The Pros and Cons of Antitrust in Deregulated Markets

Konkurrensverket

Swedish Competition Authority

Konkurrensverket
SE-103 85 STOCKHOLM
tel +46 8 700 16 00
fax +46 8 24 55 43
konkurrensverket@kkv.se

© 2004 Konkurrensverket/
Swedish Competition Authority
individual chapters © 2004
The contributors
ISBN-91-88566-34-X
Printed in Sweden by
AB Danagårds Grafiska
Ödeshög 2004

Preface

"The Pros and Cons of Antitrust in Deregulated Markets" is the third in the Swedish Competition Authority's Pros and Cons series following last year's "The Pros and Cons of Low Prices" and "The Pros and Cons of Merger Control" from 2002. The book will be officially released on November 12, at a seminar in Stockholm where the authors will present their work and senior officials from competition authorities around the world will act as discussants.

I would like to express my gratitude to all the authors who have contributed; without you we would not have a book at all. At the Swedish Competition Authority, our chief economist Mats Bergman has been the editor and Arvid Nilsson has managed the project; they both deserve due credit.

Stockholm, September 2004

Claes Norgren

Director-General

Table of contents

The contributors

Martin Cave is a Professor and Director of the Centre for Management under Regulation at Warwick Business School, University of Warwick, UK. He is co-author of Understanding Regulation (1999), co-editor of the Handbook of Telecommunication Economics (2002) and author of the Review of Radio Spectrum Management (2002). He is advisor to Ofcom and the Postal Services Commission and a non-executive advisory director of OFWAT.

Peter Crowther is an English qualified solicitor at LeBoeuf, Lamb, Greene & MacRae, and is also admitted to the Brussels Bar. Peter received his PhD from Copenhagen Business School, after completing his Masters in Law and Economics at Oxford University. Peter has advised on the application of competition law and related regulatory matters across a wide range of industries including retail, telecommunications, insurance, upstream oil and gas, energy distribution and supply, banking and capital markets, commercial finance, railways, sea transport, and motorsport. Prior to entering private practice Peter was a lecturer in EC law and Competition law. Peter has published widely and is a Section Author of Butterworths Competition Law.

Michael Harker is a member of the ESRC Centre for Competition Policy and a lecturer in Norwich Law School at the University of East Anglia. His main research interests are in the fields of UK utility regulation (which was the subject of his doctoral thesis) and competition law. In the coming years he will be participating (in collaboration with fellow members of the Centre) on research projects on consumer representation and remedies in mergers cases. In addition to this, he also pursues research interests in public law, particularly public law litigation concerning regulation.

Catherine Waddams Price is founding Director of the ESRC Centre for Competition Policy and Professor in the School of Management at the University of East Anglia, UK. Her main research interests are in the reform of utility industries, the introduction of competition into previously monopolised markets, and the effect of these changes on income distribution, both in developed and developing countries. She has advised the World Banks and regulatory bodies and government in the UK, Brazil, Argentina and the Netherlands and is a member of the UK Competition Commission.

Alison Oldale is Principal at LECG. Dr. Oldale specialises in using economic theory and empirical techniques to analyse the competitive impact of mergers and joint ventures, company conduct and agreements between companies. Dr. Oldale has advised companies in a wide range of industries, including fast moving consumer products, financial markets, transport, industrial goods and she has particular expertise in analysing telecommunications markets. Her merger work has involved providing advice and expert reports during investigations before the MTF and various national authorities. She has advised companies during competition law infringement cases involving various pricing abuses, as well as both vertical and horizontal agreements. In the telecommunications sector Dr.

Oldale has offered advice and submitted reports during consultations and proceedings involving competition authorities and telecommunications regulators. Dr. Oldale was educated at Cambridge University and the London School of Economics where she earned a PhD in Economics. She published in the European Competition Law Review and the Journal of Economic Theory and has presented papers at both academic and industry conferences.

Atilano Jorge Padilla is Managing Director at LECG. Jorge belongs to the Competition Policy group and is based in Madrid. He joined LECG in March 2004. During his years in consultancy Jorge has advised clients on a variety of competition policy and intellectual property issues, covering a wide range of industries, including retail electricity, media, telecommunications, entertainment and technology. Dr. Padilla earned M. Phil and D. Phil degrees in Economics from the University of Oxford. He is a Research Fellow of the Centre for Economic Policy Research (CEPR, London), and a member of the Steering Committees of the Nyenrode Institute for Competition (The Netherlands) and the Association for Competition Economist (ACE). Jorge has written several papers on competition policy and industrial organization in the Antitrust Bulletin, the International Journal of Industrial Organization, the Journal of Economics and Management Strategy, Journal of Economic Theory, the RAND Journal of Economics, the European Competition Law Review and World Competition. He is also a regular speaker at competition policy conference in Europe and the United States.

Frank Wolak is a Professor of Economics at Stanford University. His fields of research are industrial organization and econometric analysis. He specializes in the study of privatization, competition and regulation in network industries such as electricity, telecommunications, water supply, natural gas and postal delivery services. He is a Research Associate of the National Bureau of Economic Research and a Visiting Researcher at the University of California Energy Institute. Wolak is also the Chairman of the Market Surveillance Committee for the Independent System Operator of the California Electricity Supply Industry. He has testified several times at the Federal Energy Regulatory Commission (FERC), and at various Committees of the US Senate and House of Representatives on issues relating to market monitoring and market power in electricity markets.

1. Introduction

This volume, "The Pros and Cons of Antitrust in Deregulated Markets", is about the intersection of competition law and sector-specific regulation. When is competition law sufficient and when is sector-specific legislation necessary? What are the advantages of relying only on competition law? And which are the drawbacks?

Although the authors mainly discuss energy and telecom markets, the principles they base their discussions on are of a general nature. They all subscribe to the view that competition is desirable and that markets should be liberalised, rather than monopolised. Despite this, they hold different views on the necessity of complementing competition law with sector-specific regulation. According to some, competition law is sufficient in deregulated markets; according to others, the special properties of certain markets makes it necessary to introduce specific regulatory measures.

In the first chapter, Martin Cave and Peter Crowther discusses the relationship between sector-specific regulation and competition law, with a particular focus on the distinction between *ex ante* and *ex post* intervention and the new EU regulation concerning electronic communication. The authors begin by observing that, in principle, a main difference between competition law and sector-specific law is that Articles 81 and 82 are applied ex post, while sector-specific law is applied ex ante. In practice, however, this distinction is less clear-cut than one would think. The EU Commission, in addressing violations of Article 81 and 82, has many times negotiated agreements with the parties or the concerned government(s) and reached agreements which have in effect amounted to an ex ante regulation of the industry. For example, before France Telecom, Deutsche Telecom and Sprint were allowed to form the joint venture *Global One,* the Commission reached an agreement with the French and the German governments to liberalise the telecom markets of the two countries. In addition, the incumbent telecom operators had to consent to a number of detailed conditions and obligations akin to those typically placed on regulated companies, including third-party access. Another technique used by the Commission is to codify permissible behaviour in so-called Communications. Although formally and legally non-binding, these are ex ante recommendations on how to behave in order not to risk being challenged for violating the antitrust rules.

In this way, the EU Commission has been able to use competition law to actively enhance and improve regulation that could be regarded as defective. Often, these initiatives have been followed by subsequent regulatory reform and, Cave and Crowther argue, the Commission's ability to devise innovative solutions has been augmented by the experience it has gained in antitrust cases. The authors also note that the Commission's powers in this respect were extended in Regulation 1/2003 (the "modernisation").

The main conclusions of the chapter appear to be that competition law and regulation are complementary and that there is no "bright line" between the two. The authors point to some potential problems with using competition law to put regulation in place. For example, there is a risk that such regulation will be ad hoc, whereas full-blow regulatory reform could be based on a larger base of knowledge and experience. On balance, however, the authors appear to be in favour of the current situation, where

competition law complements sector-specific regulation, in particular in areas where desirable reform is stalled for political reasons.

The second article in the volume, by Michael Harker and Catherine Waddams Price, focus on gas and electricity retailing in the UK, outlining recent development and the regulatory strategies employed.

When monopoly and price regulation gave way to competition and competition law, it was expected that the incumbents would have to face a choice between high market shares or premium prices. However, more than five years after liberalisation, the gas incumbent and the regional electricity incumbents are still able to maintain approximately 60 per cent of their respective "home" markets, despite selling at a considerable premium over the entrants. Harker and Waddams Price report that, despite this, the regulator is still confident that market-share erosion will lead to a well-functioning market.

In the UK, the energy regulator, Ofgem, has concurrent power with the competition authority, OFT, to apply the national correspondences of Articles 81 and 82 in energy markets. Harker and Waddams Price argue that Ofgem has been hesitant to challenge behaviour by the incumbents that possibly could have been seen as examples of abuse of dominance. For example, London Electricity used win-back offers to induce consumers to switch back from entrant suppliers and British Gas offered discounts to customers that signed direct debit contracts. These types of targeted discounts, or instances of price discrimination, reduce the incumbents' cost of maintaining a price premium, since many of the customers that would have switched to another supplier will be able to use the offers to obtain better prices from the incumbents. In addition, fixed-term contracts have been used to raise consumers' switching costs.

The authors suggest that the regulators reluctance to challenge price discrimination and predatory pricing could have the paradoxical result that the regulator ends up protecting the energy incumbents from competition law challenges that otherwise would have come from the competition authorities, in the same vain as industry regulators often have been accused of siding with the industry, rather than with the consumers. They are also somewhat pessimistic as to the effect of "consumer voice", in the form of consumer organizations and government bodies. Since re-regulation is not an attractive alternative, the main avenues towards better-functioning energy retail markets appear to be active consumers and an active and innovative application of general competition law.

In the third article, Alison Oldale and Jorge Padilla argue with force that the current regulatory situation in the European telecom market is an unfortunate one. The begin by a quick tour through the history of telecom regulation since the Second World War, but focus on the so-called New Regulatory Framework (NRF) for telecom, which was enacted in July 2003. The NRF has two main objectives: achieving regulatory harmonization within the Union and promoting the development of competition, so that regulation can give way to general competition rules. The latter objective is to be achieved primarily through the promotion of facilities-based competition. In addition, sound principles from competition law, such as methods for market delineation and for evaluating market dominance, have been introduced into the regulation for electronic communication.

However, Oldale and Padilla argue, the talk about stimulating facilities-based competition is just talk. In practice, the NRF favours access-based entry, market fragmentation and micro-management of the telecom

industry, despite it paying lip-service to facilities-based entry and the principles of competition law, including a preference for relying on market-based solutions as far as possible.

Oldale and Padilla then address what they call "the Investment Ladder". This can be seen as an intellectual attempt to reconcile what can be perceived as a grave inconsistency in the NRF: that while one of its goals is to promote un-regulated facilities-based competition, in practice, it goes far in requiring access to incumbents' infrastructure. According to the "investment-ladder theory", access is necessary in order to achieve the first stage of competition: access-based competition. Only when such competition has given the entrants a foothold can these new firms be expected to invest in infrastructure. The regulator's task is then to slowly lift the incumbents' obligations, so as to make it more profitable for the entrants to build their own infrastructure. Eventually, as the dominant's market power and its control of bottleneck facilities has eroded far enough, sector-specific regulation can be dismantled completely. The authors object strongly to this theory, on the grounds that such micro-management of an industry is beyond the abilities of the regulator and that this type of policy will lead to a fragmentation of the industry. The result will be a huge number of firms that will not be willing to undertake major investments. The authors conclusion is that general competition law is sufficient to ensure healthy competition, while interventionistic regulation, such as the NRF, will be detrimental to the long-run development of the market.

Frank Wolak focus on the electricity market and, in contrast to the other contributions, draws mainly on the US experience. In a sense, his main conclusion is exactly the opposite of that of the previous article. He argues that antitrust law will be ineffective as an instrument against the exertion of unilateral market power in the electricity market and that, therefore, antitrust laws must be complemented with sector-specific legislation. The underlying reasons are the special properties of the electricity market: supply must equal demand at every instant and at every location, electricity is very costly to store, capacity constraints are absolute, demand is inelastic, production is large-scale and the industry is concentrated.

Because of these special properties, power-generating companies will often be in a position to exert *unilateral* market power, possibly resulting in prices far above the competitive level. Since antitrust law (in particular US antitrust law) is mainly concerned with *coordinated* actions (and mergers), it will typically not be able to address these types of concern. A possible response would then be to regulate prices, but that will lead to well-known problems. Another alternative would be to prohibit "market manipulation"; an alternative that has been tried in the US and elsewhere. The inherent problem with such regulation is that it seeks to curb the firms' natural tendency to maximize profit. But at which point does profit-maximizing behaviour become illegal?

Wolak's proposed solution is the implementation of a sector-specific regulation with three main ingredients. First, a requirement that electricity-generating firms provide extensive information on their market activities, including outages, to the regulator and to the public. Second, a set of well-defined rules for market behaviour, designed so as to support a well-functioning market and backed by penalties and sanctions with a sufficient preventive effect. Third, and perhaps most controversially, a "local market power mitigating mechanism". According to this mechanism, such behaviour that is detrimental to system reliability and market efficiency *and*

where the behaviour is intentional, should be considered illegal. In order to establish the effect of a certain type of behaviour, a thorough analysis is necessary. When this is done, the regulator will announce its view on the matter. Only if the firm persists, despite being given a warning, will a particular behaviour be considered as intentional.

Wolak points to two possible drawbacks with the proposed scheme. First, the regulator may yield to political pressure, e.g., by opportunistically exploiting the sunk-cost nature of power-plant investments. Secondly, extensive information dissemination in combination with a system that involves negotiations may result in price coordination. However, for the reasons given above, he arrives at the conclusion that the general antitrust rules are insufficient to deal with the particular problems of the electricity market and that, consequently, sector-specific regulation is necessary. Wolak ends the article by observing that "industry-specific regulators need not fear for their jobs, because there is much for them to do in the future".

Collectively, the four authors are able to draw on a number of real-world experiments in market designs as well as on their own contributions to the literature on regulated and unregulated markets. The papers in this volume demonstrate that it is not easy to evaluate the pros and cons of regulation, deregulation and competition policy. On the other hand, it is possible to make costly regulatory mistakes. For this reason, it is important to develop the regulatory framework so that effective competition, efficiency and consumer welfare is promoted as far as possible. This volume will help regulators to continue to improve their understanding of how regulation works – and when it does not work.

Mats Bergman

Editor

2. Co-ordinating regulation and competition law – ex ante and ex post

Martin Cave and Peter Crowther

2.1 Introduction

In this paper we investigate the roles of *ex ante* and *ex post* intervention under regulatory policy and competition law. It is widely recognised that competition law is intended to act as a deterrent to unlawful behaviour, and to that extent operates *ex ante*, but here we investigate the degrees to which regulators rely on *ex post* interventions, and to which competition law has become 'regulatory', prescribing patterns of behaviour on undertakings by way of formal or informal settlements. We also make observations on the desirability of this intermingling of *ex ante* and *ex post* competition law and regulation and *ex ante* and *ex post* interventions.

Our focus is on regulated sectors, and the starting point of our analysis in Section 2.2 is the regulatory regime applying to electronic communications services in the European Union. This sector is of particular interest because the regulatory regime is designed to approximate to competition law in ways described below, because it specifically addresses the issue of where regulation and competition law have relative advantages, and because NRAs in choosing remedies to deal with market power can under the sector specific legislation adopt pure or hybrid *ex ante* and *ex post* remedies.

In Section 2.3 we shift to competition law, showing how, in a number of cases, the Commission has adopted regulatory-style *ex ante* approaches. The implications of the trends that emerge from Sections 2.2 and 2.3 are then explored in Section 2.4.

2.2 Ex ante and ex post approaches in electronic communications services regulation

With the exception of New Zealand,[1] where until recently there has been no sector-specific regulation, the relationship between regulation and competition law has to some extent been serendipitous. However the new European regime for the regulation of electronic communication services, which NRAs are now having to implement, has attempted to impose logic

[1] For more recent developments in New Zealand, see n. 19.

on the relationship. We begin with a brief account of this regime, especially of how the scope of *ex ante* regulation is chosen by NRAs[2]; then consider how the "allowable interventions" are divided between *ex ante* and *ex post*.

Outline of a new regime

A new regulatory framework for electronic communications was adopted in 2002. The main objectives of the new framework are to simplify the previous regimes, to apply them in a technologically neutral manner, and to encourage competition while guaranteeing user rights. Certainly, the previous regime has been streamlined, through a reduction from twenty key Community law measures to just five[3]. At one level, the new régime is a major step down the transition path between regulated monopoly and normal competition, governed exclusively by generic competition law. As a result of the new regime, NRAs are no longer able to regulate the sector by issuing individual licences. Subject only to certain limited exceptions, Member states are required to establish a general authorisation regime. The conditions that may be imposed are heavily circumscribed. The new regime's provisions are applied across the range of "electronic communications services", ignoring pre-convergence distinctions. It represents an ingenious attempt to corral the NRAs down the path of normalisation – allowing them, however, to proceed at their own speed (but within the uniform framework necessary for the internal market).

Since the end state is envisaged to be one governed by competition law, the Commission proposes to move away from the rather arbitrary and piecemeal approach of the previous regulatory package towards something consistent with that law. However, competition law is to be applied (in certain markets) not only in a conventional responsive *ex post* fashion, but in a pre-emptive *ex ante* form. The new régime therefore relies on a special implementation of the standard competition triple of: market definition, identifying dominance, and formulating remedies to deal with (anticipated) competition law breaches. We examine these in turn[4].

According to the underlying logic of the legislation, the European Commission first establishes a list of markets where *ex ante* regulation is

---

[2] For further details, M. Cave "An economic analysis of remedies in network industries" in D. Geradin (ed) *Remedies in Network Industries: EC Competition Law vs. Sector Specific Regulation,* 2003, pp. 1-20.

[3] Directive 2002/19/EC on access to, and interconnection of, electronic communications networks and associated facilities [2002] OJ L 108/7 *(Access Directive);* Directive 2002/20/EC on the authorisation of electronic communications networks and services [2002] OJ L 108/21 *(Authorisation Directive);* Directive 2002/21/EC on a common regulatory framework for electronic communications networks and services [2002] OJ L 108/33 *(Framework Directive);* Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services [2002] OJ L 108/51 ("Universal Service Directive"); Decision No.676/2002/EC of 7 March 2002 on a regulatory framework for radio spectrum policy in the European Community ("Spectrum Decision").

[4] The first two of these processes are elaborated in respectively, Commission Recommendation 2003/311/EC 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 on a common regulatory framework for electronic communication networks and services. [2003] OJ C 114/45 ("Recommendation"); and Guidelines on Market Analysis and the assessment of SMP ("Guidelines On Market Analysis") [2002] OJ C 165/6. Remedies are the subject of a Common Position on the approach to Appropriate remedies in the new regulatory framework, 2004, ("the Remedies Paper") by the European Regulators Group, the college of NRAs created by the legislation to, *inter alia*, advise the Commission on the implementation of the procedures.

permissible, the markets being defined according to normal competition law principles. These markets are then adapted and analysed by NRAs with the aim of identifying dominance (on a forward-looking basis). Where no dominance is found, *ex ante* obligations may not be imposed on any undertaking in the relevant market (*ex post* competition law would still apply). Where dominance is found, the choice of an appropriate remedy must be made from a specified list. The effect of the regime is to create a series of market-by-market "sunset clauses" which reduce the level of *ex ante* regulation as the scope of effective competition expands.

Market definition

In February 2003 the Commission issued a Recommendation on relevant markets[5], to identify those markets which, in the Commission's view, may warrant *ex ante* regulation. Unlike the previous regime, markets must be defined in accordance with the principles of competition law[6]. NRAs may vary the markets subject to objection by the Commission. The Recommendation incorporates flexibility by allowing related "technical services" to be aggregated within a market definition. Member states can also add or subtract markets[7], using specified (and quite complex) procedures.

The Recommendation identifies three cumulative criteria for identifying those markets which are suitable for *ex ante* regulation: high and non-transitory barriers to entry, the expected persistence of such barriers to entry over a relevant time period, making the prospect of effective competition unlikely, and the inability of competition law adequately to address the particular issue. The second of these is simply a projection of the first (albeit difficult to apply in practice). The logic of the régime thus rests heavily on the combined operation of the first and third criteria.

