The Relationship Between Regulation and Competition Policy for Network Industries

David Newbery

March 2006

CWPE 0631 *and* EPRG 0611

The relationship between regulation and competition policy for network utilities[1]

David Newbery
June 2005

Abstract
One of the central questions facing the regulatory and competition authorities is whether to delegate the regulation of the potentially competitive elements of the utility to specialised regulatory agencies, or whether they should be solely subject to normal competition laws enforced by the competition authorities – often referred to as 'light-handed regulation'. German electricity and New Zealand utilities exemplify this approach, with the United States at the other extreme in granting considerable powers to federal regulatory agencies. Britain evolved a system of licensing the majority of network activities, including the larger part of the potentially competitive services such as electricity generation and supply, gas trading and supply, and mobile telephony. The EU, in contrast, has always placed more emphasis on the application of competition law, and has sought to make sector regulation consistent with general competition law. This is most clearly evident in the 2003 EU Communications Directives which required considerable changes on the British approach to telecoms regulation.

The paper discusses the appropriateness of this approach for telecoms and electricity. Post-modern utilities like telecoms, in which facilities-based competition is possible, lend themselves to the approach laid out in the Communications Directives. Regulators must first determine whether markets are effectively competitive or suffer from Significant Market Power (SMP). Only in the latter case is *ex ante* regulation warranted, and then only if it is necessary, justified and proportionate. The paper considers how this approach works in the case of mobile call termination, and the implications of restricting regulation solely to the market experiencing SMP.

Electricity does not fit comfortably into this approach, in part because the wholesale market is more likely to suffer from collective dominance than the single firm dominance that allows for an easier determination of SMP. Licence conditions are no longer needed under the Telecommunications Directives, but appear to retain advantages for electricity, where information and continuing market surveillance are desirable, and where it may be necessary to modify market rules in a timely and well-informed manner. The paper considers the example of the English Electricity Pool, which suffered from tacit collusion in the period 1996-99, and the unsuccessful attempts of the regulator to impose a Market Abuse Licence Condition.

JEL: G18, L94, L96
Key words: Regulation, competition policy, telecommunications, electricity, market power

[1] Paper presented at the conference in tribute to Jean-Jacques Laffont in Toulouse on 30 June 2005. The paper was originally written for the conference *How should competition policy transform itself? Designing the new competition policy*, 20 November 2003, in Japan.

Introduction

The deregulation movement that started in the United States was intended to remove detailed government control from otherwise competitive markets. The theory of regulatory capture expounded by Stigler (1971) and Peltzman (1976) argued that most industrial regulation was designed to protect incumbents against competition from entrants and had the effect of legalising monopoly power to the detriment of consumer welfare. Transport deregulation followed, most notably after Alfred Kahn's move to the Federal Aviation Administration and the subsequent deregulation of airlines. Normal competition policy would replace the detailed regulatory control exercised by state-level and federal regulators.

It is hard to see why naturally competitive industries like trucking should ever be subject to economic regulation, given that trucks meet the required safety and environmental standards. Similarly, the theory of contestability suggests that mobile assets like aeroplanes do not need to be protected against competition nor regulated to prevent excessive prices. In short, where competitive markets can work, they should not be subject to economic regulation, but only to normal competition law.

The situation is radically different for network utilities, where the network is typically a natural monopoly, delivering essential services to the mass of the voting consumer population. Electricity grids, gas distribution networks, local telephone networks, water and railways all share this property. Modern societies will not tolerate the unimpeded exercise of market power over such essential facilities. Network builders typically need rights to build networks, and local or national jurisdictions will only grant such rights in exchange for regulatory powers over prices and quality. Private investors will be reluctant to sink expensive capital investments without some assurance that they will not subsequently be prevented from charging remunerative prices. The only solution under private ownership is regulation. If such regulation is not credible, adequate investment will not be forthcoming and state or municipal supply will be the only viable default option.

Cost-of-service (or rate-of-return) regulation evolved in the United States over nearly two centuries in response to legal challenges to clarify the nature of property rights and place restraints on regulatory activism. When Britain came to privatise British Telecommunications (BT) in 1984, it was recognised that BT would have overwhelming market power in local telephone services, and for a considerable period very substantial market power in long-distance telephony. Regulation was clearly required, and Professor Littlechild was commissioned to propose a suitable form of regulation. Littlechild designed price-cap regulation as the natural solution to moving from an overwhelmingly concentrated market to what was anticipated to become a competitive telecommunications market in due course (Littlechild, 1983; Bartle, 2003). Price-cap regulation has the attractive property of simulating the competitive market even in the presence of market power.

In the United States, the Modified Final Judgement of the AT&T case in 1984 suggested that at least long-distance telephony might evolve to become a competitive industry. Long-distance telephony was no longer a natural monopoly with the advent of microwave transmission, the huge increase in capacity provided by fibre optic cables and the rapid fall in the cost of micro-electronics and switching capacity. This raised the question

why telecoms should be subject to a sector-specific regulator, rather than being subject to normal competition law. Instead of imposing a price-cap, competition authorities could fine companies holding dominant positions in any market if they abused their market power by charging excessively high prices. The threat of penalties would, on this theory, deter utilities from exercising their market power, while minimising intrusive and costly regulation.

The case for minimising regulation is powerful. The US evidence suggests that regulation evolves in response to pressures from special interest groups. On this view regulators are invariably captured, and use their regulatory powers to redistribute rents rather than to drive down costs and benefit consumers. The theory of asymmetric information suggests that just as redistributive taxation is distortionary and therefore costly, so attempts to redistribute cost savings from utilities to consumers will produce inefficiencies and dead-weight loss. Regulatory inefficiency does not end with overpricing or inefficient supply. Hausman (1997) argues that one of the most costly consequences of traditional cost-of-service regulation is the delay hindering the introduction of new services, caused by concern that these new services might be cross-subsidised from the captive customer base. Hausman gives two examples - voice messaging and cellar phones. AT&T originally wished to offer voice messaging in the late 1970s, which the FCC delayed until 1988, when it was immediately introduced. The estimated consumer surplus generated is estimated at US$1.27 billion per year by 1994, so the ten-year regulatory delay cost billions of dollars. Hausman further argues that a possible ten-year regulatory delay in introducing mobile phones might have cost consumers cumulatively as much as $100 billion, large compared with the 1995 US global telecoms revenues of $180 billion/year.

Avoiding unnecessary regulation is therefore to be desired. New Zealand was the first country to follow this logic when it came to liberalise its telecommunications market on 1 April 1989 (Bollard and Pickford, 1996). The state telephone monopoly that had been corporatised as Telecom in 1987, was sold in 1990 to a joint venture between local investors and the two American companies, Bell Atlantic and Ameritech, without establishing an independent regulatory body. The company was required to disclose financial and pricing information, and the “Kiwi share” obligation required Telecom to provide free local calls for residential customers and to not increase the (uniform country-wide) rental charge faster than the rate of inflation. It also provided an undertaking to ensure that interconnection would be provided to competitors on a “fair and reasonable basis”.

Clear Communications negotiated a long-distance access agreement with Telecom and entered the market in April 1991. By 1995, Clear had 21% of the domestic long-distance market and 24% of the international calls market. Shortly afterwards, Clear attempted to reach an agreement to access Telecom's local loops in order to be able to provide a local call service, but negotiations broke down. The matter at issue was whether Telecom could set the access charge using the Efficient Component Pricing Rule of Baumol (1983) and Willig (1979), which included the lost profit of Telecom, or whether, as Clear argued, the access price should be based on avoidable cost. Clear took Telecom to court for claimed abuse of its dominant position. The High Court supported Telecom's application of the Baumol-Willig rule but found that Telecom had indeed impeded entry. Both companies appealed and in

December 1993 the Court of Appeal upheld the dominance finding but found against the Baumol-Willig rule. The case was finally appealed to the Privy Council in Britain, who accepted the original High Court judgement and supported the Baumol-Willig rule. Their argument was that if the monopoly price element (Telecom's lost profit from each call carried by Clear) was objectionable, then it was open to the government to control those monopoly profits directly under the provisions of the Commerce Act.

The case is justly famous, both because it illuminated the economic principals of the efficient component pricing rule, and because it demonstrated the major drawback of relying on competition law, namely the time and cost involved in resolving disputes. The case started in 1991 and the government did not resolve the issue until July 1995. In 2001 the government passed a Telecommunications Act which obliges the Commerce Commission to resolve access disputes, to monitor the regulatory regime, and to recommend changes to the list of regulated services. It was not until April 2003 that the Commerce Commission, acting under the requirement of the Telecommunications Act, published an issues paper consulting on whether access to the local loop should be regulated.[2] New Zealand appears to be following an evolutionary path in clarifying the role of regulatory agencies, starting from a minimalist position, but responding to the need to resolve disputes quickly and to restrain market power more effectively than appears possible under the normal provisions of competition law.

