar (2001). The AER has not argued for limiting the considerable discretion of electricity distribution businesses to set prices flexibly (AER 2012a, p. 17) and the Rules specify the need for efficient tariff structures (clause 6.2.5(c)).
- the AER does not oppose peak load pricing and were they to do so, it would be undesirable for consumers as a whole (chapters 9–12). A major barrier to peak load pricing in electricity has been the absence of smart meters, a technological barrier, rather than an intrinsic hostility by regulators. Peak load pricing is

common among commercial customers, as is pre-agreed load shedding by large customers when total electricity demand exceeds the available supply. The AER has not expressed opposition to critical peak pricing trials as part of demand management trials (AER 2012b, p. 16).

- the AER has not questioned the value of a merits review process, but more the design of the existing arrangements. In its submission to the review of the LMR, the AER (2012l, p. 1) observed that: 'As an administrative process, it is appropriate that [utility regulation] be overseen by judicial review and some form of limited merits review mechanism.' This appears at odds with the apparent implications of hold-up for merit review as described by Biggar.

Hold-up costs are deadweight costs

Even were it to occur, the reason that hold-up is problematic is that it curtails investment, with the dynamic inefficiencies that flow from that. As pointed out by Frontier Economics (2011, p. 22), this is just one form of a deadweight loss.

B.2 In summary

The Commission does not see hold-up as the dominant rationale for regulation of natural monopolies. However, Biggar's research into hold-up problems has still been useful in highlighting the potential intertemporal aspects of deadweight costs, may be relevant to specific cases of anti-competitive conduct, and more generally adds to the burgeoning modern literature on monopoly and consumer behaviour where there is asymmetric information and strategic game playing.