The Recommendation identifies structural barriers and legal or regulatory barriers as being relevant for analytical purposes. The least problematic are those created by legal or regulatory rules. A legal barrier might take the form of a requirement that firms have a licence in circumstances where additional licences are not available. Alternatively, there may be a legal limitation on the availability of a particular input which is necessary to produce a relevant electronic communications service. These two considerations come together in the case of wireless markets, for which operators require both a licence and access to spectrum. As a consequence, entry by network providers into wireless (fixed and mobile) markets may be effectively blocked. A legal barrier does not mean that dominance will be found, as competition behind the barrier can be effective.

A regulatory barrier arises when, as a result of regulatory policy or previous practice, entry into a particular market is made to be financially unprofitable and this situation is expected to persist. This state of affairs arises, for example, when the NRA, in pursuit of other objectives – typically relating to ensuring the affordability of retail services – imposes a retail pricing structure which means that some services are individually provided at below cost.

---

[5] Op cit, n.3.

[6] Framework Directive, Article 15(1).

[7] Markets are defined using standard competition law criteria.

Structural barriers essentially consist in cost or demand characteristics that create an asymmetric relationship between incumbents and entrants. The examples given in the Recommendation are economies of scale and/or economies of scope and high sunk costs. Economies of scale do not of themselves create a barrier to entry that cannot be overcome by entrants. For scale-related barriers to exist, it is necessary that the economy of scale operate over a range which is large in relation to the market as a whole. It is also necessary that exit from the market be costly because the firm must incur substantial sunk expenditures which are not recoverable on exit; otherwise the market would be contestable[8].

A further structural barrier to entry may be associated with demand-side network effects. Network effects in certain circumstances can give an operator an enduring competitive advantage over its smaller rivals, causing the market to "tip" in its favour. However, such markets may still be subject to competitive pressures over the longer term as competitors seek to replace the dominant operator through technological advances. Other regulatory measures, such as requirements for interconnection, interoperability and end-to-end connectivity, may alleviate the risk of tipping.

The economic literature has for some time considered the presence of "strategic" barriers – that is, barriers to entry that may be artificially "manufactured" by the firm which enjoys them. Strategic barriers to entry include many types of possible conduct. They may take the form of excessive investment in product capacity, R & D or advertising, which creates in the minds of entrants the rational expectation that entry will be subject to a strongly competitive response by the incumbent. It may also raise rivals' costs, for example by seeking to pre-empt inputs, raising the price of inputs, or artificially inflate consumers' switching costs by contractual or other means. The Recommendation adopts the position that these are best dealt with by competition law.

The third, cumulative, criterion is whether competition law is sufficient to address the particular market failure. The Recommendation does not address this question at length, offering an illustration of a comparative advantage for regulation in situations where compliance requirements are high, intervention is frequent or where legal certainty is required. Similarly, the Framework Directive offers little assistance. Recital 27 provides that:

> "It is <u>essential</u> that ex ante regulatory obligations should only be imposed where there is not effective competition... and where national and Community competition law remedies are not sufficient to address the problem." (Emphasis added)

The Recommendation and the Guidelines on Market Analysis offer substantial guidance on whether there is effective competition, but little on how an NRA is to satisfy itself (as it is required to do) whether competition law is "insufficient" to resolve the particular market failure. This difficult criterion goes to the very heart of the relationship between competition law and regulation and we return to it in Section 2.4.

---

[8] On contestability, see W. Baumol et al *Contestable markets and the theory of industrial structures* 1982 and (more recently), M.A. Bergman *Potential competition: theory, empirical evidence and legal practice,* working paper, Swedish Competition Authority, 2002.

Dominance

Pursuant to Article 16 of the Framework Directive, the regulatory framework only permits the imposition of *ex ante* regulation where one or more undertakings is found to have Significant Market Power (SMP).

According to Article 14(2) of the Framework Directive:

> "An undertaking shall be deemed to have [SMP] if, either individually or jointly with others, it enjoys a position equivalent to 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".

This definition of SMP is of course identical to the standard definition of dominance determined and repeated by the European Court of Justice, ensuring in principle a major step forward towards the convergence of approaches under regulation and competition law. However, by way of extension of the competition law concept of dominance, if an undertaking is dominant on a specified market, it may be "deemed" (Article 14(3), Framework Directive) to have SMP on a closely related market.

The Commission's Guidelines on Market Analysis contain the principles to be used by NRAs in determining whether an undertaking has SMP.[9] Essentially, to determine whether one or more undertaking has SMP (i.e. whether effective competition is absent), NRAs must evaluate the structural conditions on the relevant market. Where the analysis indicates an absence of effective competition, the NRA must then examine whether the market may be "prospectively competitive".

Remedies

Under the Directives, NRAs have the power to impose obligations on firms found to enjoy SMP in a relevant market. Essentially, for wholesale markets the remedies are contained in Articles 9-13 of the Access Directive, while for retail markets the remedies are contained in Articles 17-19 of the Universal Service Obligations Directive. The wholesale remedies are, in ascending order of rigour: transparency, non-discrimination, separate accounting, mandatory access, and cost-oriented pricing. NRAs must act within a framework of duties set out in Article 8 of the Framework Directive and the measures they take must be proportionate to the policy objectives identified. This can be construed as meaning that the intervention is appropriate, no more than is necessary, and, by implication, satisfies a cost-benefit test, in the sense that the expected benefits from the intervention exceed the expected costs. Article 8 additionally specifies policy objectives, but does not go so far as to determine the weights appropriate for use in the cost-benefit analysis. For example, Article 8(2) requires NRAs to promote competition for electronic communications networks and services by maximising users' choice and value for money, eliminating distortions or restrictions to competition and encouraging efficient investment in infrastructure. Further, Article 7(4) of the Framework Directive requires NRAs to promote the interest of EU citizens by, inter alia, providing

[9] Op cit, n.4.

consumers with protection in their dealings with suppliers, and requiring transparency of tariffs and conditions for the use of publicly available electronic communications services. NRAs must also contribute to the development of the internal market by avoiding different approaches to regulation within the EU. These provisions provide an important context in which NRAs must hone their interventions.

Mandating access and non discrimination

Here we focus on two major wholesale remedies, which crosscut with competition law.[10] The first relates to the conditions on and under which a competitor should have access to a dominant firm's assets, the second to the remedy of non-discrimination.

Facilitating access

A key element in any telecommunications remedy is likely to be the mandating of access to a facility by competitors. The Access Directive contains two such remedies. In essence, one mandates access while the other mandates access at cost oriented prices. The two remedies are now described in more detail.

Article 12 of the Access Directive entitles NRAs to impose on operators with SMP obligations to meet reasonable requests for access to, and use of specific network facilities. An NRA may impose obligations on operators to grant access to specific facilities or services, including in situations when the denial of access would hinder the emergence of a competitive retail market, or would not be in the end user's interests.

This represents an obligation to be implemented in circumstances similar to, but significantly broader than, those in which the essential facilities doctrine is applied under competition law. Simplifying somewhat, the extension of the test lies in the replacement of the precondition under competition law for mandatory access, that the asset is essential and cannot be replicated, by a much broader condition that NRAs can mandate access in circumstances where its denial `would hinder the emergence of a sustainable competitive market at the retail level, or would not be in the end-user's interest.'

There is a risk that the last phrase in particular (`would not be in the end-user's interest') might open the door for extensive regulatory intervention. The obvious problem is the conflict between users' short-run and long-run interests. Short-run interests might best be furthered by the adoption of mandatory access on a wide scale. However, such a policy clearly reduces incentives to invest by both incumbents and competitors, and is likely to stifle innovation in the long term. Article 12 (2) gives a list of factors to be taken into account by the NRA, including 'the initial investment of the facility owner, bearing in mind the risks involved in making the investment.'

The obligation is silent about the pricing of such access, except to the extent that it prohibits "unreasonable terms and conditions" having a similar effect as denial of access. "Not unreasonable" pricing may therefore

[10] Articles 17-19 of the Universal Service Directive also concerns (retail) remedies. The focus on access remedies reflects a belief that SMP principally resides in network or wholesale, rather than retail markets.

encompass a range between the long run average incremental costs of the access service provided and the retail price charged to end users minus an amount equal to the cost to the access provider of the services which the access seeker now produces (so-called "retail minus prices").

Article 13 of the Access Directive entitles NRAs to impose obligations of price control and cost accounting. In assessing costs, national regulators should take into account the investment made by the operator and the risks involved.

The circumstances identified as appropriate for the application of this remedy are "situations where a market analysis indicates that a lack of effective competition means that the operator concerned might sustain prices at an excessively high level, or apply a price squeeze, to the detriment of end users." There is, however, a major distinction between these two cases. In the case of excessive pricing, the question is whether the customer is being exploited directly through an excessive price. In the case of a margin squeeze, the hypothesis is that the operator with SMP is foreclosing competition in a vertically related area by pricing that fails to cover costs. The appropriate way of dealing with the latter issue is discussed below.

It is widely accepted that cost-oriented pricing for interconnection or access to customers should only be considered when dealing with an operator with SMP which is both persistent and incapable of being dealt with by other remedies, including particularly structural remedies. A classic case for its application might therefore be access to the local loop, either for call termination or for the purposes of leasing unbundled loops – provided of course that one operator enjoys a monopoly or position of dominance in the relevant geographical area.

The ERG's Remedies Paper points to the case of replicability of the assets as being a criterion which can be used to choose an appropriate remedy.[11] If an asset is immediately replicable, the market will not exhibit SMP. If an asset is virtually non-replicable in the medium term, Article 13 and cost-oriented pricing is appropriate, especially if there are no investment issues. Article 12 can be considered for intermediate cases.

Non-discrimination

Absent structural separation, the incentives on firms to discriminate against downstream (or upstream) competitors are significant. In European telecommunications NRAs and competition authorities are already faced with a large range of discriminatory behaviour by vertically integrated firms including:

- outright denial of access to the network;
- discriminatory or excessive pricing;
- failure to link prices properly to costs (reflecting the fact that many access price régimes have been based on traditional retail pricing models rather than underlying costs - e.g. per-minute interconnect pricing, per-message pricing for wholesale SMS access);
- margin squeezes;

[11] *Op. cit.*, n. 3.

- quality differentials;
- discriminatory wholesale offers;
- failure to ensure timely disclosure of vital information;
- failure to allow equivalent customer experiences.

Under Article 10 of the Access Directive an NRA may impose obligations of non-discrimination, in relation to interconnection and/or network access. Essentially, such obligations must ensure "equivalent conditions in equivalent circumstances to other undertakings providing similar services, and services and information to others under the same conditions and of the same quality as the operator with SMP provides for its own services, or those of its subsidiaries or partners."

An example of ex ante/ex post choices – wholesale broadband access[12]

We now present an example of how the issue of choice of *ex ante* or *ex post* intervention plays out in one particularly contentious market – wholesale broadband access, which as the Recommendation on Relevant Market notes, "includes what is traditionally referred to as bitstream services," adding that "for the purpose of this Recommendation, bitstream is a service which depends in part on the PSTN and may include other networks such as the ATM (asynchronous transfer mode) network". [13]

The network inputs required to provide broadband services over copper lines using a technology known as ADSL (asymmetric digital subscriber line) are shown in Figure 1;

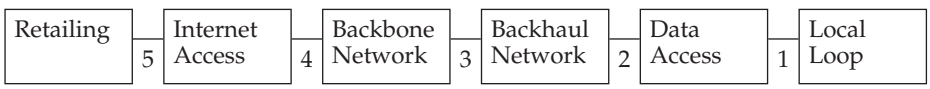

Figure 1: Broadband services

From the right, a customer is connected to the exchange by the local loop. She gets access to the data network at the local exchange using equipment known as "Digital Subscriber Line Access Mechanisms" or DSLAMs. Data are conveyed deeper into the network from the local exchange by a backhaul network, to the operator's backbone or core network which then gives access to websites around the world. The retailer provides marketing, billing, help-line and other functions.

Competition in broadband can take many forms. A cable operator, for example, can provide an end-to-end service. However, in the case of a rival telecommunication operator, such a policy would require it to duplicate the local loop. Accordingly, competitors usually purchase services from the incumbent at one of the number of points shown in the figure:

---

[12] The details of wholesale broadband access are discussed more fully in the Remedies Paper, *op. cit.*, n. 3 and M. Cave "Broadband remedies" *Review of Network Industries,* 5/1 (2004) pp. 23-50.

[13] Op. cit., n. 3.

1. Here the competitor leases an unbundled local loop from the incumbent, providing all other services itself.

2. The competitor uses the incumbent's data access facilities (DSLAMs) at the exchange as well as the loop.

3. The competitor also relies on the incumbent for backhaul into the core network.

4. The competitor also utilises the incumbent's network giving access to the World Wide Web.

5. The competitor simply "resells' (i.e. markets under its own name) the incumbent's complete broadband product.

How can regulation articulate the supply of a range of wholesale products of this kind?

Relative prices of access products

A key precondition for neutrality across different wholesale broadband products is satisfaction of relevant margin squeeze tests. An operator practises a margin squeeze in a vertically integrated production process where the price difference or margin between two vertically related products, one more inclusive than the other, fails to cover the (incremental) cost of providing the 'wider' service. The classic case is a retail margin squeeze, where firm A sells a wholesale product to firm B, which competes in the retail market with firm A. If the difference between the wholesale and the retail price fails to cover retailing costs, firm B will be squeezed out of the retail market.

This formulation begs many questions about the range of products over which the test should be conducted, the methodology for measuring costs, the firm whose costs are relevant to the test (firm A, firm B or some hypothetical efficient retailer) and the period over which costs should be calculated.

A prohibition of a margin squeeze is designed to ensure that prices are set in a way designed to prevent the dominant firm from leveraging its market power from one stage of the production process into a neighbouring one. Applying it consistently over a range of broadband wholesale (and retail) products should avoid exclusionary behaviour of this kind. In the case of wholesale broadband access, some regulators, such as Ofcom in the UK, prefer to use their competition law powers to police such squeezes, relying on a narrow interpretation of the application of the definition of bitstream in the Recommendation, while other NRAs extend the application of Article 10 to a broader range of product.

Absolute prices of access products

There then arises the question of how to tether the sequence of prices of access products, thus constrained by margin squeeze tests. Adding up all incremental costs will fail to provide a contribution to common costs (either costs common to several access products, or costs common to broadband access and other services). Firms will seek, and be entitled, to recover these

costs, and may also try to make additional profits. The tests will not prevent them from doing so because they only prescribe relative and not maximum prices.

The existence of SMP (necessary to trigger any of the remedies) implies independent price-setting ability. One natural manifestation would be excessive prices – for example implemented by the SMP operator identifying the key "least replicable" element in broadband access – for example bitstream – and extracting maximum profits from its position of dominance. In these circumstances, some form of price control may be appropriate – to sit alongside the various margin squeeze tests.

As noted above, two remedies in the Access Directive deal with controlling the price of access products – cost oriented pricing and mandatory access at reasonable prices. A combination of the latter with Article 10 (non-discrimination) would allow bitstream access to be priced at the incumbent's retail price minus retail costs, minus the cost avoided by the incumbent by not providing network elements supplied by the competitor, or the cost incurred by the entrant in doing so. In other words – a margin squeeze type test.

In this case, the theoretical and practical arguments in favour of cost-plus and retail-minus are finely balanced. Regulators may choose a path which takes them from the less informationally demanding retail minus approach, which may better reward risky investment by the incumbent and encourage investment by competitors, to cost plus at a later stage, if SMP becomes entrenched.

However, during the retail minus phase a further issue arises – whether to police that restriction on bitstream prices on an *ex ante* or *ex post* basis; in other words, whether to require all changes to bitstream and retail broadband or other access prices to be approved in advance by the regulators or subject only to *ex post* challenge.

In favour of the former, it is argued that competitors get fuller protection from potential abuses, that it solves the problem that they may not have the necessary information about the incumbent's retail prices to launch well-founded complaints, and that problems of pricing wholesale products not matched by a retail counterpart are guaranteed a solution.

In favour of the latter, it is argued that *ex ante* regulation imposes delays on price changes in a dynamic market and that the regulator can intervene by exception rather than across the board, thus reducing the regulatory burden.

Conclusions

This section has discussed an example of the interaction of competition law and regulatory approaches and of *ex ante* and *ex post* interventions in electronic communications services markets. Essentially, the new electronic communications services regime seeks at a high level to render to competition policy what is competition policy's (episodic interventions in potentially competitive markets) and to regulation what is regulation's (persistent interventions in markets with established dominance).

However, it becomes apparent from the broadband example in particular that the regulatory remedies, though available more broadly than under competition law, are in fact similar or the same, and that, at a level of detail, their implementation can be either *ex ante* or *ex post*. In other words, there is no "bright line" between the two, even when the regime is subject to comprehensive planning.

This makes it even more unlikely that a bright line will emerge from the competition law side, where the regime grows by (generally) unplanned accretion of precedents, punctuated by occasional hard or soft law. We illustrate this, with examples, in the next section.

2.3 Application of competition law in regulated industries

We now examine the relationship between competition law and regulation, as illustrated through a number of cases in the communications and energy sectors. Although we are by necessity not able to cover all regulated sectors, we believe the themes illustrated by the cases below are of general relevance.

At a simple level, competition law is applied *ex ante* through merger control, and *ex post* through Articles 81 and 82. In the former case, the Commission investigates whether a competition would be significantly impeded, while in the latter case the Commission examines whether an abuse has actually occurred. A number of trends illustrate that this simple distinction breaks down in practice. To provide a context for the discussion we begin by summarising a number of relevant cases. These cases have arisen against the background of a weak (or non-existent) regulatory framework to support competition.

Example cases

Global One[14]

EC competition policy has for a long time played a significant role in shaping the structure of regulated markets, most significantly to break up previously monopolised markets. In an early case in 1982 involving BT, the then UK monopoly telephone company, the European Court of Justice confirmed that the EC competition rules applied to the communications sector. Through subsequent radical use of EC Treaty provisions regarding competition, the Commission prised open the markets to achieve full liberalisation of infrastructure and service markets.

During the period prior to telecommunications liberalisation in Europe (i.e. pre 1998) the Commission was faced with an increased number of international joint ventures, as national incumbents prepared for liberalisation. The innovative approach taken by the Commission to the application of the competition rules is illustrated by the Global One case. This case involved a proposed alliance between France Telecom, Deutsche Telekom and Sprint (a U.S. carrier), with the aim of providing international advanced telephony services. However, the EU market was then still largely monopolised, and neither France nor Germany had committed politically to full liberalisation.

Before allowing the joint venture to proceed, the Commission required

[14] Decision 96/546/EC Atlas OJ [1996] L 239/23; Decision 96/457/EC Phoenix/Global One OJ [1996] L 239/57.

political and legal action by the French Government, the German Government, Deutsche Telekom and France Telecom. In essence, the Governments were required in large measure to liberalise their telecommunications markets, while the telecommunications companies were required to commit to a series of detailed conditions and obligations regarding their future activities. Many of the obligations were akin to those typically placed on regulated companies, and required the joint venture company and its shareholders to act as a benign network operator vis-à-vis third party service providers. Specific obligations included disclosure of technical specifications and commercial information, no bundling of services, and no cross-subsidisation. The parties were required to have their accounts audited to preserve the ringfencing measures and to keep accounts to verify compliance with all obligations. By taking this approach, the Commission effectively liberalised two national markets and imposed a new regulatory framework.

British Interactive Broadcasting[15]

British Interactive Broadcasting Limited ("BiB", subsequently renamed "Open") was a joint venture vehicle created by BT Holdings Limited (a wholly owned subsidiary of BT), BSkyB Limited (a wholly owned subsidiary of British Sky Broadcasting Group Plc), Midland Bank Plc and Matsushita Electric Europe (Headquarters) Limited (a wholly owned subsidiary of Matsushita Electric Co. Limited).

The joint venture posed a number of significant competition issues, as it involved the dominant telecom provider and the dominant satellite television channel provider joining forces to roll out a new service, interactive TV, across the U.K. The Commission identified five distinct markets for analysis:

(1) the market for digital interactive TV services, which it distinguished from (a) the market for PC based on-line services, (b) the market for retail pay-TV and (c) the market for high street retailing;

(2) the market for technical and administrative services for digital interactive and retail pay-TV services;

(3) the retail pay-TV market;

(4) the market for the wholesale supply of films and sports channels for retail pay-TV; and

(5) the local loop infrastructure market.