There are similar lessons to learn from the early experience of telecoms regulation in the UK. When BT was privatised in 1984, it faced a single small licensed rival, Mercury, that was granted protection against entry until at least 1991. Although BT's services to final customers were subject to price-cap regulation, access agreements between BT and Mercury were left to commercial negotiation. Mercury, like Clear in New Zealand, immediately disputed the terms offered by BT and appealed to OFTEL, the telecoms regulator. Economists rapidly established that interconnection agreements between symmetric network operators, each of which had sole ability to terminate calls to its subscribers, would lead to agreements that were at or above the profit-maximising level. The fact that both operators might speedily agree access charges was no guarantee that these charges would be in the public interest. Fortunately, access agreements between very asymmetric players might break down as the dominant player attempted to exclude his competitor, and in such cases disputes would come to light, as with Clear and Mercury. Since then, the theory of access pricing has been developed (Armstrong, 1996, 1998; Laffont, Rey and Tirole, 1998a, 1998b) so that interconnection is now recognised to be an area of regulatory scrutiny, if not regulatory control.

Implicit versus explicit regulation of natural monopolies

New Zealand and Germany lacked regulators for at least some network monopolies until recently.[3] In effect, the network owners were subject to implicit rather than explicit

[2] See *Network*, Issue 14, August 2003, published by the ACCC, Melbourne, Australia

[3] At the date of the conference, Germany still awaited the final passage of the relevant legislation setting up an electricity regulator, but this was anticipated by July 2005.

regulation – they could pursue profits providing they did not violate competition legislation, and specifically did not abuse their dominant position. This raises the question whether there are important differences between implicit and explicit regulation, and whether in particular there are additional advantages to implicit regulation (apart from avoiding the drawbacks of regulation already noted). Utilities might argue that any action that does not breach the rules of explicit regulation is legitimate. Poorly designed or incomplete regulation (for example, failing to address issues of quality) may then lead to unsatisfactory outcomes. Utilities might believe that the concept of abusing dominance was more encompassing and could include any outcome that damaged consumers compared to the counterfactual of a competitively supplied service, where consumers could choose according to quality and price. In that case implicit regulation might seem to have advantages.

The question is interesting but hard to answer. Poor regulation can normally be improved, although in some jurisdictions this might require new legislation that would be very hard to pass.[4] On the other hand, utilities might argue that abusing dominance was narrowly interpreted as unreasonable high pricing or unjustified discrimination. If there is no case law on other aspects, such as supplying unsatisfactory quality, then it might be hard to address such problems. There is a further and important point demonstrated in Germany. The justified (average) price for using a network will be the average cost, including a normal return on, and depreciation of, the capital. Valuing the capital in the network is therefore central to network regulation, and has been highly contentious in Britain (Newbery, 1997). The problem is whether to value the assets at historic cost, the privatisation sales value, or at written-down modern equivalent asset value (to take the most obvious alternatives). The differences can be profound: for water the sales value was £6 billion but the written down modern equivalent asset value was estimated at more than £100 billion.

In Britain, decisions by the regulator are typically appealed to and settled by the Competition Commission, thereby gradually building up suitable precedents. Without that process and the consequential development of case law, the courts examining whether prices were defensible or not might reach economically unsatisfactory conclusions, or might, particularly if the same competition authorities were involved as in regulatory dispute resolution, reach the same conclusions, but far more slowly and expensively.

The evolution of liberalisation towards structural reform

Telecommunications was the first network utility to be privatised in the UK, and the first to be liberalised in the United States, in both cases for good reason. The underlying case for liberalising network industries is that it allows competitive pressure to be put on sleepy monopolies, and restricts cross-subsidies that frequently take the form of a tax on competitive medium-sized industry to subsidise domestic consumers (and sometimes politically powerful large business). In the case of telecoms, where technical progress is rapid, competition is the best way of identifying winners and enabling them to replace losers. It is also easier to sustain

[4] Chile is a notable example, see Bitran and Serra, (1998), Heller and McCubbins, (1996)

telecoms as a private regulated utility than with most other utilities, as there is less difficulty in designing a system of regulation that will be credible and therefore sustainable.

While the United States had strong Constitutional protection of regulated utilities, reinforced by extensive case law and underpinned by Administrative Law (Levy and Spiller, 1996), Britain had a sovereign Parliament that was free to change laws, and hence to threaten the stability of any legal arrangements. The essence of credible regulation is that the private investor should have confidence that his sunk investment will not be subject to regulatory opportunism, by, for example, reducing regulated prices to non-remunerative levels. The temptation to act opportunistically is counterbalanced by the threat of the utility to not invest if it finds the regulatory compact has been broken (Gilbert and Newbery, 1996; Newbery, 2000). In the case of telecoms, the rate of growth of demand is high, the need for investment in new technology continuous, and the expertise of the telecoms company in managing the software and equipment critical. All these factors give the utility strong bargaining power in the repeated game between regulator and utility.

Telecoms industries are, however, atypical of other network utilities, in that liberalisation normally retains the initial industrial structure, with entry taking place by competing networks that require interconnection. The United States was relatively unusual in that AT&T agreed to its own break-up into Regional Bell Operating Companies (RBOCs) and the long-distance AT&T, reflecting the earlier functional organisation of the company and the vast size of the United States (Temin, 1987). BT was allowed to retain its local exchanges with their local loops as well as its long-distance network and international lines; a model followed in most modest-sized countries. Even in the United States, competition took place between different interstate networks, each of which had to interconnect with an RBOC. In that sense, long-distance networks had ceased to be natural monopolies, although they continued to be regulated by the FCC.

The quest for increased competition in the United States led to the gradual restructuring of regulation in the privately owned but regulated natural gas industry. The gas industry appeared ideally suited to competition, with over 8,000 producers, the 40 largest of which accounted for 57% of 1990 gas production. They were connected to more than 1600 Local Distribution Companies through 44 major interstate pipeline systems and hundreds of smaller pipeline companies (IEA, 1994). Despite the obvious prospect of competition in supply, well-head gas prices were regulated, creating huge inefficiencies after the oil price rises of the 1970s. The subsequent collapse of the market as long-term contracts were signed at the peak of the oil price boom created excess capacity in pipeline networks, ideal conditions for the introduction of competition (IEA, 1994). Various Federal Energy Regulatory Commission orders from 1984 on finally unbundled and restructured the industry to one of third party access that facilitated active competition.

As with the privatisation of BT, Britain was slow to learn the importance of competition for enhancing performance, and thus also slow to realise the importance of restructuring utilities to achieve effective competition. In the case of BT, the legislation had at least envisaged a review of the duopoly in 1991. The Duopoly Review recommended full liberalisation of entry, which duly occurred. By 1995 there were over 150 operators licensed to

compete with BT, including 125 cable TV companies who could offer telephony with cable, of which 80 were actually providing service, (Bell, 1995). Rapid entry was followed by substantial consolidation, but cable companies now have more than half the broadband market. The failure to create a more competitive structure at privatisation and the delay in allowing competition led Armstrong, Cowan and Vickers (1994) to the judgement that 'the duopoly policy has been detrimental to development of competition, and its main beneficiary has been BT itself' (p.240).

Unfortunately, the American liberalisation experience (in both gas and telecoms) was not exploited when British Gas was privatised as a vertically integrated monopoly in 1986, although Vickers and Yarrow (1988, p267) criticised the failure to introduce competition. Under intense regulatory pressure, supported by the Office of Fair Trading and via references to the Monopolies and Mergers Commission, British Gas was persuaded to first functionally unbundle the competitive services from the natural monopoly networks, and then finally to break up the company into competitive and monopoly parts. The latter, Transco, in turn sold off storage activities and finally merged with the already unbundled electricity National Grid Company, consolidating the regulated energy networks in one company. Even here, regulatory activism has attempted to replace regulated prices by market mechanisms of auctions, first for storage, and then for scarce entry capacity from the beach-head (to which gas is delivered by various oil and gas companies using their own pipelines). Fortunately, by the time the electricity supply industry (ESI) came to be privatised in 1990, the message of restructuring to introduce competition before privatisation had finally been learned.

Unbundling the competitive services

The new conventional wisdom is that network utilities should be unbundled, with the potentially competitive network services under separate ownership from the natural monopoly network, so that the network owner has no incentive to favour its own service provider. Such vertical separation of network service from the network risks losing economies of scope and hence raising costs. In the case of electricity and gas, these economies of scope appear small compared to the potential efficiency gains, but in the case of telecoms this is less obvious. The operator of a switched network has to create end-to-end connectivity and at the same time calculate the cost of the call in order to bill the customer. The design of switches to route and record such information have natural synergies that would take considerable investment and design to replicate at comparable cost in unbundled form. In those parts of the network where economies of scale are not significant (long-distance, international) facilities-based competition is feasible and now fairly standard, but where there are important natural monopolies (arguably still in the local loop), vertical integration between call delivery, billing and network ownership is still common. Positive steps to overcome the advantages of incumbency are then required (such as number portability and carrier pre-selection).

The attraction of ownership unbundling is two-fold - it allows competition to give companies better incentives to drive down costs, and it also offers the prospect of removing costly regulation. The new wisdom of "competition where feasible, regulation where not" suggests that regulation should be confined to the natural monopoly elements, typically the networks, with network services subject only to competition law. Indeed, in transcribing the EU

Electricity Directive 96/92 into national legislation, some Continental countries adopted this approach. Britain followed an earlier and different path. The difficulty of creating credible regulatory structures through primary legislation led the designers to propose licences that would contain the important regulatory details. These licences had the force of contracts, enforceable through the courts. They could only be changed by mutual agreement between the utility and the regulator, or if the regulator found that continuing the existing licence conditions unchanged would be against the public interest (normally at the periodic price control review). The regulator could then propose a licence modification (normally a change in the level of the price cap for the start of the next regulatory period, and/or a change in the X-factor in the RPI-X formula). This could either be accepted or appealed to the Competition Commission – an alternative costly to both utility and regulator, and hence to be avoided if possible.