To allow the joint venture to proceed, the Commission required a substantial number of commitments that created a new regulatory environment within which the Commission considered BiB would not be able to act anticompetitively. The commitments were extremely wide ranging and detailed, involving legal separation of BiB box and services

[15] Decision 99/781/EC British Interactive Broadcasting/Open OJ [1999] L312/1.

operations, removal of subscription tie between BiB boxes and BSkyB services, removal of key elements of exclusivity, commitments by BSkyB to distribute channels either with or without interactive applications to prevent bundling, and by BT to divest its cable interests, disclosure of technical specifications, third party access to interactive services on non-discriminatory terms, and so on.

Gazprom/ENI

As part of a wider investigation into restrictive supply agreements, in October 2003 the Commission announced[16] that it had reached a settlement with Gazprom and ENI concerning the supply of natural gas from Russia into the EU. In a speech to the World Forum on Energy Regulation in October 2003, Competition Commissioner Mario Monti, stated:

> "[This case] shows that during the initial delicate transition phase from monopolised to liberalised energy markets, the focus should lie, in some occasions, on the Commission's interventions improving effectively the market structure, rather than on formal procedures imposing fines".

The Commission's informal "interventions" were not limited to requiring the parties to remove restrictive clauses and undertake not to repeat the offending activity. The Commission also required ENI to offer significant gas volumes to customers located outside Italy over a period of five years (using auctions where ENI fails to meet agreed targets), to increase the capacity in its TAG pipeline used to transport all Russian gas destined for the Italian market, and to offer an improved third party access regime facilitating the use of the TAG as a transit pipeline. As reflected upon by Mario Monti, this type of intervention amounts to direct structural intervention, going beyond "merely" requiring the cessation of arrangements that the Commission considers anticompetitive. This suggests a company may be required to create competition, as part of a settlement with the Commission.

The "Marathon" cases

The Marathon cases relates to a series of informal settlements reached by the Commission as a result of a complaint by Marathon Oil in connection with alleged referral to grant access to their gas networks by Gaz de France, Ruhrgas, BEB, Thyssengas and Gasunie. In the last of the cases, *Gaz de France and Ruhrgas,* in a Press Release announcing the settlement Mario Monti said:

> "These settlements include important improvements to the functioning of European gas markets and show that competition policy complements and reinforces sector specific legislation. Effective access to gas transport networks is essential for the introduction of competition on the European gas markets to the benefit of industrial users and ultimately consumers. I call upon Gaz

[16] IP/03/1345.

> de France, Ruhrgas and other gas network operators across Europe, in cooperation with the national authorities, to continue their efforts to create an effective Third Party Access regime."[17]

As a result of the settlement Gaz de France and Ruhrgas were required to improve third party access to their respective gas transport networks. The detailed commitments differ to reflect the market situation in each country, but it is important to note that the settlement resulted in a liberalised environment that goes further than the political agreement reached in the context of the Gas Directive[18]. This approach is reminiscent of the Commission's proactive use of Article 86(3) in the telecommunications sector, but is achieved through informal legislative techniques.

According to the settlement, Gaz de France was required to undertake a number of activities designed to promote competition, including extending its offering to rural areas, reducing its balancing zones, implementing an investment programme to reduce congestion on the network, offering gas conversion services to third parties, to enable more customers to receive gas from competitors, undertaking a gas release programme, reducing the capacity reservations made by Gaz de France's trading division, further developing the secondary market in transport capacity, and increasing transparency in handling access requests.

Ruhrgas was also required to undertake a significant programme of activities, including setting up a new capacity reservation scheme, introducing new tariff zones with charges on a cost causation basis, extending the regime to all group companies, offering refined balancing services, improving transparency, and introducing a "use-it-or-lose-it" principle into all its transport contracts, including the contracts with its own downstream operations.

Commentary

The above cases illustrate the wide and flexible powers of enforcement available to the Commission, to achieve solutions appropriate to the particular problem identified. Rather more challenging is the question of how to "classify" such cases in the context of the traditional distinction between regulation and competition law[19], to which we now turn.

Classification of competition policy application

It is conventional to characterise the application of competition law as being through "advocacy" and "enforcement". Advocacy refers to the process of engaging policy-makers and legislators, to ensure that the competition rules are applied effectively throughout the economy (e.g. through removal of unnecessarily protected monopolies). Enforcement refers of course to the application of the competition rules to private activity (agreements, abuse and merger control), to prevent anticompetitive behaviour.

---

[17] IP/04/573.

[18] Directive 2003/55/EC of 26 June 2003 concerning common rules for the internal market in natural gas and repealing Directive 98/30/EC [2003] OJ L176/57.

[19] See, in the context of electronic communications, A. de Streel, Remedies in the European Electronic Communications Sector in D. Geradin (Ed), Remedies in Network Industries: EC Competition Law vs. Sector-specific Regulation, 2004.

These cases suggest however that it may be appropriate to identify a third process, which may be described as "negotiation", referring to the process by which a Competition Authority reaches private agreement with the dominant company involved, to determine the fundamental structure of an industry according to competition policy (not regulatory) prescriptions.

Negotiation may also be distinguished from sophisticated forms of enforcement. A sophisticated form of enforcement may well require complex mechanisms designed to achieve termination of the objectionable practice and avoid its repetition in the future. For instance, in ensuring that a non-discrimination obligation is enforced in practice, it may be necessary to impose a reporting obligation on a dominant company (see e.g. Atlas/Phoenix/Global One).[20]

"Negotiated" settlements might therefore be classified as private agreements created by and between the Commission and the undertakings concerned, that result in an arrangement that (usually) extends or recasts the regulatory framework into an *ex ante* one. The cases discussed above may be considered examples of this.[21]

Another informal legislative technique employed by the Commission is the use of non-binding legislation and guidance, notably in the form of Communications, to describe what, in the Commission's view, would (for example) amount to an abuse. The *ex ante* effect of such documents means that it could be wrong to characterise them as merely bringing clarity to an *ex post* prohibition. Moreover, it seems possible to distinguish documents produced by the Commission summarising the Commission's understanding of the Courts' case law, from documents that are much closer to a an *ex ante* policy document one might typically expect to emerge from a regulatory authority. A relevant example of the latter would be the Commission's Recommendation on Unbundled Access to the Local Loop[22], produced jointly by DG Information Society and DG Competition.

The Commission's strengthened powers

The flexibility afforded to the Commission in applying the competition rules is evident from the above cases, particularly as the Marathon cases were informal settlements of alleged *ex post* infringements. The fact that the vast majority of Commission competition cases have historically lead to informal settlements may of course have contributed significantly to the Commission's ability to devise innovative solutions. In this context it is interesting to note that the formal powers in relation to remedies previously granted to the Commission through Regulation 17/62, were extended in Regulation 1/2003. Article 7(1), Regulation 1/2003 now provides (the italicised words are those added by Regulation 1/2003):

---

[20] *Op. Cit.*, n. 12.

[21] Interestingly the new "light touch" regulatory arrangements in New Zealand, adopted since 2001, have some of the characteristics of such a negotiation. When threatened with regulation, a company can offer undertakings, the effects of which are then assessed in a cost-benefit analysis against a fully specified counterfactual – what would unfold in the absence of regulation. Hence instead of having a remedy thrust upon it, the firm can calibrate its offering to pass the cost-benefit test, amending it as the enquiry proceeds. See Paula Rebstock, *The New Zealand Experience in Utility Regulation,* paper delivered at Australian Competition and Consumer Commission Regulatory Conference, July 2004.

[22] Recommendation 2000/417/EC 0J [2000] L 156/44.

> "Where the Commission, upon application or upon its own initiative, finds that there is infringement of Article 81 or Article 82 of the Treaty, it may by decision require the undertakings or associations of undertakings concerned to bring such infringement to an end. *For this purpose, it may impose on them any behavioural or structural remedies which are proportionate to the infringement committed and necessary to bring the infringement effectively to an end. Structural remedies can only be imposed either where there is no equally effective behavioural remedy or where any equally effective behavioural remedy would be more burdensome for the undertaking concerned than the structural remedy. If the Commission has a legitimate interest in doing so, it may also find that an infringement was been committed in the past."*

There are, therefore, two limits on the Commission's ability to impose remedies: (a) the remedy must be proportionate to the infringement and necessary to bring it to an end; and (b) a structural remedy may only be used where there is no equally effective behavioural remedy.

Proportionality is of course a well-established and fundamental principle of EC law. The Commission's Guidelines on Market Analysis, for instance, merely provide a general description of proportionality:[23]

> "In essence, the principle of proportionality requires that the means used to attain a given end should be no more than what is appropriate and necessary to attain that end. In order to establish that a proposed measure is compatible with the principle of proportionality, the action to be taken must pursue a legitimate aim, and the means employed to achieve the aim must be both necessary and the least burdensome, i.e. it must be the minimum necessary to achieve the aim". (Para. 118)

Yet despite the limited elaboration of proportionality, "Respect for the principle of proportionality will be a key criterion used by the Commission to assess measures proposed by NRAs under Article 7 of the Framework Directive" *(Ibid.)*

A further important development under Regulation 1/2003 is the formalisation of the previous informal procedure under which the Commission accepted informal commitments from undertakings. Article 9(1) allows the Commission to adopt "Commitment Decisions", under which the Commission may by Decision make binding the commitments offered by an undertaking, in return for the Commission's formal conclusion that there are no longer grounds for action. Depending on how this new power is applied, this may well have the effect of strengthening the *ex ante* effect of competition law. Finally, a further feature of so-called modernisation of EC competition law, namely the ability of national competition authorities to apply EC competition law, may also help to free Commission resources to focus on "big cases" where serious structural deficiencies are apparent.

---

[23] *Op. Cit.*, n. 4.

2.4 Evaluation

The question of the respective roles of competition law and regulation, and the related (but not identical) question of the appropriate areas for *ex post* and *ex ante* interventions will remain a key policy question in regulated industries. Indeed, a current practical consequence of the overlap of regulatory policy and competition law can be seen in wholesale international roaming, a service which telecommunications companies offer to operators in other countries, allowing customers of the latter to make and receive mobile telephone calls when they visit another country. Aside from being one of the relevant markets in the Commission's Recommendation on relevant markets for electronic communication services (potentially subject to *ex ante* regulation) and subject in that capacity to study within the ERG, it is the subject of a current European Commission competition inquiry (into the UK market). In this case, not only is the legislation different but the identity of the enforcement agency (the NRA in one case the Commission in the other) is different. Similarly, in most Member states, the body applying sector-specific law (the NRA) will be different from the body applying competition law (the national competition authority – NCA).

Through a discussion of a number of cases, the previous sections have attempted to assess the continued importance and relevance of the conventional distinction between *ex ante* regulatory intervention and *ex post* competition law enforcement.

In the case of electronic communications services, the system has been devised to impose the distinction, with *ex ante* regulation limited to a subset of markets, "recommended" by the Commission, with opportunities for the NRAs to add to that list, subject to meeting the three criteria noted above to the Commission's satisfaction. (Remaining markets are subject to competition law only.) Where dominance is found in a market subject to *ex ante* regulation, NRAs can choose remedies from an approved list, subject to the requirements of proportionality. We have noted above that because many of the remedies, especially those associated with the avoidance of discrimination, are identical with similar proscriptions in competition law, the NRA may face a choice between *ex ante* and *ex post* enforcement. This blurs the line between regulatory and competition law approaches, but not in a major way.

The sector-specific communications legislation has explicit policy objective on which NRAs and the Commission can rely in formulating interventions. These include the "promotion of competition". Articles 81 and 82 are associated with no such explicit objectives, and this might lead to the adoption of alternative approaches when competition law is applied in regulated sectors.

Our analysis of a number of cases has suggested, however, that the Commission has been prepared to "negotiate" pro-competitive outcomes relying upon a battery of *ex ante* interventions which are very much in the regulator's arsenal, including mandated access and legal or accounting separation, as well as structural remedies such as divestiture. Generally, the cases have arisen in circumstances where the regulatory regime can reasonably be regarded as defective, as evidenced by its subsequent reform. This is the case in relation to BiB, where access to conditional access and related service was subject to the Advanced Television Services Directive; to

the Global One case, where markets were subsequently liberalised under the 1998 telecommunications package; and to the energy cases where the relevant Directives have been to some degree emasculated by the political process. Thus "regulatory antitrust" can be seen as making up for the deficiencies of regulation.

However, the use of competition law in such contexts does raise some further important questions:

- since one might expect private companies not to cooperate fully in the pursuit of a competitive market, does a negotiated settlement without the benefit of expert regulatory resources risk a sub-optimal outcome from a policy perspective?
- how should the success of an intervention be measured compared with the alternative?
- when would a regulatory-type intervention be disproportionate?

There can be no doubt that the Commission has been pro-active in its enforcement of the competition rules. A key policy question is whether, in driving forward the use and relevance of EC competition law in regulated industries, it runs the risk of excluding potential outcomes that might be preferred from the standpoint of the broader objectives of regulatory policy. In our view, the very factors which underlie the need for regulators – the existence of long-run problems in the competitive replication of assets, make an explicit *ex ante* regime desirable. The electronic communications services regime, which mimics and then defaults into competition law, has many attractions. But in its case too, a number of questions still remain unanswered:

- compared with the "negotiated" settlement procedure in "regulatory antitrust", does it impose sufficient restraints on NRAs in their choice of remedies?
- by focussing on the conventional set of telecommunications markets – voice services leased lines, interconnection, broadcasting transmission – does it deal adequately with the 'extended' communications value chain of the broadband era, where interfaces, devices and content acquire greater significance?

These questions leave a potential role for competition law in electronic communications services. In energy, where the political process for passing adequate sector-specific legislation at the EU level has been far more intractable, the benefits of a *ex ante* intervention under competition law seem clearer.

3. Consumers and antitrust in British deregulated energy markets

Michael Harker and Catherine Waddams Price[24]

3.1 Introduction

Antitrust institutions are designed to protect consumers from abuse of market power. In this chapter we explore the evolution of competitive constraints when consumers are able to exercise choice of suppliers for the first time. In particular, where such markets are deregulated in the sense of removal of *ex ante* regulatory constraints, does more general antitrust law provide sufficient protection or is additional sector specific regulation necessary, as provided in the UK? We focus on the recently deregulated British[25] residential energy markets as our "case study", referring to case law both during and after the removal of formal price regulation. Energy markets are politically sensitive, and many of the issues raised are similar to those discussed Cave and Crowther in this volume.

Academic literature and policy initiatives increasingly recognise the potential of consumers in competition policy (Waterson, 2003; Office of Fair Trading, 2004). Consumers have the power of "voice" in markets which are politically sensitive, and this affects both the operation of the market and the institutions within which it operates; we discuss the British context in the next section. Consumers also have the more traditional option of "exit", i.e. choosing an alternative supplier, if an incumbent supplier attempts to exploit them, and in part 3.3 we review the general role of switching, and its particular manifestation in energy markets. In section 3.4 we assess the empirical evidence of switching in British energy markets, and section 3.5 relates case law to the relevant economic issues; section 3.6 concludes.

3.2 The British energy market: history of competition and institutions

The British residential energy markets were opened fully to competition between 1996 and 1999. Virtually all British households are connected to the

[24] We are grateful to Mats Bergman and Morten Hviid for helpful comments on an earlier version of this paper; and to Laurence Mathieu and Rupert Sheldon for research assistance. We alone are responsible for any remaining errors.

[25] We focus on retail energy markets in England, Wales and Scotland. The piped gas industry in Northern Ireland is under development, and retail electricity competition has not yet been introduced there.

electricity network, and around 80% to gas. There is some extension of the gas network in rural areas, but the markets are generally mature. Both gas and electricity had previously been nationalised industries, gas as a national monopoly and electricity as fourteen regional monopolies. In both industries, the retail function was separated from distribution around the time that competition was introduced. Gas choice was introduced over two years on a regional basis, starting with half a million consumers in the South West of England in May 1996, and extending across Britain over the next two years. The electricity market was opened in 1998-99, with the fourteen regionally based incumbent monopolists.

The energy sector is politically sensitive; one reason for privatisation was supposedly to remove it from political influences which had resulted in direct political interference over the previous twenty years, including the level and structure of retail prices and the choice of fuels for generation plant. The sector is crucial to the smooth running of the economy and communications, with importance well above the 2% it contributed to GDP; major disruption would follow any supply failure. There are important low-income issues at the retail level, because although consumption increases with income, it does so at a decreasing rate, so that low-income households spend a higher proportion of their income on energy than do higher income groups (Waddams Price and Biermann, 1997). This issue has recently been emphasised by the development of a UK government policy to reduce the incidence of fuel poverty; this is defined to occur in households who spend more than a tenth of their income on energy. Moreover there has been considerable concern about households who are disconnected from supply because of non payment of bills. Since privatisation there has been a significant increase in the use of prepayment meters which firms have used as an alternative to disconnection for consumers in debt. Consequently a disproportionate number of low-income consumers use this form of payment as compared with others. Because of its strategic importance to the economy, and its sensitivity for poverty issues, the sector continues to attract considerable political interest, despite previous ambitions that privatisation would provide more independence. This effectively provides a strong, though not always representative, coherent or objectively argued, consumer voice.

The institutions of government control are based on the need for *economic* regulation of parts of the industry where monopoly is endemic (because of elements of "natural monopoly" in transmission and distribution,) or where competition has not yet developed sufficiently. A single regulator, the Office of Gas and Electricity Markets (Ofgem) with oversight of both monopoly and emerging competition, was created from the merger of the Office of Gas Regulation (Ofgas) and the Office of Electricity Regulation in 1999. The privatised retail energy markets were subject to *ex ante* price cap control, which the regulator gradually rolled back from the introduction of competition until they were completely deregulated, and subject only to the *ex post* provisions of the Competition Act 1998.

The regulator's role was reviewed in the Utilities Act 2000, one of the first initiatives of the incoming 1997 Labour Government. Ofgem acquired a new primary duty of protecting consumers "wherever appropriate by promoting effective competition". One of its new secondary duties was to take account of the needs of low income consumers, as well as those of pensionable age, the chronically sick and disabled and those in rural areas. The Act also provided for the government to provide guidance on social or

environmental issues to the regulator, but no requirement for its implementation. The same Utilities Act created an independent sectoral consumer "watchdog", Energywatch, with responsibility for providing information on the market to consumers, as well as an advocacy role.

We argue that deregulation has been a more subtle process, because of the sector's political sensitivity. There is ample evidence both from the prices charged and from interviews with companies (Sharratt and Waddams Price, 2003), that they are very conscious of potential adverse publicity from the consumer watchdog or the general media, to any changes which might be seen as disadvantaging vulnerable consumers. In some ways we could characterise the markets as only partially deregulated. Although formal *ex ante* price regulation has been removed, the companies still act under the shadow of potential regulatory, consumer watchdog and general public opprobrium, which effectively curtails their choices.

The privatised industries inherited from their nationalised predecessors a pricing system broadly based on average costs (Waddams Price and Hancock, 1998). The emphasis had been on universal access, rather than on reflecting costs for different consumer groups. Consumers with higher than average costs were generally subsidised by those with low costs. In retail supply such uniform pricing could not survive the introduction of competition, since entrants would target the lower cost, higher margin, consumers. The incumbent would either lose these consumers to entrants or have to lower its prices to retain them; in either case their higher margin would no longer be available to subsidise the higher cost consumers, whose prices would then rise if the incumbent were not to face financial difficulties[26]. Cross subsides were present in two significant aspects of the retail energy market before competition was introduced: in payment method, where companies argued that the additional costs of collecting frequent cash payment, relative to those of automated bank debits, were not adequately reflected in the price differential for these two methods of payment; and in the balance between the fixed ("standing") charge relative to the charges for fuel consumed, which did not reflect the full fixed costs of retail service. Rebalancing these elements of the tariff would harm those who used prepayment (predominantly lower income households) and less energy (again more low income than high income consumers, (Biermann and Waddams Price, 1997). The introduction of competition was therefore socially and politically controversial because of its potential effect on vulnerable households (Ernst, 1994; Cory at al., 1996). Against this background it is hardly surprising that political involvement in the sector remains high and that consumers' voice has a ready audience in political circles. In the next section we turn to how consumers have used exit by switching suppliers.

3.3 Consumer switching

Switching costs are disadvantages which a consumer perceives would be incurred by changing to another supplier, but not by staying with the

[26] We abstract here from covering such a revenue loss by increased efficiency, since our argument is focused on relative costs and prices between consumers.

current supplier. Such costs may be financial, involve the consumer's time, or be psychological. Switching costs make it more costly for consumers to change supplier, and affect the behaviour of firms. Klemperer (1995) shows that if firms are constrained to supply both "old" and "new" consumers at the same price, their effect on the competitive behaviour of firms is ambiguous in a dynamic setting. The additional costs which existing customers perceive in switching enables the firm to raise price to them without them leaving for other suppliers; but this very opportunity makes customers more profitable, once "captured", and so makes firms more ambitious to recruit customers so that in future they become 'locked in' by the inertia which switching costs engender. These dual incentives complicate the impact of switching costs on antitrust policy, and the interpretation of the actions of firms which competition authorities may observe in the market (see National Economic Research Associates, 2003, for a recent review of the literature). In particular, firms may appear to be charging predatory prices in recruiting new consumers, and excessive prices in exploiting their inertia, once recruited. Such actions are consistent with charging consumers an appropriate price (relative to costs) over the whole period of the firm's relationship with the consumer. In energy, dynamic considerations are complicated by the use of prices as a signal both for current consumption and longer term investment decisions, which themselves depend on expected future fuel prices. This dichotomy is particularly sharp at a time when energy prices are expected to rise considerably over the next few years to reflect environmental concerns. Such dynamic considerations may also be important in assessing any distributive effects of firms' pricing policies, effects which we have already noted are particularly pertinent in energy and other utility markets.

Consumers had had no choice of supplier before 1996, and incumbents inherited 100% of the market (there was no *ex ante* division of the market between incumbent and entrants as in some US states). Indeed, the conceptualisation of such choice was itself problematic when competition was first introduced; consumers found it difficult to distinguish the supply of energy from the pipes and wires which conveyed it (Waddams Price and Bennett, 1999).

In retail energy there are no financial costs of changing supplier, and the product is by definition homogeneous, though firms may offer different services in terms of frequency of meter reading and billing, subject to regulatory constraints[27]. Moreover there are no compatibility, learning or contractual switching costs for most consumers, though they can choose to enter a fixed term contract which may entail termination fees.

Barriers to switching can be divided into three main groups: search costs, to find the best deal for a consumer according to payment method and consumption level (mainly time, which will vary according to access to information); transaction costs for actually making the change (mainly time, though some of this will overlap with the previous category); and uncertainty and psychological costs associated with a new provider whose characteristics may be unknown.

Suppliers are required to offer general tariffs for three forms of payment: prepayment (where the supply of energy is activated through insertion of an electronic card recording advanced payment); quarterly

[27] Meters must be read at least every two years.

standard credit terms; and monthly automated deductions from bank accounts (direct debit). In addition, British Gas has offered discounts for prompt payment against the standard tariff. The same tariffs must be available to all consumers, and be published. Information about these tariffs is made available through the consumer watchdog, Energywatch, and by firms marketing door to door or by telephone.

3.3.1 Debt blocking

At market opening, suppliers were entitled to object to a customer transferring to another supplier on the grounds that the customer was in debt ("debt-blocking"), giving rise to concerns that some consumers, particularly those on low incomes, who were most likely to be on prepayment meters, were being "locked-in". These households were also most likely to be using prepayment meters, which companies could insist on installing for those in debt. Prepayment tariffs are generally higher than others, so these low income consumers, who were least able to pay, faced the highest tariffs and might be prevented from switching.

While there were administrative procedures in place for the assignment of debt from the old to the new supplier (with the latter's consent), these provisions were rarely used. In 1999, Ofgem expressed concern that switching had been blocked inappropriately and that the objection process was being used as a "customer retention strategy" (1999a, pp. 12-14).. Generally, it concluded, that "[e]xtensive use of debt blocking facilities is likely to dilute competition and provide opportunities for suppliers to frustrate customers wishing to switch to a competitor". In respect of debt assignment (an alternative to debt-blocking), Ofgem found that debt assignment was only being employed by a small minority of the companies surveyed. In contrast, debt-blocking was being activated in more than an eighth of cases, effectively introducing high switching costs for some consumers.

Firms and consumers may enter fixed term contracts (for example for energy at a set price) subject to a termination fee. This concept has been extended to enable consumers to commit to buying energy from firms who make energy saving investments in their homes (the British housing stock is notoriously energy inefficient, particularly amongst low income households). This suspension of consumers' rights to switch suppliers at 28 days notice is being undertaken on a trial basis, with considerable alternative protection through relation of the price paid to a reference group of other consumers supplied by the firm concerned (Ofgem, 2004a). Ofgem's views of fixed term contracts are discussed further in section 3.6.

Both regulator and consumer watchdog have placed considerable emphasis on making appropriate pricing information available to consumers to stimulate activity in the market. However this raises well known problems of the ambivalent effect of increased transparency; while it may encourage consumers to switch actively in the market, by reducing their search costs it simultaneously improves the conditions for co-ordinated effects (tacit collusion) between the firms in the market (Nilsson, 1999) and Møllgaard and Overgaard, 2002). The very presence of a regulator and consumer watchdog, acting overtly in the interests of

consumers, may exacerbate co-ordination between the firms which will be to consumer detriment. We return to this in section 3.5.

3.4 Evolution of British energy markets

The entrants to both gas and electricity markets have mainly been other energy suppliers, with some fringe entrants who have remained very small or been absorbed into larger suppliers. When competition was introduced in gas, the incumbent was handicapped by long term contracts which priced gas above the spot price, available to entrants. (It is likely that the incumbent now has an advantage, as wholesale gas prices rise in the early 2000s). Sources of electricity were sufficiently diverse that there was no consistent market advantage.

Consolidation in the industry, mainly by takeover, has led in 2004 to the domination of both the national gas and regional electricity markets by six main players, namely the gas incumbent and five descendents of regional electricity incumbents (see Table 1 which shows the division of customers between these six players.

Table 1: Percentage of customers supplied by energy retailers, December 2003

supplier	gas	electricity (national)	electricity incumbent 'in area'#	gas and electricity	dual fuel (gas and electricity *
British Gas	61	24	n.a.	40	44
Powergen	12	21	58	17	18
npower	9	15	56	13	13
EDF Energy	5	14	64	10	8
Southern and Scottish	7	14	73	11	10
Scottish Power	6	11	61	8	8
Others	1	1	n.a.	1	0

group percentage, unweighted average across regions; * summer 2003;

n.a. = not applicable

Source: Ofgem, 2004b, various tables

Incumbents in each market have therefore retained considerable market share of around 60% and are able to charge a significant price premium over entrants in its own historical market, shown in table 2. The first supplier in table 2 is the gas incumbent, the others are electricity incumbents

Table 2: Maximum and median price premia of incumbent over entrants, July 04, £s per year for consumer with medium demand

	direct debit premia		quarterly credit premia		prepayment premia	
Incumbent supplier	maximum	median	maximum	median	maximum	median
British Gas	77	63	104	91	83	68
Powergen	42	22	28	10	40	16
npower	44	21	38	16	52	24
EDF Energy	58	23	47	21	42	34
Southern and Scottish	49	21	43	24	48	21
Scottish Power	64	23	57	27	36	36

Source: Energywatch website and own calculations; note that these figures include price rises announced by British Gas in August 2004, but not by other companies who may follow suit shortly afterwards.

Several consumer studies report the evolution of these market shares and price premia as competition developed. Waddams Price and Bennett (1999) showed early evidence of targeting more profitable customers and tariff rebalancing was already evident. In a survey of consumers in early 1999, about eight months after the gas market was fully opened to competition, Giulietti et al. (2003) found that switching behaviour was best explained through a 2 stage decision model where awareness of the ability to switch and switching itself are modelled simultaneously. Search costs were found to be a barrier for many consumers, particularly those who had no switching experience in other similar markets; and switching costs were *perceived* to be high (relative to the reality) in terms of the time it would take. Those of pensionable age and using prepayment meters were less likely to be aware of the possibility of switching and those in rural areas less likely to make the change. Willingness to consider switching increased with income, but at a decreasing rate. Respondents were more likely to consider switching if the best entrant offers relative to incumbent charges (for their payment type and consumption level), was low and if they expected that the incumbent was reluctant to match these savings, i.e. they expected the benefits of switching to last for some time. Most of the respondents, particularly those who did not switch, were optimistic that the incumbent would match the entrants' lower prices; such expectations confer considerable market power on the incumbents (Giulietti et al., 2003). Table 2 suggests that five years later such optimism is not yet fully justified, particularly after sharp price increases by the gas incumbent in mid 2004. Overall Giulietti et al. found that search costs and perceived (rather than actual) switching costs were the main barriers to more active consumers.

Analysis of a survey of low income and predominantly prepayment electricity consumers about eighteen months later focused on *actual* gains that consumers had made from switching, rather than the greatest potential

gains (Waddams Price, 2003). Almost half the consumers had switched to a *higher* cost supplier than the incumbent for their current consumption and payment level, though both gains and losses were fairly small. Since this sample included a disproportionate number of prepayment consumers, for whom fewer good offers were available, results are likely to overstate the losses amongst switchers as a whole (see table 2 above and discussion below).

3.5 Implications for antitrust policy

In this section, we consider the legal and regulatory case law established during the transition from monopoly to competitive retail supply markets and its economic implications. We discuss each economic issue and relevant regulatory rulings as the retail market developed.

When energy competition was first introduced in the UK, the industries were subject to the terms of their respective Privatisation Acts and incorporated in their licenses, and to the Fair Trading Act, 1973. This was reformed under the Competition Act 1998, which represents a fundamental change in the general competition law of the UK, both in terms of its content and its administration. At the heart of the new regime are the Chapter I and Chapter II prohibitions modeled on Articles 81 and 82 of the EC Treaty respectively. The Chapter I prohibition makes *prima facie* unlawful agreements or concerted practices which have the object or effect of preventing, restricting or distorting competition. The Chapter II prohibition makes conduct which amounts to an abuse of a dominant position unlawful.

Since the 1998 Act came into force in March 2000, most of the UK sectoral regulators (including Ofgem) have enjoyed concurrent powers with the OFT to apply the prohibitions to their industry sectors, including the power to impose financial penalties in respect of an infringement of either of the prohibitions. This reform of Competition Law coincided with the move away from prescriptive *ex ante* regulation to *ex post* policing of emergent competition in these markets.

When competition had first been introduced for residential gas consumers, the incumbent was subject to a cap on the average revenue raised from all tariffs in this market, but with no separate control on individual tariffs. A number of issues about relative prices arose once competition was announced in 1994 (see below). In response, the regulator devised a new price control for 1997 to 2000, with separate caps on each of BGT's tariff categories. In electricity, absolute price caps operated on all of the incumbents' tariffs when competition was introduced.

From 2000, price controls on the direct debit tariffs of BGT and the "in-area" electricity incumbents were removed, since the regulator believed that competitive pressures obviated the need for consumer protection in these sub-markets. Absolute price caps remained for credit and prepayment consumers. From 2001 to 2002, the electricity incumbents were subject to absolute price controls on their credit and prepayment customers. In gas, BGT's prepayment tariff was subject to a relative price cap which provided for a maximum differential between its prepayment and direct debit tariffs. In April 2002 all *ex ante* price controls were removed from British retail energy markets.

3.5.1 Market definition

Whether electricity markets are defined as national or regional (in line with the 14 ex-PES regions) has important implications for the application of the Chapter II prohibition.[28] For the purposes of Article 82, the Community Courts have stated that dominance can be presumed where an undertaking has a market share above 50 per cent.[29] OFT guidance indicates that an undertaking is unlikely to be considered individually dominant if its market share is below 40 per cent (OFT, 1999a, para. 2.11). Table 1 shows that most incumbents have a market share of around 60% in their 'home' region, but less than 25% nationally.

Ofgem employed "price parallelism analysis" to decide this question in 2002: market conditions are viewed as comparable between regions if these conditions impact similarly on prices (Ofgem, 2002, para. 4.5 and OFT, 1999b, para. 3.6). In this case, they are assumed to operate in the same (national) market. Ofgem's analysis in 2002 did not support this proposition, so it continued to define the market regionally, a position recently confirmed (Ofgem, 2003a). The market shares of each electricity incumbent therefore indicate that they are dominant in their home region.

3.5.2 Price discrimination

The industries inherited from their nationalised predecessors a requirement that a "dominant firm shall not exercise any undue discrimination against any person or class of persons" and shall not set charges which are "unduly onerous" or "predatory". This condition was already incorporated in incumbent's licenses when competition was first introduced, and remained until 2001 when the sector became subject to general competition law. Under the 'old' regulation, the gas regulator ruled on undue discrimination in two separate contexts.

The first issue was relative charges made by the incumbent to consumers who paid by different payment methods. When competition was announced in 1994, the incumbent immediately rebalanced prices, lowering the relative prices for direct debit payers relative to prepayment. The regulator issued three rulings on undue discrimination between consumers using different payment methods (Ofgas 1995, 1996, 1998). The *level* of revenue allowed to BGT was not in question, since this was the subject of a separate average revenue setting exercise, merely the *relative* prices charged within the overall cap. Some rebalancing in favour of Direct Debit would be expected, to correct previous price structures which had not adequately reflected their lower costs of supply, particularly since British Gas had expected to retain its monopoly it undertook little rebalancing (Giulietti and Waddams Price, 2005).

The regulator interpreted her duties with respect to undue discrimination as follows. "If the tariffs do not cover the costs directly attributable to each category of customer, there is a clear case of

[28] The matter does not arise in respect of the gas market where the incumbent was national and enjoys a 61 per cent share of the market (see table 1).

[29] Case 27/76 United Brands v Commission [1978] 1 CMLR 429.

discrimination. In judging an appropriate allocation of the remaining costs there is more room for debate and therefore more room for discretion by [the incumbent]. However, in the context of a price controlled monopoly business, fairness in the recovery of "joint" costs is one of the main issues to be addressed" (Ofgas, 1995, pp. 8-9). The issue of fairness is particularly pertinent where rebalancing is likely to harm vulnerable households, as was true in energy (Waddams Price and Hancock, 1998, Bennett et al., 2002). In these particular references the regulator found that while there was some evidence of discrimination on these criteria, she did not consider them sufficient to be judged "undue".

The question of undue discrimination arose for the second time in a geographical context because gas competition was rolled out regionally across Britain over two years. Soon after the first phase of competition in domestic gas supply was introduced in the South West of England, BGT introduced a new tariff called ValuePlus, available only to direct debit customers in this first phase market, representing an average saving of around 5.5% against the regulated tariff for direct debit customers, and subject to a fixed term contract. ValuePlus was alleged to be unduly discriminatory on two counts. The first was between consumers within the phase one direct debit sub-market. Ofgas did not accept this argument holding that, since ValuePlus was available to all customers who were willing to pay by direct debit and sign a fixed term contract, it could not be viewed as discriminatory. Secondly, the issue of undue price discrimination arose against consumers in the monopoly part of the regulated market. Ofgas responded merely by observing that there would be no unduly onerous prices for customers in the monopoly markets because the price controls were designed precisely in order to prevent this. However, elsewhere in its decision it did admit that there might be the possibility of undue discrimination even within the price controls but "the implied profit margin would have to be significantly greater than the equivalent contribution from other tariffs for a regulatory tariff to be unduly discriminatory".

Any fully allocated cost distribution system is arbitrary (Braeutigam, 1989), as the gas regulator recognized in her judgements. There is clear potential for abuse where a firm prices both in regulated and competitive markets, with shared costs (Bradley and Price, 1991, Otero and Waddams Price 2001a), Armstrong and Vickers, 2001).

Similar issues arose in the only case concerning an alleged abuse of dominance in the energy supply sector, which referred to a lock-in and a win-back offer. Ofgem received a complaint against London Electricity (LE). Under the terms of the win-back offer, consumers switching back to LE would receive a voucher worth £25 after four months and another voucher worth a further £50 after nine months. Although the consumer was entitled to switch at any time, (s)he would have to remain with LE for 13 months in order to benefit from the full £75 (Ofgem, 2003a).

Abuse of dominance was alleged on two grounds. First, that the financial incentive to return to LE was considerably greater than any savings that most suppliers could offer to new customers. Secondly, the requirement to remain with the company for 13 months was a form of consumer lock-in. While LE was engaging in price discrimination (as between former and existing LE customers) which was capable of amounting to abuse, Ofgem held that the price discrimination did not have a material effect on competition owing to the "severely limited" take-up of

the offer. On this basis Ofgem made a non-infringement decision; it did not explore whether and how this type of price discrimination might amount to an abuse of dominance.

Otero and Waddams Price (2001b) explored potential price discrimination by electricity suppliers in 1999, just after the markets were fully opened. They compared the relative prices which firms charged for their electricity in the regions they were incumbents, and still subject to price cap regulation on all tariffs, with those where they were entrants and subject to no price constraints. Relative prices for prepayment and direct debit payment tariffs were analysed to allow for unobserved cost differences; within each region all competitors face the same distribution charges, and each firm would have similar generation costs for all its residential customers. The paper showed that firms offered substantially higher discounts for direct debit consumers relative to prepayment tariffs out of area. If entrants reflect relative costs where they are entrants and unconstrained, this suggests that they discriminated against direct debit consumers in regions where they are incumbent, though such discrimination was largely a regulatory imposition. Sharratt and Waddams Price (2003) found that a year after all price regulation was removed, such discrimination persisted, though to a somewhat smaller extent.

The relation between prices and costs as an indicator of market competitiveness was explored further by Salies and Waddams Price (2004). They examined the relationship between the prices charged by all electricity suppliers in each region, and observable costs, i.e. transmission and distribution charges. They found a striking difference between markets; in the direct debit and quarterly credit market, charges were closely related to costs, but incumbents charged a surcharge of between 4% and 13%. In contrast, prepayment charges were much less closely related to costs, but with no consistent difference between charges levied by incumbents and entrants. Given other evidence about the markets and relative switching rates across the payment methods, this suggests that incumbents retain market power (i.e. they can charge more) in direct debit and credit markets which are otherwise competitive (indicated by charges reflecting costs); while the prepayment markets are not very competitive, perhaps indicating some coordinated effect across firms, given that incumbents do not charge a premium in these markets. We discuss coordinated effects below.

EU law is also concerned that price discrimination may be an abuse of dominance. Article 82(c) provides that "applying dissimilar conditions to equivalent transactions with other trading partners, thereby placing them at a disadvantage" may amount to abusive conduct on the part of a dominant firm (and this is replicated in the Chapter II prohibition). For price discrimination to be feasible, it must be possible for the dominant firm to segment the market. Therefore, the question of whether a dominant firm can lawfully price discriminate between new and old customers (i.e. customers who have switched and those who have not) is crucial. If this is not permissible, then a price cut to regain (or stem the loss of) market share would have to be generalised across the dominant firm's entire customer base thereby rendering it a far less attractive strategy and, if prices are below cost, unsustainable except in the short-term.

According to OFT guidance there is no presumption that discriminatory pricing is harmful, indeed it may even be desirable where, for example, "it can increase output and lead to customers who might

otherwise be priced out of the market being served" (OFT, 1999c, paras.3.8-3.10). There are a number of EC cases which have involved undertakings targeting discounts or rebates in favour of those consumers who are most likely to switch to a competing supplier. In the Hilti case, the European Commission found an abuse where Hilti had identified its competitors' main customers and offered them more attractive discounts than were offered to firms who had remained loyal. The Court of First Instance upheld this approach agreeing that this was "not a legitimate mode of competition" for a dominant undertaking.[30] The *Compagnie Maritime Belge v Commission*[31] case concerned a liner conference whose members benefited from a block exemption permitting the fixing of shipping rates and schedules for particular services. If a non-member sought to compete with them, members would agree and designate "fighting ships" for which lower rates would be quoted to shippers. While not loss-making, the owners of the fighting ships would receive payments from other the members of the conference compensating them for any loss of profits. The European Court of Justice ruled that the targeted use of a pricing strategy to eliminate competition was clearly an abuse *even if the prices were not below cost.* While not a criticism of the particular case, one commentator noted that banning this form of price discrimination, where prices are not below costs, "would chill many legitimate and competitively benign forms of price competition that had no significant exclusionary effects" (Ridyard, 2003, p. 293).