Licences were clearly central to the British style of price-cap regulation of the network natural monopolies, but were also retained for the potentially competitive network services. Their licence conditions typically contained rules of market behaviour, and provided for information disclosure to the regulator, who was thus able to exercise a market surveillance role. The position now is that, after the New Zealand experience, it is widely accepted that *ex ante* regulation is required for network natural monopolies, but there is less agreement whether such regulation is required for network services, or under what circumstances and in what form such regulation might be desirable.

Regulating network services

Regulatory design in Britain, the United States and in the EU has followed quite different routes. In the United States, utilities were already for the most part investor owned and regulated as vertically integrated franchise monopolies, under existing legislation. To take a good example, the Federal Energy Regulatory Commission, FERC, has extensive powers under the 1935 Federal Power Act. Individual States have devolved regulatory powers, but FERC has a statutory obligation to ensure that wholesale prices are “just and reasonable”. This gives FERC considerable powers to intervene if liberalisation fails to deliver satisfactory outcomes. Britain, as already noted, had to design a system of regulation at short notice and evolved a system of licences. Where the industry was liberalised or restructured, the major companies were typically required to hold licences. On the Continent, competitive services were frequently subject only to normal competition legislation. The Single European Act mandated the European Commission to propose policies to bring about a single market in gas, electricity, transport and telecommunications. The Commission has been remarkably successful in issuing Directives with that intent, and is following a strategy that builds on existing articles of the Treaty of Rome and therefore already have been incorporated into each member state’s legislation.

Applications to regulating telecommunications

The four EC Communications Directives that came into force on 24 April 2002 are an instructive example of how this process works for delivering a single market in telecommunications operating under a consistent regulatory framework. The Commission

Guidelines on market analysis explains how they will work (CEC, 2002). The new regulatory framework requires that "the markets to be regulated are defined in accordance with the principles of European competition law." (*Guidelines*, para 4). Later on it notes that "to ensure consistency, these guidelines are based on (1) existing case-law of the European Court of Justice concerning market definition and the notion of dominant position within the meaning of Article 82 of the EC Treaty and Article 2 of the merger control Regulation; …" (para 24).

The four directives are the *Framework* Directive, the *Access and Interconnection* Directive, the *Authorisation* Directive, and the *Universal Service* Directive, which member states are required to implement from 25 July 2003. The *Framework* Directive requires National Regulatory Authorities (NRAs) to conduct market reviews and to determine which markets shall be deemed to be effectively competitive, and which markets are susceptible to *ex ante* regulation. A market is effectively competitive when no operator in that market possess Significant Market Power (SMP), which is newly defined in Article 14 of the Framework Directive to be equivalent to the competition law concept of dominance: "An undertaking shall be deemed to have SMP if, either individually or jointly with others, it enjoys a position of dominance, that is to say a position of economic strength affording it the power to behave to an appreciable extent independently of competitors, customers and ultimately consumers."[5]

The final stage is that the Commission issued a *Recommendation* identifying "those product and service markets in which *ex ante* regulation may be warranted." (CEC, 2003, para 2). This sets out criteria for identifying markets where regulation may be warranted: high and non-transitory entry barriers, those whose structure does not tend towards effective competition within the relevant time horizon, and those where the application of competition law would not adequately address the market failure concerned (para. 9). "It is considered that markets that are not identified[6] in the Recommendation will not warrant *ex-ante* regulation," (*Guidelines*, para 29). The Recommendation then defines 18 markets that NRAs are required to review.

The *Guidelines* "requires NRAs to ensure that the measures they impose on SMP operators under Article 16 of the Framework Directive are justified in relation to the objectives set out in Article 8 and are proportionate to the achievement of those objectives." Later, (at para. 118) "the means used to attain a given end should be no more than what is appropriate and necessary to achieve that end."

The cumulative effect of these Directives is to limit regulation to those markets where competition law would be inadequate, and then further to restrict the scope of regulation to the minimum justifiable level. Whereas the American approach is to gradually relax an all-encompassing ability to regulate, the European approach is to propose the minimum degree of regulation that could be justified in a court of law. The British approach started rather closer to the American concept in normally requiring network services to hold licences (except for some new and small entrants). Each approach reflects the evolutionary history of

[5] EU Framework Directive 2002/21/EC, *Official Journal*, 1L 108/33-49, 24.4.2002

[6] as having SMP

utility ownership and regulation. The United States had evolved a system of regulation that was then enshrined in Acts, Britain had to design a system *ab initio*, while the European Commission had to seek consensus among countries with very different patterns of ownership, utility duties, and central government control, starting from the pre-existing EC Treaty and its competition law.

Despite the limitations that consensus-seeking might suggest, one can argue that this approach has attractions for harmonising telecommunications markets, where companies not only need to communicate across borders but where mobile (cell-phone) operators also frequently operate in many European countries. One can draw a distinction between the classic and 'post-modern' network utility. In the classic network utility, the network is a natural monopoly subject to limited technical progress and supplying a mass service. In the 'post-modern' network utility facilities-based competition is possible, failure of a single network may not be critical, innovation is important, and services supplied are useful rather than essential. In the first case, regulation is inevitable, while in the second case, competition law has potentially strong attractions and should be given maximum scope. Competition law, not regulation, should be the default option for post-modern utilities.

The interpretation of Significant Market Power, or SMP, in a post-modern utility needs to take account of the dynamics of technical progress. The concept of single dominance involves balancing a number of considerations, no one of which is likely to be decisive by itself. Thus market shares are themselves not conclusive, although there is a presumption that SMP is unlikely with a market share of less than 25%, and likely with a market share of more than 50%.[7] Normally a share of 40% or more would be required, but in addition the analysis should allow for market shares that are fluctuating, markets that are emerging and/or rapidly growing, and should be suspicious of persistently high market shares. Barriers to entry through control of infrastructure, essential facilities, because of economies of scale or scope, and exercised by vertical integration, would suggest a greater possibility of exercising SMP.

Mobile telephony

All of these qualifications to relying on market shares alone apply with considerable force in determining whether mobile network operators (MNOs) have SMP in any of the relevant markets. Thus in Britain there were four MNOs in 2002, and five in 2003, with subscriber market shares that fall close to or below the threshold for dominance of 25%. The markets were initially unregulated (although all MNOs had licences), were clearly very dynamic, growing at very high rates, and most of the MNOs were not making profits. Nevertheless, Oftel monitored the various telecoms markets (both mobile and fixed-line), and in particular those markets experiencing SMP. In March 1998, the Director General of

[7] The US Department of Justice applies a screen based on the Hirschman-Herfindahl Index, based on the market shares of all participants (and equal to the sum of the percentage market shares, with 10,000 representing a single firm). Markets with HHIs above 1,800 attract attention, and this number could be reached by one firm with 40% share, and two firms of 10% shares each, all the rest being small.

Telecommunications (DGT, i.e. the regulator) referred the then two largest MNOs (Cellnet and Vodafone) and BT to the Monopolies and Mergers Commission relating to call termination charges (MMC, 1999). The DGT considered that the MNOs were able to set the termination charges without significant market pressure, that as a result they were excessive, and should be regulated to be cost-reflective. Each MNO has an effective monopoly in delivering calls to customers on its network, as at present there is no commercially feasible alternative way in which different networks can compete to deliver calls to any particular mobile customer. In addition, because the calling party pays rather than, as in the United States, the receiving party pays, the call recipient has little incentive to choose an MNO who offers lower call termination charges.[8] The market for mobile call termination is therefore an effective monopoly and exhibits SMP.

The DGT also argued that BT's retail charges for fixed-to-mobile calls ere not adequately restrained by competitive pressure, as BT carried 83% of fixed-to-mobile call minutes from residential users and 66% from businesses. As the cost of fixed-to-mobile calls was still a modest share of total bills, BT was not under enough competitive pressure to moderate these charges, and hence also possessed SMP in the fixed-to-mobile call market. As a result, the profit BT was making on fixed-to-mobile calls was far higher than on other comparable calls. As BT was subject to price control for other services where it was deemed to have SMP, it was natural to propose a price-cap for BT's fixed-to-mobile calls. The more interesting question was whether there should be a separate price cap for fixed-to-mobile calls, or whether to include fixed-to-mobile calls in the basket of services subject to an overall price cap. The latter arrangement has the attraction that (provided the basket is base-weighted) the utility has the correct incentives to choose the most efficient or Ramsey pattern of prices (Armstrong, Cowan and Vickers, 1994, p81). The MMC rejected placing fixed-to-mobile calls in the existing price-cap basket, as they were concerned that BT might be tempted to favour its then-subsidiary, Cellnet, and instead proposed a separate price-cap on BT's retention (i.e. the profit it made on fixed-to-mobile calls).

At the time of the 1998 inquiry, Vodafone and Cellnet between them had 72% of total mobile subscribers, which had fallen from 100% five years earlier (MMC, 1999, p147). The DGT therefore concluded that the two new entrants should be left unregulated (apart from the standard licence conditions), and they were not referred to the MMC, and not included in the proposed call termination charge price cap. This price cap would run until 2001/02.