3.5.3 Predatory pricing

In the ValuePlus and EU cases discussed above, the issues extended beyond price discrimination to include predation. This hinged on whether the dominant firm(s) intended to exclude rivals. In the ValuePlus case, BGT's license condition provided that supply terms would be predatory if "charges in accordance with those terms would not cover such avoidable costs as they ought reasonably to cover" and had the purpose of, or were likely to have the effect of, unfairly limiting or excluding competition. Ofgas interpreted "avoidable cost" as follows. BGT's (regulated) Direct Debit tariff was set to "yield a satisfactory margin above... attributable costs". BGT's take-or-pay contracts meant that it was paying about 5 pence per therm for gas more than the market price paid by new entrants. The ValuePlus tariff represented a reduction of 2.5 pence per therm for an average customer. The difference between BGT's contract price and the market price for gas was apparently viewed as a "sunk cost" and not, therefore, which "avoidable" (Ridyard, 1999). Ofgas considered there to be no evidence that ValuePlus either represented an attempt to eliminate competition or had such an effect, because all but one of the new entrants retained a price advantage over BGT. Selective price cuts, such as ValuePlus, would only be permitted if competition was established in the relevant sub-market, otherwise BGT would have to generalise the price cut throughout the country, i.e. a price cut on a regulated tariff.

The general question of predation in energy markets arose again under

[30] Case T-30/89 Hilti AG v Commission [1992] 4 CMLR 16, para. 100.

[31] Case C-395/95and 396/95 P, [2000] 4 CMLR 1076, paras. 112-121.

the Competition Act 1998. In March 2001, the OFT and Ofgem had published guidance on the application of the 1998 Act to the gas and electricity markets (OFT, 2001). Because of the extent of incumbent market power in supply (para. 3.4), Ofgem are committed to paying "particular attention to the possibility of pre-emptive behaviour by incumbents... designed adversely to affect the development of competition in... the relatively newly opened domestic gas and electricity markets" (para. 3.6).

The guidance concentrates on predatory pricing as the key example of pre-emptive behaviour and identifies three factors of particular significance: the intentions of the undertaking; the feasibility of the undertaking recovering the losses it incurs; and the level of the undertaking's prices relative to its costs (para. 3.28). In respect of the last point, Ofgem will apply a "relatively strict cost-based test" having particular regard as to whether the undertaking is recovering its "avoidable costs" which are defined as including "elements of costs that are often described as fixed costs that would not be included in a variable cost test" (para. 3.29).

This avoidable cost test is similar to that used by Ofgas in the ValuePlus decision, but differs from the cost floor normally used in assessing predation. The Community Courts have established two general cost based tests for this purpose. Where pricing is below average variable cost (AVC), there is a presumption that the pricing is predatory; where pricing is above AVC, but below average total cost, pricing may be predatory where there is evidence that the undertaking intended a competitor to be eliminated.[32] However, it is generally recognised that in industries such as utilities where there are high fixed costs and low incremental costs, alternative cost floors may be considered. The long-run incremental cost test may be more appropriate, including both capital and operating costs, but not common costs. This is the test applied to the UK telecommunications sector: see OFT 2000, para. 7.8. The "avoidable cost" test is a further alternative, and is defined by OFT as including fixed and variable costs, but not including common costs and sunk costs (OFT 2001, para. 4.10).[33] However, in assessing whether an incumbent is engaging in predatory pricing, evidence of pricing below avoidable costs will not be as strong as evidence that prices are below AVC (ibid., para. 4.11).

Ofgem have recently published new draft guidance for the energy sector which is less clear about which cost floor will apply. It is "likely" to presume that pricing below AVC is abusive, though it concedes "there are legitimate commercial reasons for pricing below [AVC]" which might justify otherwise abusive conduct (Ofgem, 2004c, para. 3.24).[34] It also states that the avoidable cost test may be more appropriate in some markets, but is not specific. There seems to be a grey area between AVC and avoidable costs, and whether pricing within that area is an abuse will depend upon the intentions of the dominant firm and/or the effect of the pricing decision on competition.

---

[32] Case C62/86 AKZO *Chemie BV v Commisson* [1993] 5 CMLR 215; Case C-333/94P *Tetra Pak II* [1997] 4 CMLR 662.

[33] There may be exception in respect of sunk costs if they are shown to have been incurred as part of a predatory strategy where the undertaking could have avoided them.

[34] This appears to be in response to *Aberdeen Journals Ltd v OFT* [2003] CAT 11, para. 357 where the Tribunal conceded that there may be "rare cases" where the pricing below AVC is not predatory.

The uncertainty of antitrust authorities over tests for predation mirrors that in economic literature. In their review chapter, Ordover and Saloner (1989) provide a summary of the various tests of predatory conduct and conclude that none necessarily leads to higher social welfare. In theory, as in practice, there is no single "bright-line" standard for defining predation. The economic theory of limit pricing by an incumbent monopolist distinguished between lumpy and continuous entry (Kamien and Schwartz, 1971). In energy, the single and well advertised date of each energy market's opening was an opportunity for lumpy entry, and in practice, most effective competitors entered on that initial date. Since the incumbent's prices were still capped, this constrained prices upwards, but not necessarily downwards, and there was some under-recovery, relative to the cap, in the most competitive of the gas incumbent's markets (Ofgem (1999b), pp. 13-14).

Even when prices were deregulated, incumbents did not immediately raise their prices, and vigorous marketing by entrants, based mainly on price, suggested to many observers a competition to survive as one of a handful of players (Centrica, 2001). In 2004 suppliers have raised their prices against a background of increasing upstream costs. This may signal a change in the market, as it consolidates with six established players who are competing less aggressively with each other. We return to this issue below in considering coordinated effects.

3.5.4 Switching as a competitive constraint

Section 3 above discussed the potential for consumers to constrain the incumbent's power through their willingness to change supplier. Both models and subsequent experience suggest that switching is a rather weak constraint. Giulietti et al. (2003) asked consumers what price savings would induce switching, and deduced that the profit maximising strategy for the gas incumbent was to keep price about £100 a year above that of competitors. Though this would mean losing about 45% of the market, the profits from the remaining 55% of loyal consumers would render this profitable. Using the same question for electricity, Waterson (2003) found a similar outcome. Six years after the gas market was fully opened to competition, and two years after all *ex ante* price regulation was removed, we note from tables 1 and 2 that incumbents do indeed retain just over 60% of their markets, with prices about £45 per year higher than the median competitive entrant (depending on payment method and consumption level). Moreover in the gas market the incumbent's price rises announced in August 2004, raised the premium to around £85, and commentators predicted that the incumbent would lose a further 5% of the market (Citigroup, quoted in the Independent, 26th August 04). A mark-up of £85 and a market share of 45% mirrors closely Giulietti et al.'s results, based on BGT's market dominance. Furthermore, awareness of competitive opportunities has fallen slightly, as the same model predicted (Ofgem, 2004b). Switching rates were higher in the early stages of the electricity market than in the corresponding stages for gas, but have levelled off at a very similar level, with similar incumbent mark ups (tables 1 and 2).

These discussions have been based on market dominance by a single

firm. In the next section we discuss joint dominance, or coordinated effects. Since energy is effectively a necessity, households will need to purchase it from some supplier, which makes it difficult for switching to act as a competitive constraint on any joint dominance of suppliers.

3.5.5 Coordinated effects

Consumers who are reluctant to switch suppliers may contribute to the sustainability of coordinated effects. Coordinated effects may arise in markets where firms recognize that it is in their mutual interests not to compete actively against each other. In its general guidance (Competition Commission, 2003), the UK Competition Commission (which examines mergers and alleged abuse of dominance referred by the Office of Fair Trading) identifies three conditions for coordinated effects: a significantly high market concentration for firms to be aware of each others' actions; an ability to change prices quickly so deviant behaviour can be punished; and weak competitive constraints from entrants or fringe players outside the core group. Ofgem quotes high Hirfindahl Hirschman Indices (HHIs), well above the 1800 figure which the OFT regards as highly concentrated (Ofgem, 2004b; OFT, 2004). In gas, the figure is falling very slowly but is still above 4,000, a "numbers equivalent" of 2.5. Most electricity regions have similar HHIs, though the figure rises to nearly 7000, a "numbers equivalent" of less than 2 in the north of Scotland region. These figures have risen recently because of consolidation between electricity suppliers.

Other characteristics of the sector might also facilitate coordinated effects. Gas and electricity are homogeneous products, facilitating coordination. Electricity incumbents have broadly similar market profiles, though the gas incumbent's national market presence gives it a different structure. There is repeated interaction between the firms, both in each of the regional electricity markets and in the national gas market, strengthened by the high proportion of switching to dual fuel supply. Bernheim and Winston (1990) show that such multimarket contact can facilitate collusion, especially when costs vary between firms in different markets. Such extensive multi market contact may also help overcome the difficulties of punishing deviant firms which Ofgem (2004b) identifies. We have seen that consumers perceive there to be significant switching costs. The regulator itself assists transparency in the market, another condition for coordinated effects, by publishing tariff schedules. While we have seen that such information may reduce consumers' search costs, it also provides information to firms about each others' behaviour. Indeed the very market analyses which the regulator undertakes are themselves likely to help the participating firms understand better the actions of competitors and the implications for themselves. Such examples of government facilitation of anti competitive behaviour are well recognized in the economic literature (for example Ordover and Saloner, 1989, Albæk, Møllgaard and Overgaard, 1997). Moreover prices can be changed quickly, another condition for coordinated effects. There has been little successful entry from outside the group of incumbent firms. Entry seems to be only a weak competitive constraint, perhaps because of the importance of brand knowledge identified by the regulator (Ofgem 2004b). There is therefore both

theoretical and empirical evidence that these markets are prone both to abuse of dominant power by incumbents who retain very high market shares, and to coordinated effects among similar firms which interact repeatedly in 15 closely related markets.

3.6 Conclusions

We have traced the evolution of competition in the UK residential energy markets as they have moved from monopoly, through nascent and established competition to what the regulator terms effective competition. In this period they have been deregulated, in so far as *ex ante* regulation of price caps has been replaced by *ex post* policing of behaviour under the Competition Act 1998. However the industries retain sectoral regulatory institutions, in particular a regulator and a consumer watchdog whose role remains very powerful, and incumbent market shares remain above 50%.

Ex ante price regulation and sector regulation was a substitute for consumer ability to exit monopoly markets; consumer voice is provided both by the regulator and by consumer bodies (in the case of gas, independent of the regulator; in the case of electricity, part of the regulators' office). As the ability for consumers to switch provider was introduced, so these ex ante price controls were removed. But only half the UK households have exercised this choice, and switching seems to act as a relatively weak competitive constraint. Although the general trend in prices is for incumbents to price above the new entrants and suffer a gradual decline in market share, incumbents retain market shares of around 60%, something which would certainly have drawn these markets to the attention of more general competition authorities in other markets.

Ofgem forecasts that, as market share declines the "strategy of trading market share for profit will look increasingly short-sighted" with the consequence that there will come a point where "it will be rational for the incumbent to start competing on price for every customer" (Ofgem, 2003b, para. 3.16). The regulator propagates a largely optimistic view of competitive forces in these markets on the grounds of the dynamic movement towards competition. However, in ruling on ValuePlus, Ofgas did not include an analysis of switching costs, and it seems to take a permissive approach to win-back strategies. These involve price discrimination between consumers who have switched and those that have not, so incumbents may be able to adopt a two-tier strategy to retain and even increase market share. Selective discounts would be available to attract back consumers who have switched away in combination with fixed-term contracts, which in themselves create switching costs. If price discrimination of this type is permitted, the only requirement that the incumbents need satisfy is that their win-back offers do not offend the avoidable cost test of predation. Ofgem seems reluctant to take decisions on price discrimination and predatory pricing which would provide solid case law as guidance to the limits of incumbent behaviour. Five years after the introduction of competition in both markets, we still see that the incumbents can maintain substantial mark-ups while retaining more than 50% market share, and with continuing consolidation in the industry, there are real concerns about coordinated effects between suppliers.

If the regulator's optimism about the speed and direction of change is misplaced, it could find itself in the role of protecting these industries from the fierce winds of more general competition law, rather than of protecting the consumers from the exercise of market power, a perception which earned US regulators such a bad reputation in the last century. The counterargument is that the regulator is justified in applying different criteria to a newly opened market to ensure that competition continues to develop. Like all 'infant' arguments for special arrangements, there are real dangers that the 'infant' will never be allowed to mature. This possibility provides an argument for integrating the functions of the energy regulator which deal with competitive markets with more general Competition Authorities, where the focus is on competitive markets, to ensure consistency of approach. For these aspects, sunset clauses for regulatory responsibility should be considered unless effective competition develops rapidly.

Recent price rises by the gas incumbent exacerbate concerns about its dominance, and the regulator and the consumer watchdog urge consumers to switch. Here the consumer "voice" is encouraging consumer exit. If exit is still undeveloped fully as a competitive constraint, is consumer voice acting as a surrogate? The regulator's primary duty has been changed to be protection of consumers, and there is an independent consumer watchdog, Energywatch. Moreover there are other consumer groups, such as the National Consumer Council (NCC) and the Consumers Association (CA), who take an interest in these markets; and the political sensitivity of energy gives these groups a ready audience amongst the media and politicians. We have seen evidence that incumbents do feel constrained by such informal arrangements. While welcoming the protection which such bodies provide, particularly for vulnerable consumers, there are real concerns about their accountability, particularly since they are not required to consider the costs of implementing the changes and protection which they seek. Such concern is exacerbated for non statutory bodies (e.g. the NCC and CA) which are not part of the government system, have no statutory duties for balance or objectivity, yet exercise considerable political influence.

This confusion of voices is not an adequate substitute for effective competition, or single minded regulation. Indeed the very vacuum of the regulator's case law makes their roles both more important and potentially more dangerous. Such consumer voice could cause as much damage as good because the organisations are not necessarily operating within a framework of guidance, and they lack full investigative powers to discover all the relevant facts. Their 'voice', though loud, may be unrepresentative or confused, and may stimulate changes which would raise costs for consumers as a whole. While the market itself lacks sufficient competitive constraints for single or joint dominance via consumer exit, the partial regulation exercised by consumer voice may make things worse rather than better for consumers themselves.

By mid 2004, consumer exit (switching) had been introduced as an alternative to consumer voice, but only half the households had exercised this choice. *Ex ante* regulation had been replaced by general competition law, but single and joint dominance remained high. Consumer interests themselves were expressed through a variety of institutions, some of which had little clear framework or accountability for their campaigns. This may just be a necessary phase en route to fully fledged competition, as the regulator would have us believe. But the mixture of consumer choice,

regulation and lobby groups seems somewhat ad hoc. Mixing strategies, each of which works well on its own, will not necessarily deliver the greatest benefits for consumers or the economy. For example, the regulator's presence may assist coordination in the market, increasing rather than restraining, joint dominance.

Retaining sector regulatory institutions for jurisdiction in these deregulated[35] markets may protect, rather than address, anti-competitive behaviour within the sector. The impact of institutions is not necessarily benign because the intention is good, and there are dangers that they may both protect monopoly power and facilitate coordinated effects within the market. Pessimism about the energy markets is shared by Helm (2001), particularly at a time when energy prices are expected to rise sharply; these will have higher impact on low income households because they devote a higher proportion of their expenditure to these products

The experiment to deregulate residential energy markets in the UK, was a bold one and the outcome is not yet clear. Deregulation was necessary to implement the experiment fully, and if the regulator's optimism is correct, we could see lower prices and more choice for consumers as a result. We already see some adjustment of relative prices to reflect costs more closely and increase greater efficiency; though since low income consumers were largely the beneficiaries of previous cross-subsidies this process may have adverse distributive implications. If the worrying signs of dominance and high incumbents' prices continue, there are real concerns about the effect on consumers who are reluctant to switch providers, and particularly low income and other vulnerable households.

Other countries, particularly in Europe, are committed to opening their markets over the next few years. The UK experiment shows one route to deregulation with an institutional arrangement based on sectoral regulators and consumer watchdogs, and a host of informal lobby groups. Until consumers themselves become more active in the market, it is unclear that the current arrangements can effectively constrain either single or joint dominance. The UK experience is both a good example of what can be done, and a warning of potential economic and political pitfalls.

[35] They are a useful institution for sectors which will retain price control because of natural monopoly characteristics.

References

Albæk, S., Møllgaard, P. and Overgaard, P.B. (1997), *Government-assisted oligopoly coordination? A concrete case,* Journal of Industrial Economics, 45, 4, 429-443.

Armstrong, M. and Vickers, J. (2001), *Competitive price discrimination,* RAND Journal of Economics, 32, 4, 579-605.

Bennett, M., Cooje, D., and Waddams Price, C. (2002), *Left out in the cold? The impact of new energy tariffs on the fuel poor and low income households,* Fiscal Studies, 23, 2, 167-194.

Bernheim, D. and Whinston, M.D. (1990), *Multimarket contact and collusive behaviour,* RAND Journal of Economics, 21, 1, 1-26.

Bradley, I. and Price, C. (1991), *Partial and Mixed Regulations of Newly Privatized U.K. Monopolies,* In: Weigel, W., ed., Economic Analysis of Law - A collection of Applications, Schriftenreihe der Bundeswirtschaftskammer, Vienna.

Braeutigam, R.R., (1989), *Optimal Policies for Natural Monopolies* In: Schmalensee, R. and Willig, R., eds., Handbook of Industrial Organization, Volume I, Elsevier Science Publishers B.V., Amsterdam.

Competition Act (1998), HMSO, London.

Competition Commission (2003), Merger References: Competition Commission Guidelines.

Corry, D., Hewitt, C. and Tindale, S. (1996), Energy '98: *Competing for Power,* Institute for Public Policy Research, London.

Ernst, J. (1994), *Whose Utility?: The Social Impact of Public Utility Privatization and Regulation in Britain* (Public Policy and Management), Open University Press, Buckingham.

Fair Trading Act (1973), HMSO, London.

Giulietti, M. and Waddams Price, C. (2005, forthcoming) *Incentive regulation and efficient pricing,* Annals of Public and Cooperative Economics, issue 1.

Giulietti, M., Waddams Price, C. and Waterson, M. (2003), Consumer Choice and Industrial Policy: *A Study of UK Energy Markets,* Working Paper 112, Centre for the Study of Electricity Markets, University of California.

Helm, D. (2001), *The assessment: European networks-competition, interconnection, and regulation,* Oxford Review of Economic Policy, 17, 3, 297-312.

Kamien, M.I. and Schwartz, N.L. (1971), *Limit pricing and uncertain entry,* Econometrica, 398, 441-454.

Klemperer, P. (1995), *Competition when consumers have switching costs: An overview with applications to industrial organisation, macroeconomics and international trade,* Review of Economic Studies, 62, 513-539.

Møllgaard, P. and Overgaard, P.B. (2001), *Market transparency and competition policy*, Rivista di Politica Economica, 91, 4-5, 11-58.

National Economic Research Associates (2003), *Switching Costs,* Economic Discussion Paper 5, Office of Fair Trading, London.

Nilsson, A. (1999), *Transparency and Competition,* Stockholm School of Economics, Mimeo, Stockholm.

Ofgas (1995), Referral by the Gas Consumers Council Relating Discounts for Customers Paying by Direct Debit: The Director General's Decision, Ofgas, London.

Ofgas (1996), Gas Competition: Phase 1 - Research Study Conducted by MORI for Ofgas, Ofgas, London.

Ofgas (1997), ValuePlus: British Gas Trading's Pricing to Direct Debit Customers in the South West of England: A Decision Document, Ofgas, London.

Ofgas (1998), Review of British Gas Trading Domestic Supply Tariffs: A Decision Document, Ofgas, London.

Ofgem (1999a), *Consumers in Debt and their Access to the Competitive Market,* Ofgem, London.