In July 2000, Oftel began a review of competition in mobile termination to see whether the markets still experienced SMP. By then the two new entrants had secured half the market. Oftel concluded that there had been no technological change making call

[8] Some users (particularly businesses) are concerned to choose a network offering lower termination charges to those whose calls are valuable to the recipient (friends, customers, etc), and networks compete by charging lower on-net (i.e. calls originating and terminating on the same network) than off-net calls. Such price discrimination actually provides an incentive to raise off-net termination charges as these raise rivals' costs, and so further reduce competitive pressures to moderate call termination charges.

termination contestable, and the continuation of calling party pays reduced any competitive pressure that might exist on call termination. Oftel therefore concluded that each of the four MNOs had SMP in their respective markets for call termination, and proposed a new price cap for call termination charges. The MNOs appealed and the DGT referred the case to the Competition Commission (the successor to the MMC) in January 2002.[9]

The case took a year, and spanned the publication of the EC Telecommunications Directives. These Directives required Oftel to review the telecommunications market listed in the *Recommendation*. Oftel published its own guidelines in August 2002, although it did not publish its *Review of mobile wholesale voice call termination markets: EU Market Review* until 15 May 2003, after the publication of the Competition Commission report (CC, 2003). By then there were five MNOs, one of which, called 3, was planning to roll out a 3G network later in 2003.[10] Oftel's *Review* concluded "that each mobile network operator ("MNO") in the UK has significant market power in a separate market for voice call termination on its network, and in the case of 3, wholesale 2G voice call termination provided to its subscribers, and can be expected to have market power for at least the next three years."

The Competition Commission agreed with Oftel that the MNOs had SMP, that mobile call termination charges were excessive, and, as in the earlier inquiry, that the most appropriate remedy was a price cap on the call charges. The issue then resolved into how that price cap should be set. MNOs offer three services, subscription, associated with owning a hand-set, making and receiving calls. It is possible to measure the long-run incremental costs (LRIC) of providing each of these services, but pricing at LRIC would not cover the fixed and common costs of the mobile network. Even a competitive market would therefore have to mark up prices above LRIC if it were to break even. The MNOs argued that the admittedly high mark-up over long-run incremental costs of termination was used to subsidise hand-sets, which stimulated the development of the market and accelerated penetration. This in turn created beneficial network externalities to all callers (fixed as well as mobile), as it increased the number of people they could call and hence increased the value of each network. The combination of calling party pays (CPP), which facilitated hand-set subsidies and pay-as-you-go, dramatically increased the rate of market growth compared to the US approach of receiving party pays (RPP), so that Mexico decided to switch from RPP to CPP in 1999, and then experienced an acceleration in market growth.

Such network externalities had been recognised in the earlier MMC inquiry and justified an additional subsidy financed by call termination. The main issue was how to allocate the fixed and common costs (including internalising this network externality). The Competition Commission supported Oftel's recommendation that these costs should be financed by an equi-proportional mark-up on all LRICs. The MNO's argued for a Ramsey mark-up, which all parties

[9] The author should declare an interest in that he was at the time of writing an advisor to Vodafone in its case before the High Court.

[10] Before then, the MNOs were using second generation ('2G') technology, but all five companies had secured spectrum in earlier auctions, and were preparing to introduce third-generation or 3G packet-switched (always on) technology, with greater bandwidth and data handling capability.

accepted would be efficient if applied to all three mark-ups.[11] Ramsey mark-ups would be inversely proportional to market demand elasticities, and as these were lower for call termination than the other services, would justify a higher termination mark-up. Oftel argued that,[12] "even if termination charges were set at the Ramsey level, there would be no guarantee that MNOs would set their (unregulated) retail prices at the Ramsey levels. In Oftel's view the key reason is that the retail market for mobile services is not effectively competitive. If the retail market were perfectly competitive and if termination charges were regulated at the Ramsey level, it would be expected that firms would set Ramsey retail prices. The intuition is that in a perfectly competitive market firms are forced by competitive pressure to set prices that maximise consumer surplus."

The Competition Commission defended equi-proportional mark-up on essentially two grounds: that it was not fair to fixed-line callers to have a higher mark-up on mobile call termination, the benefits of which would flow back to subsidising mobile use, and that in any case the evidence for establishing the relevant demand elasticities was not sufficiently robust to justify a departure from equi-proportional mark-ups. The MNOs then appealed to the High Court for a Judicial Review. A judicial review of a Competition Commission case is essentially a test that the Commission had followed due process and given reasoned arguments for its conclusions, and does not as such normally reconsider whether the contestants' arguments were more compelling. Justice Moses delivered his judgement in favour of the Competition Commission on 27 June 2003.

The case is interesting in showing the effect of this competition law approach on regulation, but it also raises some important and not properly resolved issues. On the positive side, although the case was brought before the Telecommunications Directives came into force, Oftel was already conducting its market analyses and basing its case for regulation on very similar principles, particularly in defining markets and determining whether they suffered from SMP. Second, the idea of confining regulation only to those markets suffering from SMP and leaving other markets to competition law has now been accepted. From 25 July 2003 telecommunications licensing was abolished in the UK (and the rest of the EU) as a direct consequence of the new EU communications regime. Licenses will be replaced by General Conditions made under the British Communications Bill (which had not then been passed into law).[13] Third, regulation now has to be *justified* in relation to the objectives of the Directives, *proportionate*, and no more than that *appropriate* and *necessary* to achieve that end, as the *Guidelines* cited above requires.

---

[11] Oftel set out its views in *Ramsey Prices and Network Externalities: Dr. Rohlfs' Analysis*, 23 May 2002: "Ramsey prices, including the implications of externalities, are relevant in theory. But for practical reasons they are unlikely to provide a reliable basis for setting regulated charges."

[12] In "Ramsey pricing – Oftel's response to letter of 4 July" dated 12/7/2002.

[13] Oftel published a consultation document on continuing licence conditions after 25 July 2003 until the new law come into effect at http://www.oftel.gov.uk/publications/licensing/2003/cont0703.htm

The theoretical disagreements relate to how regulation should be designed to be compatible with competition policy.[14] Article 13(2) of the EU Access Directive (2002/19/EC) requires that "National regulatory authorities shall ensure that any cost recovery mechanism or pricing methodology that is mandated serves to promote efficiency and sustainable competition and maximise consumer benefits." An economist would interpret this as requiring, if possible, the regulation to simulate the effects of a competitive market. For a lawyer like Justice Moses matters were more complicated: "…the term 'maximum economic efficiency' is itself protean. … The concept involves value judgements … As Professor Geroski of the Competition Commission says, at paragraph 175 of his first statement, there is no unique definition of economic efficiency and no necessary dichotomy between efficiency and the principles of equity which the Commission sought to apply." (Moses, J. 2003, at para 122).

This argument seems somewhat disingenuous. Competitive markets deliver efficiency, not equity or fairness. Equity is better pursued by explicitly designed redistributive taxation, and under reasonable assumptions (set out in Diamond and Mirrlees, 1971), taxation should not distort pre-tax market prices, except for correcting for externalities. Under additional and also reasonable assumptions taxes on goods and services should be uniform, with redistributive taxes concentrated on income. This optimal form of redistributive taxation leaves relative market prices exactly as they would have been under competition with no taxes. In practice, it is more effective to redistribute real income through the expenditure side of the budget.[15] More to the point, the spirit of the Communications Directive and, more generally, of achieving consistency between competition policy and regulation, is that regulation should only be employed where markets do not deliver competitive outcomes. Where regulation is required it should attempt to produce the result that markets would have delivered if they had been competitive.

To return to mobile call termination, economic principles provide useful guidance for setting the price cap to simulate the effect of a competitive market. First, the efficient way of recovering fixed and common costs is Ramsey pricing for all the various services, in which the mark-up over LRIC is inversely proportional to *market* demand elasticities. Second, in a market where subscribers must choose between different suppliers to purchase the whole range of bundled services, effective competition in *all* service markets would force the suppliers to charge Ramsey prices for each service offered (i.e. subscription, making and receiving calls). Third, termination is a bottleneck service and thus not subject to the same competitive pressures as subscriptions and calling. Mobile operators can raise termination charges to fixed operators who are subject to regulation and unable to exercise countervailing

[14] There are other significant disputes to do with matters of fact, of which the most important is the size of the fixed and common costs, which in turn relates to the extent of economies of scale in mobile networks.

[15] Subsidising goods consumed by the poor normally leads to a larger subsidy to the rich, who buy more of the good, whereas equal lump-sum transfers (in kind or in cash) avoid this drawback.

power, and will have incentives to do so if the surpluses above cost thereby generated can be applied to fund competitive activities against other mobile operators.

Mobile operators also have incentives to raise charges to each other, since this disadvantages callers from other networks and encourages them to switch networks to the MNO charging the higher termination fee. This tends to over-reward mark-ups on termination relative to the efficient level. As a result, competitive forces will not ensure that MNOs set termination mark-ups to either fixed or mobile operators at Ramsey levels without regulation. Finally, new subscribers confer a network externality on existing mobile and fixed line subscribers, which they will undervalue in making the subscription decision. There is therefore a case for subsidising subscription relative to the incremental cost, and this adjustment forms part of the identification of appropriate Ramsey mark-ups.

If the remaining mobile markets are competitive, and if call termination charges are regulated at the Ramsey level, then all mark-ups would be set efficiently as though all services were competitive. Oftel disputed that the retail market for mobile services was competitive in its submission to the Competition Commission, but also accepted in its Market Review that the retail market was effectively competitive, in the sense that no MNO has SMP. That does not mean that the market corresponds exactly to the textbook definition of perfectly competitive (which, *inter alia*, requires constant or decreasing returns to scale, ruled out by the very need to recover fixed and common costs by a mark-up). It does mean that there are good reasons for not regulating that market.

This raises an important philosophical point. A welfare economist might argue that given one instrument (the mark-up in the regulated market), and one objective (maximise social welfare subject to private ownership, i.e. a profit constraint), the mark-up should take into account all market failures, including possibly imperfect competition in the mobile retail market. The whole thrust of the Communications Directives is that regulation cannot be justified if markets are deemed effectively competitive. That would seem to imply that regulators should not be able to leverage their authority to regulate some markets into other markets that are not subject to regulation (because they do not exhibit SMP). As regulation is costly (particularly in dynamic markets), it is only justified where there is no suitable competition law remedy. Competition law does not aim to make all markets perfectly competitive by assuming that it is costless to intervene, but leveraging regulation into other markets to improve their workings would seem to come very close to trying to achieve efficiency in these other markets which are almost, but not quite, competitive.

Another way of putting the same point is that the key principle of British regulation of network utilities is that competition should be encouraged where possible, and where natural monopoly precludes this, regulation should mimic the effect of competition. Price-cap regulation was developed precisely for this reason. Once it is accepted that Ramsey pricing would in fact mimic the workings of competitive markets for bundled services, it would follow that Ramsey pricing is appropriate (indeed, required) for the service (termination) that requires regulation. The principle has the additional advantage that as competition develops in the unregulated markets (through the market maturing, under competition from new

services such as 3G, or as a result of competition policy), there is no need to re-adjust the form of regulation in the call termination market.

Assessment on regulation and competition policy in telecoms

The Communications Directives have had a considerable impact on the form of telecoms regulation in Britain by ending the licence regime, and restricting the scope of regulation to markets suffering from Significant Market Power, and then only where regulation is required to achieve efficiency and sustainable competition. The evidence from telecoms liberalisation is that competition is indeed the key determinant of performance (Newbery, 2003) and so it seems natural to ensure that regulation does not impede the workings of otherwise competitive markets. The next question is whether the same approach would work as well for other network utilities. The most challenging such utility is the electricity supply industry, where, in contrast to telecoms, liberalisation is viewed with increasing scepticism in Europe and North America.

Electricity regulation and competition policy

Electricity markets differ in one obvious way from telecoms markets. Telephone users wish to speak to a particular person, whereas electrons are anonymous and are equally attractive from any source. This distinction makes vertical integration much less important.[16] The fact that economies of scope between generation, transmission and distribution appear small suggests that it is desirable to separate the potentially competitive generation and supply (or retailing) functions from the natural monopoly high and lower tension networks. These networks should then be regulated The updated Energy Directive (2003/54/EC), which came into effect in August 2003, requires member states to set up national regulatory authorities, and to impose regulated Third Party Access on transmission and distribution. The traces of earlier objections from some member states remain in that the regulator can either set tariffs *ex ante* or publish the methodology under which they can be set. Nevertheless, the principle that charges for using the networks should be regulated has now been widely accepted. The relevant question is whether the potentially competitive segments of generation and supply should be dealt with under normal competition law, as in many countries, subject presumably to tests of significant market power, or whether they should be subject to stronger regulatory controls.

Where the generation market is dominated by an incumbent, as in France (where the state-owned EdF controls about 95% of generation), or in Belgium (where the private company Electrabel has a similar share), it is clearly easy to demonstrate SMP. Tacit regulation, in which the threat of an investigation for abuse of market power, may then be sufficient. Many electricity markets are, however, characterised by the largest generator having less than 40% of the market, while still remaining fairly concentrated. The concept of

[16] There is a similar distinction between the circuit-switched networks used for voice telephone, where end-to-end continuous real-time connectivity is required for service quality, and packet-switched networks for data and internet, where vertical integration is no longer so valuable.

single dominance is then not sufficient, and something more inclusive is required. Under Article 82, a dominant position can be held by one or more undertakings ('collective dominance'). The jurisprudence on establishing collective dominance is still evolving, but it is now accepted that "a dominant position may be held by two or more economic entities which are legally and economically independent of each other." (*Guidelines*, 3.1.2.1).

The Court of First Instance gives as an example of conditions conducive to collective dominance: "the relationship of interdependence existing between the parties to a tight oligopoly within which, in a market with the appropriate characteristics, in particular in terms of market concentration, transparency and product homogeneity, those parties are in a position to anticipate one another's behaviour and are therefore strongly encouraged to align their conduct in the market, in particular in such a way as to maximise their joint profits by restricting production with a view to increasing prices." (Case T102/96, *Gencor v Commission* [1999] ECR II-753).

This list of conditions conducive to collective dominance maps almost exactly on to the characteristics of wholesale electricity markets. Consider as the leading example the English Electricity Pool, set up as the wholesale market for England and Wales at the time of restructuring and privatisation in 1989/90. The Pool continued until it was replaced by the New Electricity Trading Arrangements in 2001. The Pool was a compulsory, day-ahead, last-price auction in which generators had firm rights to transmission but no firm obligations to generate. Generators submitted bids the day before and the System Marginal Price (SMP) was set each half hour as the bid price of the most expensive unconstrained generation unit called on to operate. Unconstrained generators were all paid the same SMP plus a capacity payment, making up the Pool Purchase Price, PPP. Generators available but not dispatched also received the capacity payments, while constrained generators received their bid price, if they were required to generate within an import-constrained zone and bid above the SMP, or their lost profit (SMP *less* bid price) if they were in an export-constrained zone. The prices for each half-hour were published by 5pm the day before so that consumers could adjust any consumption decisions in the light of these prices.

The Pool prices were overlaid with financial contracts to reduce risk. The typical contract was a two-sided Contract for Differences (CfD), with a strike price s for a volume M MWh of electricity. If the Pool price were p in that time period, the buyer paid p per unit to the Pool and $(s-p)M$ to the company issuing the contract, thus effectively ensuring delivery of the contracted quantity of electricity at the strike price. The generator similarly received p per unit dispatched from the Pool, but was assured of receiving in total the strike price per unit contracted. It is readily proved that the best bid for a fully contracted generator is the marginal cost of generation, m. The generator would receive $(s-m)M$ total profit if he generated, and a larger amount, $(s-p)M$, if $p < m$ and he were not called to generate.

At privatisation, the Central Electricity Generating Board of England and Wales had been unbundled into a separate transmission company, National Grid, two fossil-fuel generating companies, National Power and PowerGen, and a nuclear generating company, Nuclear Electric, which remained in public ownership until 1996. The two fossil generating companies set the price of electricity in the Pool over 95% of the time in the first four years,

but as they had been sold almost fully contracted, they had little incentive to raise prices until they came to renegotiate the contracts over the next three years. As the original contracts expired and had to be renegotiated, so they raised the prices relative to the rapidly declining fuel prices.

Fig. 1 shows the evolution of the wholesale electricity price. The cost of generating by coal and gas are also shown, so the margin of the price over the short-run avoidable fuel cost is readily seen. The prices are yearly moving averages of the half-hourly pool prices or the quarterly fuel prices, and are deflated by the Retail Price Index to the 2001 price level. The line with diamonds gives the Hirschman-Herfindahl Index (HHI), defined as the sum of the squared percentage shares of coal-fired plant capacity (which set the price most of the time until 1999) owned by different companies, read on the RHS scale. An HHI of 5,000 represents a symmetric duopoly, and the initial concentration was somewhat higher than this as National Power was 50% larger than PowerGen.