Ofgem (1999b), Review of BGT's Price Regulation: Initial Proposals, Ofgem, London.

Ofgem (2002), *Electricity Supply Competition:* An Ofgem Occasional Paper, 83/02, Ofgem, London.

Ofgem (2003a), The Gas and Electricity Market Authority's Decision under the Competition Act 1998 that London Electricity plc has not Infringed the Prohibition Imposed by Section 18(1) of the Act with Regard to a 'Win Back' Offer, Ofgem, London.

Ofgem (2003b), Domestic Gas and Electricity Supply Competition: Recent Developments, Ofgem, London.

Ofgem (2004a), Testing Domestic Consumer Take-up of Energy Services: Direction to Initiate Trial Suspension of 28 Day Rule: Explanatory Document, May, Ofgem, London.

Ofgem (2004b), Domestic Competitive Market Review 2004, A Review Document, April, Ofgem, London.

Ofgem (2004c), The Competition Act 1998: Application in the Energy Sector, Ofgem, London.

OFT (1999a), The Competition Act 1998: Assessment of Market Power, OFT 415, London.

OFT (1999b), The Competition Act 1998: Market Definition, OFT 403, London.

OFT (1999c), The Competition Act 1998: Assessment of Individual Agreements and Conduct, OFT 414, London.

OFT (2000), Competition Act 1998: The Application in the Telecommunications Sector, OFT 417, London.

OFT (2001), The Competition Act 1998: The Application in the Energy Sector, OFT 428, London.

OFT (2004), A Strategy and Framework for Consumer Education: A Consultation Paper, July, OFT, London.

Ordover, J.A. and Saloner, G. (1989), *Predation, Monopolization, and Antitrust,* In: Schmalensee, R. and Willig, R., eds., Handbook of Industrial Organization, Volume I, Elsevier Science Publishers B.V., Amsterdam.

Otero, J. and Waddams Price, C. (2001a), *Incumbant and entrant response to regulated competition: Signaling with accounting costs and market prices,* Journal of Business and Economics, 53, 209-223.

Otero, J. and Waddams Price, C. (2001b), *Price discrimination in a regulated market with entry: the residential UK electricity market,* Bulletin of Economic Research, 53, 3, 161-175.

Ridyard, D. (1999), *Regulation of Price Discrimination and Predation by Dominant Firms: Lessons from the Ofgas ValuePlus Decision* In: McCrudden, C., ed., Regulation and Deregulation: Policy and Practice in the Utilities and Financial Services Industries, Clarendon, Oxford.

Ridyard, D. (2002), *Exclusionary pricing and price discrimination abuses under Article 82* – an economic analysis, European Competition Law Review, 23, 6, 286-303.

Salies, E. and Waddams Price, C. (2004), *Charges, costs and market power: The deregulated UK electricity retail market,* The Energy Journal, 25, 3, 19-35.

Sharratt, D. and Waddams Price, C. (2003), *Deregulating Prices and Social Obligations,* Company Policy and Outcomes, Mimeo, University of East Anglia.

Utilities Act (2000), HMSO, London.

Waddams Price, C (2004), *Spoilt for Choice? The Costs and Benefits of Opening UK Residential Markets,* Working Paper 123, University of California Energy Institute.

Waddams Price, C. and Bennett, M. (1999), *New gas in old pipes,* Energy Policy, 8, 1-15.

Waddams Price, C. and Biermann, A. (1998), *Fuel Poverty in Britain: Expenditure on Fuels* 1993-94 to 1995-96, Centre for Management under Regulation, Warwick Business School, University of Warwick, Gas Consumers Council, London.

Waddams Price, C. and Hancock, R. (1998), *Distributional effects of liberalising UK residential utility markets,* Fiscal Studies 19, 3, 295-319.

Waterson, M. (2003), *The role of consumers in competition and competition policy,* International Journal of Industrial Organization, 21, 2, 129-150.

4. From state monopoly to the "investment ladder": competition policy and the NRF

Alison Oldale and A. Jorge Padilla[36]

And he dreamed, and behold a ladder set up on the earth,
and the top of it reached to heaven:
and behold the angels of God ascending and descending on it.

Jacob's Ladder, Genesis 28, 12

The New common Regulatory Framework for electronic communications networks and services (the NRF) was launched by the European Commission in a series of Directives adopted in March 2002 and enacted in July 2003.[37] These Directives imported concepts of competition law into a regulatory context and the apparent consensus is that by doing this the new regulatory regime was given a solid intellectual foundation.[38] The NRF has been presented as a flexible approach to regulation, which will make possible the transition from regulated markets to sustainable competition.[39] It has been described as the last chapter in the history of telecoms regulation

[36] Alison Oldale and A. Jorge Padilla are economists with LECG. We have benefited from the comments and suggestions from the participants at the 2002 Lake Como Summit, and especially from Bernard Amory, Pierre Buigues, Tim Cowen and Antoine Winckler. We have also benefited from numerous conversations with Matthew Bennett, Matthew Cherry, David Evans, Karen Frazer, Inmaculada Gutierrez, Ciara Kalmus, Robert Hahn, Paul Richards, and Andy Tarrant, and the invaluable research support of Daniel Gomez and Joachim Keller. We have also benefited from the valuable comments from the editors of this volume, Mats Bergman and Arvid Nilsson. Of course, we alone are responsible for the views expressed in this paper. Please e-mail comments to: aoldale@lecg.com and jpadilla@lecg.com.

[37] The main package is composed of four directives adopted in March 2002: *Directive 2002/21/EC of the European Parliament and of the Council of 7 March 2002 on a common regulatory framework for electronic communications networks and services*, OJ 24.4.2002, L 108/33; *Directive 2002/20/EC of the European Parliament and of the Council of 7 March 2002 on the authorisation of electronic communications networks and services*, OJ 24.4.2002, L 108/21; *Directive 2002/19/EC of the European Parliament and of the Council of 7 March 2002 on access to, and interconnection of, electronic communications networks and services*, OJ 24.4.2002, L 108/7; *Directive 2002/22/EC of the European Parliament and of the Council of 7 March 2002 on universal service and users' rights relating to electronic communications networks and services*, OJ 24.4.2002, L 108/51.

[38] See Mario Monti European Commissioner for Competition Policy, Speech/03/604, *Competition and Regulation in the Telecom Industry: The way forward*, ECTA Conference Brussels, December 2003; and ERG Common Position on the approach to Appropriate remedies in the new regulatory framework, ERG (03) 30rev1, 23 April 2004, p. 23.

[39] See, e.g., Alexandre de Streel, *The Integration of Competition Law Principles in the New European Regulatory Framework for Electronic Communications*, 26(3), World Competition, 489, 494 (2003). See also Martin Cave & P. Larouche, *European Communications at Crossroads*, report of a CEPS working party, Brussels, October 2001, and Pierre Buigues, *The Competition Policy Approach*, in The Economics of Antitrust and Regulation in Telecommunications (P. Buigues & Patrick Rey eds., Edward Elgar, 2004).

in Europe – allegedly, in the future, sector-specific regulation will be replaced by competition law.

We take a more sceptical view of the role that competition law plays and has played in framing the NRF. We are also more sceptical about the role that the NRF will play in the development of sustainable competition in the telecoms markets of Europe. In our view, the most important aspect of the NRF is that it embraces a particular (but not very new) perspective on telecoms regulation – one that prioritises the short-run benefits of fragmentation at the expense of the option of long-run competition and efficiency.[40] It does so by promoting access-based entry through obligations on large telecoms companies to supply a long list of inputs to their rivals - i.e., by creating, sometimes rather artificially, a large number of so-called "wholesale markets".[41] It is true that competition law also imposes obligations to supply, but the NRF may impose these obligations even in situations where there would be no grounds for intervention under competition law.[42] While the NRF relies on competition law for market definition and the assessment of dominance, it ignores the limiting principles embedded in the competition law rules that are used on both sides of the Atlantic when specifying the remedies to be used.

The promotion of access-based entry that is embodied in the NRF is not the only possible approach to regulation. The Commission could have chosen a regulatory framework geared towards facilities-based entry where intervention would limited to the removal of legal barriers to entry, or remained more neutral and limit regulatory intervention to preventing the exploitation of consumers. If anything, this last approach would be more consistent with the principles of competition law, as now generally understood.[43]

Some endorse the NRF, and its emphasis on access-based entry, by claiming that there is *no* trade-off between promoting access-based entry in the short run and maintaining the option of sustainable competition in the long run – the two are reconciled by the so-called "ladder of investment" theory.[44] This is seen as supporting the view that the NRF – and the active promotion of fragmented markets that it embodies – will bring about a new world of generalised and sustainable competition in European telecoms where regulation will be phased out without trauma.

---

[40] See *Commission guidelines on market analysis and the assessment of significant market power under the Community regulatory framework for electronic communications networks and services,* OJ 2002/C 165/03, p. 19.

[41] The list of obligations include access to specified network elements, including unbundled access to the local loop; open access to technical interfaces, protocols or other key technologies indispensable for the interoperability of services or virtual network services; co-location or other form of facilities sharing; and access to operational support systems *(Directive 2002/19/EC of the European Parliament and of the Council of 7 March 2002 on access to, and interconnection of, electronic communications networks and services,* OJ 24.4.2002, L 108/7, article 5 and 12).

[42] See, e.g., Martin Cave, *Economic Aspects of the New Regulatory Regime for Electronic Communication Services,* in The economics of Antitrust and Regulation in Telecommunications 35-40, (P. Buigues & Patrick Rey eds., Edward Elgar, 2004).

[43] See, e.g., Massimo Motta, Competition Policy: Theory and Practice, (Cambridge University Press, 2004).

[44] Martin Cave & Ingo Vogelsang, *How access pricing and entry interact*, 27, Telecomm. Pol'y, 717, 717-727 (2003). The idea of a ladder of investment was also presented in an earlier paper, Martin Cave & Luigi Prosperetti, *European Telecommunications Infrastructures,* 17(3), OXFORD REV. OF ECON. POL'Y, 416, 421 (2001).

The "ladder of investment" theory rests on two presumptions.[45] The *first* is that regulators have enough information and competence to micromanage the evolution of competition in electronic communications markets. The *second* is that the fragmented form of competition that access based entry will promote can be sustained in the long run without the support of perpetual regulation. These presumptions are at the very least debatable and cannot be settled simply by an appeal to competition law. Indeed, the main role of competition law in the framing of the NRF may well have been to divert attention away from these two propositions, thus precluding the necessary debate on their practical validity. A major goal of this essay is to consider whether these two presumptions are justified as a matter of fact. That is, to investigate whether the "ladder of investment" theory – i.e., the view that there is no need to choose between service-based and facilities-based competition – is something more than a pleasant dream, as *Jacob's ladder* was.

This paper is, therefore, an attempt to reopen the debate about the appropriate regulation of telecoms in Europe. To do so, we first review the history of telecoms regulation up to the NRF in Section 4.1. This is a history of changing ideas about the role of competition and the state, of continued improvements in regulatory theory and practice, of conflicting views on the right way to regulate and, more precisely, on the right way to transition European telecoms from regulation to competition. In Section 4.2, we summarise the approach embodied in the NRF and its relationship with the principles and tools of competition law. We show that the obligations to supply imposed by the NRF may go well beyond those that competition law requires and, therefore, may thwart the incentives to invest in new technologies and infrastructure – a negative effect that the competition courts on both sides of the Atlantic have taken seriously into account when considering mandating access. Section 4.3 describes how the "ladder of investment" theory is being used to claim that regulators do not face a trade-off between the short and the long run. Section 4.4 considers the first of the two presumptions underlying this claim: that regulators can micromanage the sector. Section 4.5 considers the second: that fragmented competition is sustainable. Section 4.6 concludes.

4.1 An overview of the history of telecoms regulation in Europe

The history of regulation in telecoms, especially in Europe, is one of conflicting and changing ideas about what the goals of regulation should be and how those goals should be achieved. For expositional reasons we have divided the history of telecoms regulation in three periods, using three (European) milestones to separate them: (a) from the Second World War to the 1987 Green Paper,[46] (b) from the 1987 Green paper to the NRF, and (c) the NRF.

The first period was characterised by state monopoly in Europe and

[45] The two conditions are described in Martin Cave, *Remedies for Broadband Services,* Paper prepared for DG InfoSoc, 2003.

[46] European Commission (1987), *Green Paper on the Development of the Common Market for Telecommunications Services and Equipment.*

classical forms of price and non-price regulation on both sides of the Atlantic, and it resulted in high prices and inefficient investment. The response to this state of affairs was complex and multi-dimensional: some argued for relying less on regulation and more on competition, others advocated moving from public to private ownership, and still others considered that the solution was not less regulation, but better regulation. Some European countries, most prominently the UK, liberalised and privatised their telecoms markets while at the same time introducing incentive-compatible regulations. Others, like Spain, privatised but failed to liberalise. And a few others, like France and Greece, did nothing of the sort. The second period starts with the Commission's 1987 Green paper, where the Commission, with the opposition of some Member states, defended the gradual introduction of competition into the European telecoms markets. Between 1987 and the adoption of the New Regulatory Framework in 2002, the Commission issued a series of 26 Directives – in what is known as the "1998 regulatory package"[47] – aiming to liberalise the telecoms markets of the EU. This complex regulatory building was impossible to navigate and created a thousand conflicts without resolving any of the inefficiencies that it was supposed to eliminate.

The NRF represents an attempt to simplify the regulatory thicket that resulted from the 1987 Green Paper and subsequent Directives, harmonising regulation across Member states. It is also an attempt to promote competition in the European telecoms markets, so that one day regulation is no longer needed. With the NRF, the European Commission seeks to establish a regulatory framework that ensures that the European telecoms markets feature low prices, high quality and optimal investment during the transition from regulation to competition. It also seeks to reduce the scope of regulation as effective competition thrives. The NRF appears to be regarded as the "end of regulation history".

From the Second World War to the 1987 green paper

In the decades from the Second World War to the 1980s virtually all telecoms companies in Europe were state-owned monopolies.[48] This state of affairs was a response to the prevailing views about the social and economic nature of telecoms, but it created problems that would eventually lead to calls for the system to be dismantled. State-monopoly provision was partly the result of a belief that competition in telecoms was both impossible and wasteful – telecoms, it was felt, was a natural monopoly.[49] In many countries, state ownership of telecoms also reflected a view that telecoms was a public service with important social benefits, and that its provision should depend on government policy.[50] For example, in France, at the end of the 1960s, policy shifted from regarding the telecoms service as a luxury to a policy of rapid expansion and universal provision.[51] The combination of a lack of competitive pressure with the pursuit of social

---

[47] On the 1998 package, see e.g., Cave & Prosperetti, *supra* note 44.

[48] Finland is an exception to this.

[49] Damien Geradin & Michel Kerf, *Controlling Market Power in Telecommunications: Antitrust v. Sector-specific Regulation 5* (Oxford University Press, 2003).

[50] OECD, Competition in Telecommunications, OCDE/GD(96)114, Paris, 1996.

[51] OECD (1991), *Universal Service and Rate Restructuring in Telecommunications*, page 28.

objectives meant that prices were often set with little relationship to the underlying costs of provision.

> *"Operating as monopolies, [public telecoms operators] have been able to fix prices that average the costs of supply geographically and of certain types of services. Cross-subsidies have been accepted as good practice to achieve social and economic objectives..."*[52]

Meanwhile policy makers in the United States acted on the same beliefs – that telecoms were both a natural monopoly and a public service – but in a very different way. The monopoly, AT&T, was in private rather than public hands, and state managers regulated its conduct.[53] Regulation focussed on three areas. The public service aspect of telecoms was dealt with through imposing public service obligations, such as a universal service obligation, on AT&T. The monopoly element, and the perceived risk that an unregulated company would charge high and inefficient prices, was dealt with by imposing a regulated rate of return. Finally the need to make sure that AT&T invested in upgrading the network was also addressed by guaranteeing the company a specified rate of return on any investment. This regulatory regime became known as "rate of return regulation".[54]

- **The pitfalls of state monopoly and classical regulation**

Neither approach was satisfactory. On both sides of the Atlantic commentators began to point to bloated costs, inefficient investment, and high prices.[55] Markets did not allocate resources efficiently because social imperatives often dictated complex and distortionary cross-subsidies.[56] Geographically averaged prices, for example, created a cross-subsidy from dense urban regions that were cheap to supply to sparse rural ones, where telecoms infrastructure was expensive to install. To encourage a high take up the price of access was often much lower than the cost of provision, and the shortfall was made up through higher call charges. This created a cross-subsidy from high volume users to low volume ones. These cross-subsidies interfered with the role that prices play in making sure that consumers only buy something when the benefit they receive is greater than the cost of supply.

The nationalised monopolies had little incentive to pursue productive efficiency either. Since prices and revenues depended on social objectives, and not on costs, managers were not rewarded for keeping costs low. In fact the opposite was often the case, as governments used state owned industries to manage employment, as well as to provide services. All across

---

[52] Ibid., page 15.

[53] See AT&T: A brief history available at http://www.att.com/history/history1.html and Nicholas Economides, *Telecommunications Regulation: An Introduction,* NYU Centre for Law and Business Research Paper No. 03-25, June 2004.

[54] A review of theories of regulation is presented in John Vickers & George Yarrow, *Privatization: an economic analysis,* (MIT Press, 1988).

[55] The draw-backs of ROR regulation are addressed in a set of papers presented in the Symposium on Price-Cap Regulation and published in The Rand Journal of Economics in 1989. For an overview see Jan Paul Acton & Ingo Vogelsang, *Symposium on Price-Cap Regulation: Introduction,* 20(3) The Rand J. of Econ., 369, 369-372 (1989).

[56] Cross-subsidies are not inefficient per se. Access subsidies may be efficient in the presence of network externalities.

Europe, the number of employees in state telecoms companies fell significantly after privatisation. For example, Deutsche Telekom was able to reduce its workforce by a quarter in the years after privatisation.[57]

Finally, the record of nationalised telecom companies at delivering high quality services was mixed. Decisions of quality and investment in new technologies were driven by government policy, rather than competitive innovation or consumer demand. In France, a government decision in the 1960s led to France investing in an advanced data network, Minitel.[58] Conversely, in the UK, one of the main motivations behind privatisation was to raise public funds for investment in digital switching equipment, which would otherwise not be possible without breaching government borrowing constraints.[59] The form of regulation adopted in the US also created problems, notably the so-called "Averch-Johnson" or "gold-plating" effects. AT&T was guaranteed a total return on investment. This meant that the more it invested, the higher would be its profits, which gave it an incentive to overstate its investment needs.[60]

In short, telecoms markets were allocatively, productively and dynamically inefficient. state ownership had done more evil than good and the textbook approach to the regulation of a natural monopoly had not fared much better.

• Calls for change – the debate leading to the 1987 green paper

By the early 1980s, the calls for change began. The first challenge was to the prevailing wisdom that telecoms constitute a natural monopoly, the second to public ownership, and the third to the economics of regulation. The first challenge led to the promotion of competition in some of the markets that previously had been regarded as natural monopolies. The second challenge was behind the privatization wave of the mid to late 80s. And the third challenge gave rise to the birth of the so-called "incentive-based" regulation. Let us consider each of them in order.

Limiting the scope of natural monopoly. Since the Second World War, technological advances had reduced progressively the cost of networks and, as a result, policymakers and commentators began to envisage competition between rival networks - at least to some degree.[61] The critical milestone in this process was the break-up of AT&T in 1982.[62] In January 1982, following a lengthy legal process with the Federal Communications Commission (FCC), AT&T agreed to divest itself of the wholly owned Bell operating companies that provided local exchange service. This would, the government believed, separate those parts of AT&T (the local exchanges)

---

[57] Its employees fell from 230,000 on privatisation in 1995 to 170,000 by 2000. As cited at: www.eiro.eurofound.eu.int/1999/12/study/tn9912201s.html. Downloaded, 16/08/04.

[58] OECD (1991), *Universal Service and Rate Restructuring in Telecommunications,* page 28.

[59] At www.johnkay.com/political/249. Downloaded 16/08/04.

[60] The Averch-Johnson effect was presented in Averch, H. & Johnson L, *Behaviour of the Firm under Regulatory Constraint,* 52 Am. Econ. Rev., 1053, 1053-1069 (1962). It is reviewed in Jean-Jacques Laffont, *The New Economics of Regulation ten Years After,* 62(3) Econometrica, 507, 507-537 (1994).