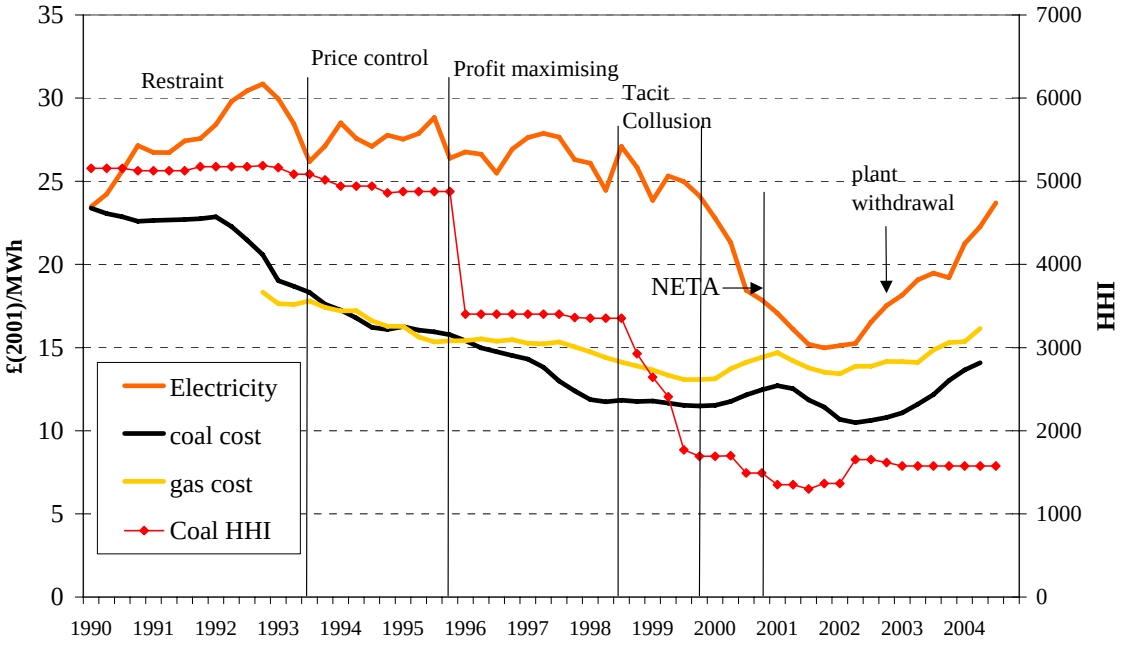

Fig. 1 Evolution of the wholesale electricity market in England and Wales

The half-hourly and daily prices were dramatically more volatile than these annual average figures, and the numerous price spikes attracted particular attention. The early Pool inquiries by the Director General of Electricity Supply, the electricity regulator, found that the price-cost margin was justified (Offer, 1992), but by 1994, it was clear that the continuing fall in coal prices was not leading to a comparable fall in the Pool price. The subsequent inquiry concluded with a price cap on the annual average Pool price, to be ended by the sale of 6,000 GW of plant, as shown. That sale lowered concentration (from an HHI of just under 5,000 to just under 3,500), but did nothing to reduce the price-cost margin. Part of the explanation was that the plant had been sold under a leasing arrangement to the buyer, who had to pay £6/MWh for

each unit sold. The defence of this was partly one of sharing risk (as the payment is only made when the plant successfully sells electricity) and partly to pay for the sulphur credits handed over with the stations, but had the effect of raising rivals' costs, encouraging them to support a high bid-price strategy in the Pool.

The Electricity Pool up to mid-1998 would seem to satisfy most of the criteria for collective dominance that would be likely to lead to tacit collusion. The market was concentrated (HHI above 3,000) and transparent (all prices were published the day ahead, and anyone could subscribe to Pool data giving the bids of each participant for each day). It was mature (with very low rates of growth); the product, electricity, was homogenous, produced by firms with similar or even identical plant and hence very similar cost structures. These costs were also transparent (mainly fuel costs whose prices were published). The demand elasticity was very low (even for those large customers who faced the Pool price), while a large fraction of electricity was sold at prices that were kept constant over periods of a year, despite huge variations in half-hourly prices.[17] The only criteria for likely collective dominance that the Pool failed to satisfy was that barriers to entry were low, and indeed entry was rapid. However, entry was by gas-fired plant running on base-load with contracts to distribution companies, and mostly bid a zero price into the Pool. They therefore did not participate in price setting. Over time the increase in capacity would undermine the prospects for tacit collusion, but the impact was gradual and not noticeable for the first seven years.

The other evidence to examine for tacit collusion is whether the prices remain high even when costs fall, and that was certainly the case. Fortunately, the extraordinary wealth of data about bidding behaviour in the Pool makes it possible to test for tacit collusion directly. Sweeting (2001) does this using bid data from each firm to see whether their bids are best responses to the bids of other firms. Fig. 1 divides up the time since privatisation into the relevant periods. During the first, from 1990-94, the duopolists offered to supply more at each price than would have been individually profit maximising, so if they were tacitly colluding it was to drive the prices lower, not higher. This is consistent both with the high initial contract coverage, and of tacit regulation. The duopolists might have reasoned that the exercise of market power would attract adverse regulatory activity, as it finally did in 1994. The two years of price control demonstrated the considerable ability of the companies to meet the price cap to within 1% despite a huge price spike caused by the loss of a nuclear station at the time the French decided to import 2,000 MW instead of, as normal, exporting 2,000MW to England.

The period following the end of the price cap until 1998 was a period in which each of the three firms appeared to be individually profit maximising, without any need for collusion (given their considerable individual market power). Finally, the last period for which Sweeting had data covers the period of rapid sales of generation plant by the incumbents, as they traded horizontal market power for permission (granted by the relevant Government minister) to

[17] Customers holding CfDs nevertheless had an incentive to respond to the Pool price, as the contract were financial and would involve payments regardless of actual consumption. Patrick and Wolak (1999) find very low price-elasticities (on average –0.025) even for large customers facing Pool prices.

vertically integrate into supply. During this period concentration fell rapidly, spare capacity increased, and bidding behaviour was consistent with tacit collusion to reduce supply at each level of price. There was a strong motive for this, as the generators wished to sell their plant for the highest price, and to do that they needed to convince prospective buyers that margins would remain high despite the excess capacity and falling market concentration. Once the sales had been made, the new owners, less familiar with the previous market game, bid more aggressively and competitively, thus undermining collusion and causing the price-cost margin to collapse.

Addressing cases of collective dominance

Electricity wholesale markets thus exhibit characteristics that are likely to support collective dominance, and specifically allow prices to remain well above competitive levels, even when the market does not appear to be particularly concentrated. In some cases (1996-8 in Britain) high price-cost margins can be sustained even without tacit collusion, given the technical characteristics of electricity. These include the very low demand elasticity, the inability to store electricity, and the effect of transmission constraints in fragmenting the market. If the capacity is not instantly available to deliver into a constrained region, then the generators operating in that region can demand extremely high prices. Even relatively unconcentrated markets can exhibit collective dominance, as the English Pool demonstrated for a brief period after 1998.

Competition authorities ought therefore to be able to establish collective dominance for many electricity markets. On the Telecommunications Directive model, it would then be open to argue that regulation were required to address the problem, if there were no better solution. The obvious competition policy remedy would be to require companies to divest generation capacity in order to reduce market concentration. There are difficulties with this structural solution, and attractions in retaining licence conditions for generation and supply. The obvious difficulty with structural remedies is that the burden of proof must be set reasonably high, if competition law is not to chill the very driving force of competition, which is the pursuit of profit. It is one thing to deny a merger that might increase concentration to the point that there were real concerns about future collective dominance, and quite another to move in the opposite direction, and argue that an existing industrial structure is suddenly inappropriate. There is good evidence that on average mergers destroy total shareholder value, and since they rarely increase competition, that must mean that they typically reduce *social* value. The burden of proof is then on the merging partners to demonstrate otherwise. In the present case, the burden of proof would lie squarely with the competition authorities, and might be difficult to establish.

The main problem of establishing an abuse of a collectively dominant position is to define the counterfactual of normally competitive behaviour. In electricity, variable costs are only half the average cost, the balance being the return to the very substantial fixed capital, as well as the fixed operating costs. A competitive market would need very high prices as the margin of spare capacity reduced, to compensate for prices being close to variable costs when capacity were adequate. Some commentators have argued that there are severe market failures on the demand side, in that final consumers cannot collectively exhibit their willingness to pay for the right degree of risk of loss of load (Stoft, 2002). Competitive wholesale markets may therefore under-provide security of supply without additional institutional requirements. The

English Pool had a capacity payment to encourage adequate reserves, but that was criticised for distorting the market (Offer, 1998). The initial proposals for the US Standard Market Design (for electricity markets) required Load Serving Entities to contract up to four years ahead for adequate capacity (i.e. enough for peak demand plus the specified reserve margin).

Large electricity generating companies argue that market power provides the ability to finance and hold adequate reserve margins, and leads to less price volatility and hence more politically sustainable outcomes than a market as fragmented, competitive and un-remunerative as the British market after 2000. There the combination of initially excess capacity and fierce competition caused the bankruptcy of a number of generators. The largest single generating company, British Energy, that holds all the privately owned nuclear plant had to be bailed out by the Government to prevent a similar fate. In response, generators have withdrawn capacity, causing concerns that there would be power shortages in Winter 2003/4. The reserve margin fell from over 32% to under 17% within a very short space of time, although the forward prices then induced mothballed plant to be returned to service and no shortages occurred (Roques, Newbery and Nuttall, 2005). On this view, too much competition is undesirable, and market power is required to deliver secure supply.

There are further problems, for if prices are above variable costs, the question is over what time period the profits should be averaged to determine whether average prices are above average total costs. Plant lasts 30 or more years, so a run of a few years of low prices might require rather high prices for an extended period to compensate. The generators under investigation could also argue that if there were a short-run problem of high prices, free entry would correct the problem, as the market is contestable.

There are other obvious problems with the competition policy route to creating workably competitive electricity markets. Where the industry was initially in public ownership, it was presumably open to the Government to restructure to increase the number of competing generation companies. If anything, most Continental governments moved in the opposite direction, consolidating previously separate companies to create 'national champions', as in Spain, and as was proposed for the Netherlands (but fortunately resisted). Where generation companies were privately owned, as in Germany, there was little or no resistance to proposed mergers that rapidly concentrated the European electricity industry (Codognet et. al., 2002).