[61] Littlechild S (1983), Regulation of British Telecommunications' Profitability, London, Department of Trade and Industry, and Damien Geradin & Michel Kerf, *Controlling Market Power in Telecommunications: Antitrust vs Sector Specific Regulation 7* (Oxford University Press, 2003).

[62] Taken from AT&T: A brief history available on http://www.att.com/history/history1.html.

where the natural monopoly argument was still seen as valid from those parts (long distance, manufacturing, research and development) where competition was considered a possibility. Divestiture took place in January 1984, resulting in a new AT&T and seven regional Bell operating companies, commonly known as the "Baby Bells".

This development did not go unnoticed in Europe. In 1983 the Littlechild Report was published in the UK.[63] This report concluded that technological developments meant that national and international calls should no longer be considered as natural monopolies. If competition was possible then it should be allowed, it was claimed. Competition would avoid the problems that were becoming apparent under state provision or regulation and would protect consumers against monopoly power. The Littlechild Report recommended that competition should be introduced into any area that did not constitute a natural monopoly, noting that:

> *"Competition is indisputably the most effective – perhaps the only effective means – of protecting consumers against monopoly power. Regulation is essentially the means of preventing the worst excesses of monopoly; it is not a substitute for competition. It is a means of 'holding the fort' until competition arrives."*[64]

From public to private ownership. At the same time there were calls to free European telecoms companies from state control. The UK was the first of the major countries to privatise its telecoms network in 1984. Commentators were very optimistic about the potential effects of this change in ownership.[65] One hope was that privatisation would lead to lower costs. Relative to managers in state owned companies, it was suggested, those in private companies have clear incentives to maximise profits and minimise costs, and are less likely to come under pressure to pursue other objectives.[66]

By the 1980s, technological progress in telecoms was leading to the development of new technologies such as digital networks, ISDN, fax and videoconferencing. It was hoped that privatisation would lead to faster investment in these technologies. One reason was profit maximisation – if consumers would pay a sufficient premium for new services then a company in private hands had an incentive to offer them. Even more so, if the privatised company was subject to competitive threats. Another reason was the improved access to capital markets that companies in private hands were supposed to have. In the early 80s, many European governments were reluctant to invest in state owned companies, since doing so would have an impact on the level of public borrowing. This was a particular constraint in the UK where BT had tried, but failed, to raise private financing while still in state hands. The hope was that private shareholders would be more willing to invest money to exploit the new opportunities afforded by technological progress.

---

[63] Littlechild S (1983), *Regulation of British Telecommunications' Profitability*, London, Department of Trade and Industry.

[64] Ibid, para. 4.11.

[65] John Vickers & George Yarrow, *Privatization: An Economic Analysis* (MIT Press, 1988).

[66] See J. D'Souza & W. Megginson, *The Financial and Operating Performance of Privatized firms during the 1990's*, 54(4) THE J. OF FIN., 1397-1438 (1999).

From classical regulation to "incentive-based" regulation. Despite the optimism in some quarters about the potential benefits of liberalisation and privatisation, few people claimed these would lead to a fully competitive market straight away. As Vickers and Yarrow pointed out, the simple change from state to private ownership did not in itself alter the fact that telecoms was being provided by a company with substantial market power.[67] The response was renewed effort to apply scientific principles to the question of how regulation could be improved. Academic advances in microeconomics using incomplete information and contract theory would eventually lead to the *"new economics of regulation"*.[68] Whilst traditional regulation assumed that the regulator was informed about firm's costs and incentives, academics began to consider the more realistic possibility that regulators were not omniscient.[69] Instead of assuming that regulators could monitor effort and costs, for example, researchers considered ways to make sure that the regulatory scheme rewarded desirable behaviour, and so provided the right incentives.

In the UK a practical contribution to the debate was the introduction of price caps, which changed year on year by a specified amount given as the RPI (Retail Price Index) – X. These price caps required BT to reduce its prices below inflation by an "X" percentage each year. This type of price regulation was supposed to provide the right incentives from a productive and dynamic efficiency viewpoint. BT could keep any profits deriving from cost savings beyond the level necessary to achieve the RPI-X price cuts. In practice, the incentive benefits of the price cap are limited, in particular if the cap is revised all too often. The first telecommunications price cap covered the period from 1984-89.[70] BT would know that it could keep the benefits of cost reductions until 1989, but would still expect that the regulator would take them in the form of a lower cap in the period after that. Its incentives to invest in cost improvements with long-run maturities were, therefore, not affected significantly.

From the 1987 green paper to the NRF

The situation by 1987 was quite confusing. The UK had liberalised its telecoms market, privatised the incumbent and introduced new, incentive-compatible regulations. Yet in other countries, like France and Greece, the markets remained closed to competition and there was no sign of reform. The 1987 Green Paper established the Commission's position that there should be a gradual introduction of competition into the European telecoms markets. Focusing on the need to develop the telecoms market beyond basic voice services, the Commission expressed concerns that Europe's nationalised monopolies would be left behind by competition from the United States and Japan, and advocated a liberalised telecoms market as the solution. In its 1987 Green Paper, the Commission noted:

---

[67] See supra note 65.

[68] Also known as "incentive-based regulation". See David Sappington & Dennis Weis, *Designing Incentive Regulation for the Telecommunications Industry* (MIT Press and AEI Press, 1996).

[69] Some of these variables can be observable, but monitoring by third parties is very costly making them impossible to include in the terms of a contract between the regulator and the firm.

[70] The first instance of price cap regulation was for contraceptive sheaths in 1982, which at the time were produced by an unregulated monopoly. The Monopoly and Mergers Commission set the initial level of RPI-X at 1.5% in an attempt to eliminate the problem of excessive pricing.

> *"an open, competitive market for new service providers and terminal manufacturers can make a substantial contribution to the right spread of the new services, under the conditions of rapid development of technology and market opportunities."*[71]

Some Member states, however, remained unconvinced, and those that were in favour entertained different views about how the process towards increased competition should be articulated in practice. Between 1987 and the launching of the New Regulatory Framework in 2002 the Commission issued a series of 26 Directives embodying an evolving view about the role of regulation in a liberalised market.

- **To liberalise or not to liberalize**

The package of directives that followed the 1987 Green Paper – the 1998 package – stipulated that the provision of public voice telephony services should be liberalised throughout the Union by 1 January 1998. However, some of the details of implementation were left to the individual NRAs. To monitor progress the Commission published a series of implementation reports, which showed that different countries had very different interpretations of what liberalisation really means. For example, whilst the Scandinavian members were relatively fast (by 1994) to allow entry into their national markets, Greece was happy to set up a telecoms regulator (1992) but took until 2001 to license competitors other than the state monopoly. In setting up a regulatory institution, Belgium installed the Post minister as head of the NRA as well as representing the state's interest in the incumbent, whilst in Luxemburg the chairman of the incumbent was also nominated as the head of regulation by the government.

Some Member states were reluctant to liberalise because they did not want to privatise, and thought liberalisation would undermine the pursuit of social goals through public provision by a strong publicly owned company. Others resisted because, although they did want to privatise, thought liberalisation and the introduction of competition would reduce the revenues they could earn from the sale – a company facing less competition could earn higher profits, and thus fetch a higher price when sold to private buyers.

- **Allowing entry v. promoting entry**

The commitment to open up markets to competition was not the only area of disagreement. During the early 80s, regulation had been seen by some practitioners as a necessary evil to be imposed on recently privatised telecoms markets.[72] However, as the decade wore on most policy makers began to emphasise an altogether different role for regulation – as a means of *actively* changing the prevailing market structure in a desired direction. The US authorities had already used regulation in that way when they broke up AT&T. The Baby Bells (who were left in control of local networks) were barred from manufacturing equipment and from the provision of long-distance service in order to foster entry into these markets from non-

---

[71] European Commission (1987), *Green Paper on the Development of the Common Market for Telecommunications Services and Equipment*, p. 52.

[72] See supra note 63.

integrated suppliers. AT&T, on the other hand, was prevented from cutting prices on long distance calls to make entry by rivals more profitable. This tampering continued with the Telecommunications Act of 1996,[73] which stipulated that, subject to a long list of conditions, the Baby Bells could only enter long distance after they had opened their local businesses.[74] In Europe, there was agreement that regulation had a role to play in shaping the market, but there was disagreement about which direction to choose.

- **Promoting facilities-based entry v. service-based entry**

Throughout much of this period, the UK regulator – Oftel (now Ofcom) – used regulation to promote competition between competing providers, each with their own network, being:

> *"convinced that the key to achieving a vibrant market for services provided over telecommunication networks is the promotion of fair, efficient and sustainable network competition."*[75]

When BT was privatised, only one other telecommunication license was granted – to Mercury. The intention was to encourage investment by Mercury by limiting its competition, thereby promoting a single, strong infrastructure competitor to BT. Even when the perceived failure of Mercury lead the government to open the market to all players in 1991, it maintained a commitment to using regulation to foster facilities-based competitors. Whilst the cable companies were allowed to provide telecommunications services, it was not till 1998 that BT was allowed to offer broadcast services to the public. This "asymmetric" treatment was thought to facilitate the entry of the cable companies – the infrastructure providers that were expected to discipline BT in the marketplace.

Those that favoured infrastructure competition noted that it had the advantage of introducing competition into all levels of the network, potentially allowing the regulator to finally bow out of regulation. But this was not the only view; others advocated using regulation to promote entry by companies that used other people's networks to access their customers – i.e., to facilitate service-based entry. The key advantage of this type of entry was thought to be that it offered the prospect of fast reductions in the market shares of incumbents in selected retail markets. Competitors could enter without the need for so much time-consuming and costly investment. If there were sufficient retail entry then regulation to protect consumers would no longer be needed, it was argued. The key disadvantage was that possibly more and more demanding regulation was needed, at least for some period of time, to make sure the new entrants could get access to the incumbent's network.

The conundrum of telecommunications regulation, at least as it was seen at the time, was the intractability of pursuing both these goals

---

[73] *Telecommunications Act of 1996, Pub. L. No. 104-104, 110 Stat. 56, 108 (1996) (codified at 47 U.S.C. § 336(d)).*

[74] See Nicholas Economides, *The Telecommunications Act of 1996 and Its Impact*, Japan and the World Economy, (forthcoming). See also Robert W. Crandall, *The Remedy for the "Bottleneck Monopoly" in Telecom: Isolate It, Share It, or Ignore It?*, University of Chicago Law Review, (forthcoming).

[75] Oftel: *Promoting Competition in Services over Telecommunication Networks*, June 1996.

simultaneously. Allowing an incumbent to exploit its dominant position by allowing it to set high access prices would create supra-normal profits and encourage entry in the longer term, but in the short term this would entail high consumer prices and resulting welfare losses. However, regulating prices or giving low (possibly cost-oriented) priced access to service competitors would erode the incentives to invest in access infrastructure. The Commission increasingly favoured using regulation to promote a service-based model of competition, which led to disputes with those that disagreed. As late as 1996, Oftel's commitment to facilities-based competition led it to try to resist Carrier Pre-Selection, as advocated by the Commission, on the basis that:

> *"[CPS's] introduction could discourage operators from developing alternative access networks if they risked the benefits of their investments to competing operators. Oftel concluded, on balance, that there is no case for directing BT to provide equal access."*[76]

The Commission eventually imposed CPS. More fundamentally, Oftel's previous distinction between network providers (who had cost-based access to BT's network), and resellers (who had "retail minus"-based access) was blurred by Commission Directives.

The trend towards using regulation to provide service providers with cheaper and more flexible access to their customers over the incumbents' networks continued with the Commission requiring NRAs to mandate partial leased lines and local loop unbundling. Oftel and others had at one time, resisted both of these. The stage was set for the New Regulatory Framework.

A regulatory framework for the 21st century

In 2002, the New Regulatory Framework – comprising just five main Directives – replaced the 26 Directives that had been published since the 1987 Green Paper.[77] The architects of the NRF stressed its simplicity,[78] its greater flexibility, and its solid foundation on competition law principles.[79] The NRF has two main goals: achieving *regulatory harmonization* and ensuring that competition develops in the relevant markets so that regulations give way to the pure application of competition law.

---

[76] Oftel, *Policy on Indirect Access, Equal Access and Direct Connection to the Access Network – Statement*, July 1996.

[77] Interestingly for a package that is meant to clarify regulations, there is some debate on what the actual number of directives is! The French regulator ART has commented on 6 directives (including the competition directive) whilst Oftel considers there are the 5 directives listed at supra note 37. Other documents even point to 7 directives, with the Radio Spectrum Policy directive included as the last one.

[78] In practice, this is unclear. While it is true that, in principle, the NRF allows for the rolling back of regulation in narrow markets where it is no longer needed, in practice the more likely effect will be to increase the number of regulated services. The use of competition law concepts of market definition mean that NRAs must analyse, and potentially impose remedies, in 18 different markets. The previous package covered just our major markets: fixed telephony, mobile telephony, and interconnection and leased lines.

[79] See, e.g., European Competition Law Annual 1998: *Regulating Communications Markets* (C. D. Ehlerman & L. Gosling eds., Hart Publishing, 2000).

- **Regulatory harmonization**

In practice, the process set up by the NRF seeks to restrict the freedom of the NRAs to pursue regulatory agendas that differ from the Commission's.[80] The NRF is an attempt to restrict an NRA's choice of *where* and *when* to intervene.[81] The Commission first adopts a *Recommendation* that identifies the markets it believes require *ex ante* regulation.[82] The main role of the NRAs at this stage is to assess whether there are any companies with significant market power (SMP) in these markets,[83] and it must do so according the competition law standards.[84] If there is an SMP operator, then the NRA must intervene on the market. If there is no SMP operator, then it should not intervene. The NRAs can depart from the list of markets identified by the Commission within certain bounds. Moreover the Commission also has the ability to veto an SMP designation in NRA decisions that affect trade between Member states. The NRF also seeks to restrict an NRA's choice of *how* to intervene. Having identified an SMP operator, the NRA must impose a remedy from a menu of options laid out in the Directives. For wholesale markets the options are given in the *Access Directive*.[85] For retail markets they are given in the *Universal Service Directive*.[86]

Hitherto, it remains unclear whether the Commission will succeed in its attempt to restrict the ability of the NRAs to choose where, when and how to intervene. A few NRAs appear to have views of their own as regards the goals of regulation and the means to achieve them. How will the Commission respond to those challenges remains to be seen. So far, the Commission appears to take its new role very seriously. Its actions over local loop unbundling (LLU) have shown that the Commission is eager to enforce its views about how to regulate electronic communications. The Commission launched the LLU sector inquiry in 2000 to look into the slow

---

[80] See Pierre Buiges, *The Competition Policy Approach,* in *The economics of Antitrust and Regulation in Telecommunications 30* (P. Buigues & Patrick Rey eds., Edward Elgar, 2004).

[81] This description draws on that in Alexandre de Streel *The Integration of Competition Law Principles in the New European Regulatory Framework for Electronic Communications,* 26(3) World Competition, 489, 489-514 (2003).

[82] Commission Recommendation of 11 February 2003 on relevant product and service markets within the electronic communication sector susceptible of ex ante regulation in accordance with Directive 2002/21/EC of the European Parliament and the European Council on a common regulatory framework for electronic communications networks and services, C (2003) 497. The Recommendation hereinafter.

[83] The NRAs must define the geographic extent of the markets on their national territories, and also in some cases assess whether certain classes of customer form distinct markets. But this gives only a very limited amount of discretion.

[84] Under the previous regulatory package, the SMP designation was applied in a fairly mechanical way – companies with 25% of some rather broadly defined markets were designated as having SMP. Under the NRF, the SMP designation is made on the basis of a competition law style analysis. Essentially, a company has SMP if it is "dominant", as defined by the European Court of Justice in Hoffman-Laroche, on one of the relevant antitrust markets listed in the Recommendation.

[85] Directive 2002/19/EC of the European Parliament and of the Council of 7 March 2002 on access to, and interconnection of, electronic communications networks and services, OJ 24.4.2002, L 108/7.

[86] Directive 2002/22/EC of the European Parliament and of the Council of 7 March 2002 on universal service and users' rights relating to electronic communications networks and services, OJ 24.4.2002, L 108/51.

take up of unbundled loops.[87] The Commission identified two problems. The first was that call and access tariffs had not been rebalanced in many countries, with the result that the retail price for line rentals was lower than the price of an unbundled loop. To remedy this problem, the Commission took actions against Italy, Spain and France. The second problem concerned the conduct of some of the companies supplying the loops. As a result the Commission opened two formal proceedings under Article 82: one against Wanadoo – France Télécom's Internet subsidiary – concerning a possible predatory pricing strategy in high-speed Internet access,[88] and one against Deutsche Telekom related to a possible problem of margin squeeze in local access.[89]

- **Promoting access – promoting fragmentation**

The other major goal of the NRF is to create a regulatory framework that facilitates the transition from regulation to competition. To achieve that goal, the NRF embraces the view that regulation should be used to promote actively service-based competition by facilitating access to existing infrastructure. By encouraging entry of service-based competitors, the NRF promotes market fragmentation over the option of long-run competition. The promotion of access-based competition is reflected in many different ways. For example, if a wholesale market is listed in the Commission's *Recommendation* and the NRA finds there is an SMP operator, then according to the *Access Directive* the NRA has little real choice but to ensure that third parties have access to the wholesale services in question.[90] Since the markets listed by the Commission in its *Recommendation* come in pairs – for each retail market there is an associated wholesale market – the practical effect of the Commission's policy will be to require NRAs to deal with problems of SMP in retail products through providing access to wholesale inputs.[91]

The rationale for this choice of regulatory approach is the belief that access-based competition is the *proper* response to market power – a response that is consistent with the principles of competition law,[92] but that successfully regulates market power when antitrust law would be insufficient to do so.[93] More recently, this choice has been defended by

---

[87] See press release IP/00/765, Commission asks incumbent operators to provide information on local loop access, Brussels, 12 July 2000.

[88] Press release IP/01/1899, High-speed Internet access: Commission suspects Wanadoo (France) of abusing its dominant position, Brussels, 21 December 2001.

[89] *Speech/02/323,* Speech by Mr. Mario Monti, European Commissioner for Competition Policy, Getting competition in local access, Public hearing sector enquiry local loop unbundling, Brussels – 8 July 2002.

[90] *Directive 2002/19/EC of the European Parliament and of the Council of 7 March 2002 on access to, and interconnection of, electronic communications networks and services*, OJ 24.4.2002, L 108/7, art 12.

[91] *Commission Recommendation 2003/311/EC of 11/02/2003 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 communication networks and services*, OJ 8.5.2003 L 114/45.

[92] *See Directive 2002/21/EC of the European Parliament and of the Council of 7 March 2002 on a common regulatory framework for electronic communications networks and services*, OJ 24.4.2002, L 108/33, recital 27.

[93] See Pierre Buiges, *The Competition Policy Approach,* in The economics of Antitrust and Regulation in Telecommunications 24, (P. Buigues and Patrick Rey eds., Edward Elgar, 2004).

reference to the so-called "ladder of investment theory",[94] which denies that a conflict between the promotion of short-run competition and the viability of long-run competition exists.[95] In the next Sections, we investigate whether these rationalisations are indeed justified.

4.2 The NRF and competition law

The features of the NRF that have received most attention are (a) that it uses concepts and tools taken from competition law to identify SMP, [96] and (b) that regulatory intervention is intended to be limited to those situations where "national and Community competition law remedies are not sufficient to address the problem".[97] The fact that NRAs cannot impose regulatory obligations on companies unless (a) they have SMP and (b) competition law remedies are deemed insufficient, is taken to imply that under the NRF regulatory intervention will take place only where it is absolutely necessary.[98] This self-restraint is to be praised. However, we remain sceptical about whether regulatory practice will exhibit such restraint for two reasons.