The contrast with the United States is also instructive. European countries lacked the duties and powers available to the Federal Energy Regulatory Commission, FERC, under the 1935 Federal Power Act. FERC has a statutory obligation to ensure that wholesale prices are "just and reasonable". If an electric utility wishes to sell at market-determined wholesale prices, this will be allowed providing "the seller (and each of its affiliates) does not have, or has adequately mitigated, market power in generation and transmission and cannot erect other barriers to entry."[18] Even then, the authority to sell at market-determined prices can be withdrawn and replaced by regulated prices if there is "any change in status that would reflect a

[18] *Heartland Energy Services, Inc*, 68 FERC • 61,223, at 62,060 (1994), cited by Bogorad and Penn (2001).

departure from the characteristics the Commission has relied upon in approving market-based pricing."[19]

FERC therefore assumes that market pricing is "just and reasonable" so long as it is competitive. The reason for its concern to ensure that prices remain competitive is that any FERC-approved form of pricing greatly restricts the competition authorities from intervening. At the same time, existing antitrust laws are relatively powerless to enforce competitive outcomes in the energy industry as "the antitrust laws do not outlaw the mere possession of monopoly power that is the result of skill, accident, or a previous regulatory regime. ... Antitrust remedies are thus not well-suited to address problems of market power in the electric power industry that result from existing high levels of concentration in generation." (DOE, 2000, cited by Bogorad and Penn, 2001). This last point is particularly telling, and suggests that structural remedies may be difficult to secure.

If competition policy lacks the power to restructure (except in egregious cases of abuse), the next recourse would seem to be a regulatory solution. Britain provides an example, where the Director General imposed a price cap on wholesale prices until the companies divested some plant to improve competitiveness. Price caps on markets where prices are supposed to be determined by supply and demand only work if there is substantial market power, and the ability to moderate the supply price in the face of effectively a perfectly elastic demand schedule. However, the Californian electricity debacle of 1999-2000 demonstrates that poorly designed price caps can make matters far worse when there is genuine shortage combined with market power (Joskow and Kahn, 2002; Wolak, 2003). There are also problems where trading takes place in multiple markets (e.g. bilateral, OTC and balancing markets as well as power exchanges), and where the power exchange only handles a few percent of total production. Capping one market when there are uncapped close substitutes is likely to have perverse effects. The English Pool price cap applied to the annual average price, so some of these problems were avoided, and the fact that all electricity had to be traded through the Pool made the prices capped representative, avoiding the bypass problem. Regulators, mindful of the Californian lessons and the limited coverage of electricity wholesale markets in most countries, are likely to be wary of simple price caps.

The Latin approach to market power is to require bids into wholesale markets to be cost-justified (i.e. subject to audit), and to be maintained for lengthy periods (possibly indexed to relevant fuel prices). That is the approach followed in Chile, where generation is relatively concentrated (Bitran and Serra, 1998; Heller and McCubbins, 1996). System operators in some US power pools reserve the right to substitute cost-based bids for market bids if the latter lead to an inefficient dispatch, or clear signs of market power. There are several problem with this approach. The market design has to find a method of paying for, and charging consumers for, capacity. The resulting price formation rapidly becomes a regulatory artefact rather than a competitive market, with the players optimising against the regulatory price rather than competing to deliver best value. In concentrated markets the results may be

[19] *Heartland* 68 FERC at 62,066, cited as above.

less bad than unregulated price formation, although the latter might at least attract entry with the prospect of eventually reducing concentration.

A preferable alternative might be to require that generators be willing to offer contracts at prices related to the average cost of new entrants. That has the attraction of leaving the spot market to reflect supply and demand, while encouraging consumers to seek contracts where they are more attractive than relying on the spot market (i.e. when the spot market has prices higher than the entry price). The higher the degree of contract cover demanded, the less the incentive of the generators to exercise market power in the spot market. Wolak (2001) has suggested this as a remedy for market power in California.

The other alternative is to retain the licenses that the Telecommunications Directives abolished, recognising that electricity differs in important ways from telecoms. License conditions can require detailed operating information to be supplied to the regulator for better market surveillance, and can also impose additional restraints. Thus British generators above a certain size were required to offer for sale plant that would otherwise be scrapped, to ensure that the scrapping was not used to reduce capacity to increase prices above the competitive level. Grid codes in some countries specify rules for proper market behaviour and are clearly important. They may be an adequate substitute for licenses, although much depends on the institutional arrangements for modifying and enforcing them.[20]

Licenses can be used to prevent companies vertically integrating or having interests in related activities (such as transmission or distribution for generation companies), and to include more restrictive interpretations of market abuse than normal competition legislation. Again, Britain provides an interesting and salutary example. Fig. 1 shows that although market concentration was falling quite rapidly by 1999, the price-cost margin remained high. In May 1999 Ofgem (the gas and electricity regulatory agency) published a decision document about price spikes in the Pool during winter 1998/99, warning that it would continue to monitor prices. In July 1999 Ofgem investigated the high prices at the beginning of that month. Ofgem was concerned about the limited competition in price setting despite the increase in the number of generators selling through the Pool (from 8 in 1990 to 38 in 1999). The gas-fired plant of the new entrant Independent Power Producers did not compete at the margin and only rarely set prices (3 percent of the time in 1998/99). Three companies (National Power, PowerGen and Eastern) set Pool price 86 percent of the time in 1998/99. Ofgem expected the recent divestments to increase competition in price setting, but remained concerned about the ability of certain generators to influence the price setting mechanism (Ofgem, 1999).

In October 1999 Ofgem concluded that a Market Abuse Licence Condition (MALC), better known as the “good behaviour condition”, needed to be introduced into the licenses of

[20] The English Pooling and Settlement Agreement is a case in point, as it was a multilateral contract that could only be changed by mutual agreement, and over which the regulator had limited powers. It proved so difficult to adapt that it was eventually abolished by primary legislation which also replaced the Pool with the New Electricity Trading Arrangements. The new Balancing and Systems Code has a properly designed procedure for making modifications.

the seven generators most likely to have market power. This provided that "The Licensee shall not engage in conduct, whether alone or with one or more other undertakings, which amounts to an abuse of a position of substantial market power in the determination of wholesale prices for electricity under the relevant trading arrangements." Possible examples of such abuse would be price bidding strategies, capacity withholding, manipulation of complex market rules and using the influence of contractual positions.

To minimise regulatory uncertainty about the operation of the condition, Ofgem issued Guidelines about its interpretation. These provided that a Licensee will be regarded as having a position of substantial market power if it has the ability to bring about, independently of any changes in market demand or cost conditions, a substantial change (over £30 million) in wholesale electricity prices. Substantial changes in price might refer to a few very large effects or a series of lesser ones. For example, it would include a change of 5 per cent for more than 30 days (1440 half hours), or 15 per cent over 10 days (480 half hours) or 45 per cent over 160 half hours (3 1/3 days), all within a one year period (Ofgem, 2000a). It is worth noting that the required increase in prices would amount to about one-third of one per cent of turnover in the wholesale electricity market.

After a lengthy discussion process lasting until April 2000, five generators consented to the condition: Magnox Electric, TXU Europe (formerly Eastern, the purchaser of the 1996 divested coal-fired plant), Edison Mission Energy, National Power/Innogy and PowerGen. Two generators (AES and British Energy) did not consent, and Ofgem referred them to the Competition Commission in May. Ofgem argued that other possible remedies and developments (further divestment or revised generation market structure, modification of Pool and NETA rules, the introduction of NETA itself, the Competition Act 1998 and Financial Services Regulation) would not suffice to prevent such abuse. This was because of the specific conditions of electricity: the need to match demand and supply instantaneously, the non-storability of electricity and the limited demand side response.

Shortly afterwards, Ofgem carried out its first investigation under the market abuse condition, in the licence of Edison First Power (Ofgem, 2000b). The company had withdrawn 480MW of capacity from the system. Ofgem concluded that the company had substantial market power, which it had exploited to the detriment of consumers. Specifically, the continued withdrawal of capacity was not justified on the basis of avoiding losses, and had materially increased prices in the Pool and the forwards markets. This was to the detriment of over 200 large customers purchasing electricity on Pool-related terms, and other respondents were harmed by the increase in Pool Uplift. Because the company announced that it would return the capacity to the system, Ofgem took no further action.

The Competition Commission published its decision in December 2000 and rejected Ofgem's requirement that AES and British Energy's licences be modified to include the MALC (CC, 2001). The Competition Commission was critical of Ofgem, noting that it did not define the relevant market and did not demonstrate that it would be profitable for either company to attempt to exercise market power by withdrawing capacity. AES had a long-term contract with TXU that made reducing output unprofitable, while 90 per cent of British Energy's potential output came from inflexible base-load nuclear power plant. The

Competition Commission considered that a condition prohibiting abuse of market power "would cause uncertainty, because of the difficulty of distinguishing between abusive and acceptable conduct, and would risk deterring normal competitive behaviour." (CC 2001, 1.12). "We are mindful of the disadvantages of a broad, effects-based prohibition … such a prohibition would not be suitable for dealing with manipulation of market rules." (para 1.18). "We see manipulation of the market as conduct for which a sufficient remedy would in principle be the modification of the market rules or mechanisms." (para 1.13).

Assessment of regulatory policy in electricity

There is no doubt that a specialist energy regulatory authority is valuable in dealing with the natural monopoly networks, as the new EU Energy Directive recognises. The issue is whether generation and supply should also be regulated or whether normal competition law is adequate. The *AES and British Energy* case is instructive. The case required a time extension as the Competition Commission found it very difficult to understand the complexities of the wholesale electricity market, even though Ofgem supplied a large amount of explanatory material. Electricity market surveillance is a critical element in ensuring that market participants do not abuse positions of collective dominance, and such surveillance requires timely information that is commercially sensitive and not normally made available without a licence (or grid code) condition making it a condition of operation. Market surveillance needs to be a continuing activity if market participants are to be assured of non-abusive behaviour, and it needs to be analytically well-founded if market rules found wanting are to be improved.

All this suggests that the energy regulatory authority should take on the task of market surveillance, and should be adequately empowered to monitor the relevant markets, and make rule changes. Competition authorities are likely to be far less well-equipped to do this, and will be far slower to act. The Californian electricity industry went from crisis to claimed bankruptcy in the space of a few weeks, while FERC made regulatory decisions that, according to an expert witness "allowed a manageable problem to develop into an economic disaster" (Wolak, 2003). Whether regulators need additional powers in addition to being able to make rule changes will depend on market circumstances. Britain suggests that the ability to impose conditions (price caps, cost-based contracting or even cost-based bidding) might be a useful lever to achieve market restructuring, although in the hands of a less trusted regulator might discourage private investment in generation, which might be far worse than accepting some market abuse.

Conclusions

Post-modern utilities like telecoms, in which facilities-based competition is possible, lend themselves to the approach laid out in the EU Telecommunications Directives. Regulators must first conduct market reviews to determine whether markets are effectively competitive or suffer from Significant Market Power (SMP). Only in the latter case is *ex ante* regulation warranted, and then only if it is necessary, justified and proportionate. Electricity does not fit comfortably into this approach, in part because the wholesale market is more likely to suffer

from collective dominance than the single firm dominance that allows for an easier determination of SMP. Licence conditions are no longer needed under the Telecommunications Directives, but appear to retain advantages for electricity, where information and continuing market surveillance are desirable, and where it may be necessary to modify market rules in a timely and well-informed manner.

Competition policy remedies are likely to be too slow in electricity, where rapid, knowledgeable and intelligent actions may be necessary to avoid very costly outcomes. That said, one should not underestimate the difficulty of ensuring that wholesale electricity markets work well, in delivering competitive outcomes that are sustainable in the long run. FERC, with extensive powers under the 1935 Federal Power Act, failed to avoid the Californian crisis, and Ofgem's Market Abuse Licence Condition was found unsuitable for at least some of its target clients. The fall in concentration coming after excess entry induced by the earlier exercise of market power caused a collapse in British wholesale electricity prices and a subsequent withdrawal of plant that halved the reserve margin and has caused concern for supply adequacy in the winter months. All this suggests that ensuring workable competition in electricity supply continues to remain a challenge to regulators and competition authorities.

References

Armstrong, M. 1996, 'Network Interconnection', DPET 9625, Southampton

Armstrong, M. (1998), 'Network Interconnection in Telecommunications', *Economic Journal*, 108 May, 545-564

Armstrong, M., S. Cowan and J. Vickers, (1994) *Regulatory Reform - Economic Analysis and British Experience*, Cambridge: MIT Press.

Bartle, I. (ed.) (2003) *The UK Model of Utility Regulation: A 20th anniversary collection to mark the 'Littlechild Report' – retrospect and prospect*, Bath: Centre for the study of Regulated Industries

Baumol, W.J., (1983) 'Some subtle pricing issues in railroad regulation', *International Journal of Transport Economics*, 10: 341-55

Baumol, W.J., J.C. Panzar, and R.D. Willig, (1982) *Contestable markets and the theory of industry structure*, Sydney and Toronto: Harcourt Brace Jovanovich, Academic Press

Bell, A. (1995) `The Telecommunications Industry 1994/95, ch 5, pp 93-104 in *CRI Regulatory Review 1995*, ed. P. Vass, London: CIPFA

Bitran and Serra, (1998), 'Regulation of Privatized Utilities: The Chilean Experience', *World Development*, Vol.26, No.6, pp.945-962.

Bogorad, C.S. and D.W. Penn (2001), 'Cost-of-service rates to market-based rates to price caps to ?!#?#!?', *The Electricity Journal*, May, 61-72.

Bollard, A. and M. Pickford (1996) 'Utility regulation in New Zealand', London: Institute of Economic Affairs *Lectures on Regulation*

CC (2001), *AES and British Energy*, Report 453, London: Competition Commission

CC (2003), *Mobile phone charges inquiry*, London: Competition Commission

CEC (2002), 'Commission guidelines on market analysis and the assessment of significant market power under the Commission regulatory framework for electronic networks and services', *Official Journal*, 11.7.2002

CEC (2003), 'Commission recommendations of 11 February 2003 on relevant product and service markets within the electronic communications sector susceptible to *ex ante* regulation in accordance with Directive 2002/21/EC of the European Parliament and of the Council on a common regulatory framework for electronic networks and services', *Official Journal*, L 114/45-49; 8.5.2003

Codognet, M-K., J-M. Glachant, F. Leveque and M-A. Plagnet (2002) *Mergers and Acquisitions in the European Electricity Sector*, Paris: CERNA, Ecoles des Mines.

Diamond. P.A. and Mirrlees, J.A. (1971) 'Optimal Taxation and Public Production, Part 1: Production Efficiency' *American Economic Review*, 61(1), 8-27.

DOE (2000) *Horizontal Market Power in Restructured Electricity Markets*, U.S. Department of Energy Office of Economic, Electricity and Natural Gas Analysis

Hausman, J.A. (1997), 'Valuing the effect of regulation on new services in telecommunications', *Brookings Papers: Microeconomics*, 1-38

Heller and McCubbins, (1996), 'Politics, Institutions and Outcomes: Electricity Regulation in Argentina and Chile', *Policy Reform*, Vol.1, pp.357-387.

IEA (1994) *Natural Gas Transportation: Organisation and Regulation*, IEA/OECD, Paris.

Joskow, P. and E. Kahn (2002), `A quantitative analysis of pricing behavior in California's wholesale electricity market during summer 2000', *Energy Journal*, Dec, pp1-35.

Laffont, J-J, P. Rey and J. Tirole, 1998a, Network Competition: Overview and Non-discriminatory Pricing, *Rand Journal of Economics* **29**(1):1-37

Laffont, J-J, P. Rey and J. Tirole, 1998b, Network Competition: Price Discrimination, *Rand Journal of Economics* **29**(1):38-56

Levy, B and P. Spiller (eds.) (1996) *Regulations, Institutions and Commitment*, Cambridge: CUP

Littlechild, S. (1983) *Regulation of British Telecommunications profitability*, London, HMSO, reprinted in Bartle (2003) .

MMC (1999) *Cellnet and Vodafone*, London, Monopolies and Mergers Commission

Newbery, D. M. (1997) `Determining the Regulatory Asset Base for Utility Price Regulation', *Utilities Policy*, **6**(1), 1-8.

Newbery, D.M. (2003) 'Privatising network utilities', mimeo, Cambridge (paper for CESifo conference on privatisation, Cadenabbia, Oct 31-Nov 3, 2003)

Offer (1992), *Review of Pool Prices*, Office of Electricity Regulation, Birmingham, December 1992.

Offer (1998) *Review of Electricity Trading Arrangements: Background paper 1 - Electricity Trading Arrangements in England and Wales*, February, Birmingham: Office of Electricity Regulation

Ofgem (1999) *Rises in Pool Prices in July: A Decision Document*, October, London: Office of Gas and Electricity Markets

Ofgem (2000a) *Introduction of a 'market abuse' condition into the licences of certain generators, Ofgem's initial submission to the Competition Commission*, May, London: Office of Gas and Electricity Markets

Ofgem (2000b) *Ofgem's investigation of Edison First Power under the market abuse licence condition: Initial Findings*, July, London: Office of Gas and Electricity Markets

Oftel (2002) *Otel's market review guidelines: criteria for the assessment of significant market power*, London: Oftel, 5 Aug

Oftel (2003) *Review of mobile wholesale voice call termination markets: EU Market Review* London: Oftel, 15 May

Patrick, R.H. and F.A. Wolak (1999), 'Customer Response to Real-Time Prices in the England and Wales Electricity Market: Implications for Demand-Side Bidding and Pricing Options Design under Competition' in Crew, M. A., (ed.) *Regulation under increasing competition*. Boston; Dordrecht and London: Kluwer Academic, p155-82.

Peltzman, S. (1976) `Toward a more general theory of regulation', *Journal of Law and Economics*, 19:211-240.

Roques, F.A., D.M. Newbery and W.J. Nuttall (2005) 'Investment incentives and electricity market design: the British experience', *Review of Network Economics*, **4** (2), 93-128

Stigler, G.J. (1971) `The Theory of Economic Regulation', *Bell Journal of Economic and Management Science*, 2:3-21.

Stoft, S (2002). *Power System Economics*, New York: Wiley-Interscience.

Sweeting, A. (2001) 'The effect of falling market concentration on prices, generator behaviour and productive efficiency in the England and Wales electricity market', MIT Department of Economics.

Temin, P., with Galambos, L. (1987) *The fall of the Bell system: A study in prices and politics*, Cambridge; New York and Sydney: Cambridge University Press.

Vickers, J. and G. Yarrow (1988) *Privatization: An economic analysis*, Cambridge: MIT Press

Willig, R.D. (1979) 'The theory of network access pricing', in H.M. Trebing, ed. *Issues in Public Utility Regulation*, Michigan State University Public Utilities Papers

Wolak, F.A. (2001) 'A comprehensive market power mitigation plan for the California electricity market', California ISO Market Surveillance Committee, Apr. 24 available from http://www.stanford.edu/~wolak

Wolak, F.A. (2003) 'Diagnosing the California Electricity Crisis', *The Electricity Journal*, Aug/Sep, 11-37.