First, while the NRF relies on competition law concepts and tools to identify the markets to be subject to *ex ante* regulation and to assess whether undertakings have SMP, it does not follow competition law precedent when setting up the conditions under which access to the assets of an SMP operator must be granted. Therefore, we believe that the NRAs may end up imposing remedies that violate the principles of EC competition law, in which case intervention likely will be excessive and unjustified. [99]

For example, if in a given country a company is found to have SMP in the supply of a service included in one of the wholesale markets listed in the *Recommendation,* the NRA may impose on the dominant operator an obligation to provide that service to third parties

---

[94] Mario Monti European Commissioner for Competition Policy, Speech/04/37, Remarks at the European Regulators Group Hearing on Remedies, Public hearing on remedies under the new regulatory framework for electronic communications networks and services, 26 January 2004.

[95] Note that we do not claim that this theory constituted the guiding force behind the NRF. On the contrary, we regard the investment ladder theory as an *ex post* rationalization of the NRF's choice in favour of access-based competition. The theory is now playing a major role in current discussions about remedies and appears to have been embraced by the Commission. See Section 4.2 below.

[96] The Economics of Antitrust and Regulation in Telecommunications (P. Buigues & Patrick Rey eds., Edward Elgar, 2004).

[97] See *supra* note 92.

[98] Pierre Buiges, *The Competition Policy Approach,* in The Economics of Antitrust and Regulation in Telecommunications 17, (P. Buigues & Patrick Rey eds., Edward Elgar, 2004).

[99] We do not seem to be the only ones concerned by the potential over-interventionist implications of applying the NRF. See, e.g., Alexandre de Streel *The Integration of Competition Law Principles in the New European Regulatory Framework for Electronic Communications,* World Competition, 26(3), 2003, at p. 497.

> "..., inter alia, *in situations where the [NRA] considers that denial of access or unreasonable terms and conditions having a similar effect would hinder the emergence of a sustainable competitive market at the retail level, or would not be in the end-user's interest."*[100]

These circumstances are significantly broader than those applying under competition law,[101] which requires access only in *exceptional circumstances.*[102] These are (a) the asset to which access is requested is indispensable to compete; (b) the refusal is such as to reserve to the asset owner a secondary market by eliminating all competition on that market, and (c) the refusal is not justified by objective considerations.[103] These limitations are not the result of practical limits on the application of competition law because of, for example, inability to obtain reliable information about costs. On the contrary, they are there to balance short-run and long-run considerations.[104] As Advocate General Jacobs stated in *Bronner:*[105]

> *"[I]n the long term it is generally pro-competitive and in the interest of consumers to allow a company to retain for its own use facilities which it has developed for the purpose of its business. For example, if access to a production, purchasing or distribution facility were allowed too easily there would be no incentive for a competitor to develop competing facilities. Thus while competition was increased in the short term, it would be reduced in the long term. Moreover, the incentive for a dominant undertaking to invest in efficient facilities would be reduced if its competitors were, upon request, able to share the benefits. Thus the mere fact that by retaining a facility for its own use a dominant undertaking retains an advantage over a competitor cannot justify requiring access to it."*[106]

The need to balance short-run and long-run considerations and to promote investment and innovation is just as important in

---

[100] *Directive 2002/19/EC of the European Parliament and of the Council of 7 March 2002 on access to, and interconnection of, electronic communications networks and services,* OJ 24.4.2002, L 108/7, Article 12.

[101] See Alexandre de Streel, *The Integration of Competition Law Principles in the New European Regulatory Framework for Electronic Communications*, 26(3) World Competition, 489, 505 (2003).

[102] The expression "exceptional circumstances" in the context of refusal to supply violations of Article 82 of the EC Treaty was first used by the European courts in *Magill,* Joined cases C-241/91 and C-242/91 *RTE and ITP v. Commission* [1995] ECR I-743, upholding an earlier Commission decision IV/31.851 *Magill* TV Guide/ITP, BBC and RTE OJ 1989 L78 p 43 and CFI judgement (Case T-69/89 *RTE v. Commission* [1991] ECR II-485 and Case T-76/89 *ITP v. Commission* [1991] ECR II-575).

[103] If the essential facility is an intellectual property (IP) right, there is the undertaking which requested the license intends to offer products or services not offered by the IP owner and for which there is potential consumer demand.

[104] See Christian Ahlborn, David Evans and A. Jorge Padilla, *The Logic and Limits of the `Exceptional Circumstances Test* in *Bronner, Magill* and *IMS,"* August 2004, mimeo.

[105] Case C-7/97 *Oscar Bronner GmbH v. Mediaprint Zeitungs und Zeitschriftenverlag GmbH* [1998] ECR I-7791.

[106] Ibid., at p. 57.

telecommunications industries as elsewhere. Thus, there is no economic justification for a more expansive interpretation of the obligation to deal – or a narrower definition of property rights – in the telecoms industry, in particular in those markets where innovation and investment plays a paramount role. Access under the NRF should have been restricted to the very same exceptional circumstances found by the European Court of Justice in *Magill, Bronner* and *IMS Health*.[107]

Second, we cannot see how the NRAs can in practice limit their regulatory interventions to "those markets for which competition-law remedies are insufficient to effectively redress possible market failures".[108] We are not sure that this is a meaningful statement. This is because:

- As noted by Cave and Crowther in a paper published in this volume, the NRF offers "nothing on how an NRA is to satisfy itself (as it is required to do) that competition law is 'insufficient' to resolve the particular market failure".[109]

- Under the NRF, the NRAs must intervene in markets where they find SMP, whether or not there is evidence of abuse. In contrast, competition law interventions are limited to situations where a firm enjoying a dominant position engages in abusive conduct. So regulatory measures might be adopted in markets where the competition authorities would not find any reason to intervene *even if* competition-law remedies would prove effective if deployed.

- Finally, the imposition of regulatory remedies, such as caps on access prices, may limit the scope of rivalry and, thus, foreclose the development of market conditions where competition policy would be most effective.[110] In other words, regulatory intervention may raise the barriers to entry that are then used to justify the use of (perpetual) regulation rather than competition law.

4.3 The NRF and the "ladder of investment"

As we saw above, the NRF emphasises access-based remedies and so unambiguously promotes a model of service-based competition. We have seen that competition law principles do not justify this emphasis. What does then? Whatever the origins of this policy, it is currently being justified by claiming that regulators do not need to choose between promoting long-

---

[107] See *supra notes* 102 and 105, and Case C-418/01 *IMS Health GmbH & Co. OHG v. NDC Health GmbH & Co. KG.*

[108] Ibid. at p. 16.

[109] Martin Cave and Peter Crowther, *Co-ordinating regulation and competition law – ex ante and ex post,* in The Pros and Cons of Antitrust in Deregulated Markets, (Mats Bergman & Arvid Nilsson eds., Swedish Competition Authority, forthcoming 2004) at page 15.

[110] See Philip Lowe, *Introduction to the Roundtable on the Economics of Antitrust and Regulation in the Telecom. Sector,* in The Economics of Antitrust and Regulation in Telecommunications 240 (P. Buigues & Patrick Rey eds., Edward Elgar, 2004).

run infrastructure competition and investment and ensuring consumer benefits in the short run – we can have both by promoting service-based competition, we are told.

The idea that there is *no* trade-off between the long run and the short run is a radical departure from the past. The history of regulation has shown that, while the participants have often had different views about the goals of regulation and the means to achieve them, they have agreed that there is a choice between policies that emphasize static efficiency and those that give preference to dynamic efficiency. This choice has been the subject of heated debate since at least 1987, when the Commission published its Green Paper on the development of the European telecoms markets. The theory used to underpin the new claim that there is no trade-off is the "ladder of investment" theory. This section explains the logic and limits of this theory, drawing on two documents: the statement of the European Regulators Group (ERG) on remedies[111] and Commissioner Monti's speech at the ECTA regulatory conference in 2003.[112]

The "ladder of investment" theory hinges upon four Propositions. If one of these Propositions failed to hold, then theory would be incorrect, the ERG and the Commission would be wrong, and their policies might cause serious problems for European telecoms markets in the long run. These Propositions are:

Proposition 1. Infrastructure competition is the only means to sustainable competition in telecommunications.

Proposition 1 asserts that infrastructure competition is the most important goal for regulators. Both the ERG statement and Commissioner Monti claim that achieving self-sustaining competition in the long run constitutes the key goal for regulation. Commissioner Monti, for example, states: *"The aim of regulation is, or should be, creating a pro-competitive environment in the long run."*[113] Likewise, the ERG says,

> *environment in the long run."*[114] Likewise, the ERG says,
>
> *"While NRAs have to protect consumers against exploitative behaviour and inefficiencies where significant market power exists, the ultimate goal is to promote self-sustaining competition...."*[115]

Moreover, they both appear to believe that self-sustaining competition is facilities based – infrastructure competition. Commissioner Monti claims,

---

[111] *ERG: ERG Common Position on the approach to Appropriate remedies in the new regulatory framework,* 23 April 2004.

[112] On the "ladder of investment" theory, see also *supra note* 44.

[113] Mario Monti European Commissioner for Competition Policy, Speech/03/604, Competition and Regulation in the Telecom Industry: The way forward, ECTA Conference Brussels, Conrad Hotel 10 December 2003.

[114] Mario Monti European Commissioner for Competition Policy, Speech/03/604, Competition and Regulation in the Telecom Industry: The way forward, ECTA Conference Brussels, Conrad Hotel 10 December 2003.

[115] *ERG: ERG Common Position on the approach to Appropriate remedies in the new regulatory framework,* 23 April 2004, p. 15.

> *"In the longer term the regulatory framework should privilege operators which base their competitive advantage on building their own infrastructure, simply because they are those who are likely to best improve the competitive conditions of the market."*[116]

The ERG also sustains that remedies must be chosen to foster facilities-based competition because:

> *"Competition over competing infrastructure has many advantages. The pressure to minimise costs is exerted over the whole value chain. This will induce greater scope for innovation, process innovation etc, which creates a downward dynamic for costs. Consumers also benefit from more diversified offerings, which correspond more closely to their individual needs. There is general agreement that a great potential harm to welfare occurs when replication is feasible but not promoted."*[117]

Proposition 2. Service competition is a necessary prerequisite for infrastructure competition.

Proposition 2 makes the strong assertion that without the development of service competition it would not be possible to obtain competition between alternative infrastructures, which from Proposition 1 is the only type of competition that is sustainable. It is clear that the Commission is behind this Proposition,

> *"Competition would never be able to develop, in the short term, if entrants were not able to gain access to the incumbent operator's network to start offering services."*[118]

And for the ERG service competition is just a preliminary phase on the road from monopoly to sustainable competition:

> *"[The new framework] is a major step in the transition path between the vertically integrated monopolies of the past and the normal competition process (governed exclusively, where appropriate, by competition law)."*[119]

Proposition 3. Service competition is only possible through a regulated infrastructure.

This Proposition asserts that the only means to achieve service competition is through regulating the existing infrastructure. If it were not to hold, there would be no need to regulate infrastructure and service competition. Both

---

[116] Ibid.

[117] ERG: *ERG Common Position on the approach to Appropriate remedies in the new regulatory framework,* 23 April 2004, p. 67.

[118] Mario Monti European Commissioner for Competition Policy, Speech/03/604, Competition and Regulation in the Telecom Industry: The way forward, ECTA Conference Brussels, Conrad Hotel 10 December 2003.

[119] ERG: *ERG Common Position on the approach to appropriate remedies in the new regulatory framework,* 23 April 2004, p. 19.

the Commission and the ERG appear to believe that Proposition 3 holds true in practice. For Commissioner Monti,

> *"Competition would never be able to develop, in the short term, if entrants were not able to gain access to the incumbent operator's network to start offering services".*[120]

The ERG concurs,

> *"Excessive prices on the retail market should first be addressed at the wholesale level, e.g. by ensuring access at cost-oriented prices."*[121]

Proposition 4. The trade-off between service competition and infrastructure competition is reconciled by the "ladder of investment"

This is the key of the "investment ladder" theory, as this is the Proposition that states that it is possible to simultaneously gain the benefits from service competition and infrastructure competition. Service competition and infrastructure competition are reconciled, because the former will give way over time to the latter through the workings of a "ladder of investment" carefully managed by regulators. For example, Commissioner Monti states, "In order to reconcile access-based and facilities-based competition it is necessary to take account of the time dimension. NRAs should provide incentives for competitors to seek access from the incumbent in the shorter term, and to rely increasingly more on building their own infrastructure in the longer term."[122]

The ERG agrees,

> *"Service competition based on regulated access at cost-oriented prices can be (and in general is) the vehicle for long-term infrastructure competition. With this new entrants can decide on their investment in a step-by-step way and can establish a customer base (critical mass) before they go to the next step of deploying their own infrastructure."*[123]

Both documents rely on the "ladder of investment" theory to reconcile the short-term and long run goals of regulation. In their view, there is *"... not necessarily a contradiction between access-based and facilities-based*

---

[120] Ibid.

[121] ERG: ERG *Common Position on the approach to appropriate remedies in the new regulatory framework*, 23 April 2004, p. 109.

[122] Mario Monti European Commissioner for Competition Policy, Speech/03/604, Competition and Regulation in the Telecom Industry: The way forward, ECTA Conference Brussels, Conrad Hotel 10 December 2003.

[123] ERG: ERG *Common Position on the approach to Appropriate remedies in the new regulatory framework*, 23 April 2004, p. 68.

competition."[124] The central premise of this theory is that service competition will lead to access infrastructure competition *if it is managed correctly* – e.g., by fine-tuning access prices over time. Although this notion has been around for some time, the idea of managing competition through a complex tangle of regulatory products towards full infrastructure competition has been put forward more recently by Martin Cave and Ingo Vogelsang.[125]

The mechanism underlying the ladder of investment is in theory rather simple. Initially, regulation will encourage entry into the wholesale market where SMP is found to exist by setting low (possibly below cost) access prices for those assets that entrants find difficult or costly to replicate. *"At the outset this might include a large number of assets, which initially are complements to the entrant's investment, but with time become substitutes."*[126] Over time, once entrants consolidate their beachheads and start to earn positive rents, regulators will proceed to increase access prices. They will have to do so in descending order of asset replicability: those assets that are easier to replicate will experiment the price increase first. Entrants are expected to respond to those price increases by investing in assets that were previously regarded as non-replicable.[127] In this way, what started as service-based competition will blossom into self-sustaining infrastructure competition.

Are these Propositions plausible? Is the "ladder of investment" that is at the centre of Proposition 4 just a theory with no factual support? We certainly agree with Proposition 1 for the reasons discussed in Section 4.5 below. Proposition 3 is important only if Propositions 2 and 4 hold true, so we will not discuss it here. As regards Propositions 2 and 4, the answer to the questions above depends in turn on how one answers the following two questions. First, do regulators have the information, time and competence to micro-manage the evolution of a market from a service-based to a facilities-based? Or, in other words, are they in a position to fine tune the evolution of access prices over time as the ladder of investment theory would require? Second, access-based remedies will tend create a relatively fragmented industry. The companies offering services at one vertical layer may not be the same as the ones active in another. And there may be many companies at each layer. Is this fragmented structure a good basis for sustainable competition?

Propositions 2-4 will be satisfied if the answers to these two questions are affirmative – i.e., if one believes that regulators have enough information and competence to micro-manage the evolution of competition, and that the very fragmented form of competition that access-based entry will promote can be sustained without the support of perpetual regulation so that it eventually gives rise to a market characterised by infrastructure competition. But if the answer to at least one of these questions is negative, then the Commission's and the ERG's shared belief in a regulatory free

---

[124] Mario Monti, European Commissioner for Competition Policy, Speech/04/37, Remarks at the European Regulators Group Hearing on Remedies, Public hearing on remedies under the new regulatory framework for electronic communications networks and services, 26 January 2004.

[125] Martin Cave & Ingo Vogelsang, *How access pricing and entry interact*, 27 Telecomm. Pol'y., 717, 717-727 (2003). The ideas in this paper where presented to the Commission in: Martin Cave, *Remedies for Broadband Services* DG InfoSoc, September 2003.

[126] Ibid. p. 724.

[127] *"The entrant passes progressively through several stages of infrastructure competition, as it ascends a "ladder' of infrastructure construction."* Ibid. p. 10.

lunch will be unjustified. The "ladder of investment" theory is therefore based on the *presumptions* – or prior beliefs – that both questions should be answered affirmatively. The next two Sections investigate whether those presumptions are supported by economic theory and evidence.

4.4 First presumption: regulatory omnipotence?

Those who suggest that regulators no longer need to choose between the short and the long run because the two can be reconciled via the "ladder of investment" have a very optimistic view of the resources, information and competence available to regulators. The "ladder of investment" theory places on regulators a heavy responsibility – not only must they act to make sure that consumers are protected in the short term, but they must also manage the evolution of market structure. This would be a challenge even in a well-understood and stable industry. And even more so in industries, such as the electronic communications industry, that are neither well understood nor stable. The past 20 years are full of examples where regulators and industry participants have acted on expectations that turned out to be wrong. For example, in the late '90s commentators believed that 3G, with its ability to transmit data and video over mobile links, would revolutionise telecoms. Companies bid huge sums for 3G licences – $125 billion in Europe alone – and planned to spend as much again on infrastructure. However, technical difficulties have delayed the start of fully-fledged 3G services, and in the meantime alternative technologies such as Wi-Fi and further improvements of fixed line services (e.g. ADSL) are undermining the profit expectation of 3G mobile operators.[128] Given the difficulty of knowing how technology will evolve, and what customers really want, the task facing regulators that rely on the "ladder of investment" is truly daunting.

Getting it right is difficult...

As we saw above, the "ladder of investment" presupposes that the regulator will lead entrants through a clear sequence of investments. It will first identify the bottom rung – a replicable asset that it considers a suitable basis for entry. Then the regulator will encourage a cohort of new suppliers to invest in that asset and start providing services by making sure they have cheap access to all other assets of the incumbent (including especially those that are non-replicable) which are needed to complement the one they have invested in themselves. Once the cohort of entrants have finished building their first asset, the regulator will decide what assets they should invest in next and raise the price of access to them while keeping access to the remaining ones low. And so on.

The choice of the initial rung and of the accompanying access remedies will obviously have a big impact on the direction of investment. Subsequent choices will add to this impact. Unfortunately, it is complicated to get it

[128] Other examples include the mixed success of infrastructure competition in the UK, the failure of the big global alliances and JVs of the 80's, the disappointing extent of LLU take up, and the collapse of the US competing providers of long distance conveyance.

right. Regulators will find it very difficult to decide on the right sequence of rungs to construct the ladder, in part because there are so many different elements in a communications network, which give rise to many conceivable access services. But there is a more fundamental problem. Different sorts of entrant are favoured by different ladders. For example, urban local loop companies would find it easier to justify their investments if the starting rung (the one for which access is not regulated) was local access while at the same time long distance conveyance was cheap. On the other hand, a Web portal, which wants to use its brand as a content provider to extend its services to access provision, will want to be able to buy cheap broadband local access. None of this would be a problem if regulators could construct a number of different ladders and let entrants choose which one to use. But they cannot, because the different ladders would conflict with each other: it is impossible to grant cheap local access to please the Web portal while *at the same time* denying it to please the local loop companies.

Regulators cannot avoid committing to a given investment path by adopting a flexible strategy of first trying out one ladder and, if it does not work, moving to a different one. The most obvious reason for this is that entry under the first ladder would make it costly for the regulator to remove that ladder at a later date. Those entrants would have created assets that would be wasted if the regulator pursued a new strategy that undermined their business case. It would also be politically difficult to remove that ladder once companies have become dependent on it. The "ladder of investment" can only work if the regulator is committed to its choice. The regulator must commit to fostering investment in subsequent rungs of the ladder by raising the price of some access products after a while. This will encourage early entrants to move on from their initial service-based business model to one that is more based on facilities - i.e., to build new rungs in the ladder. Regulators must be committed to raising access prices *even if* the initial entrants still depend on them so that the increases lead some of them to exit. There are two reasons for this. The *first* is that otherwise there would be too much and inefficient entry by access-based competitors – companies would invest knowing that they could not afford to move further up the ladder at a later date, but confident that the regulator, when it comes to it, would not remove the low access prices on which they rely. The *second* reason is that, absent commitment, entrants would have perverse incentives. Regulators are likely to honour a commitment to make access more expensive if entrants are well established, but to renege on that commitment if entrants are still dependent on the cheap access. But this means entrants are punished if they do well, which gives them incentives to shirk.

Despite the importance of commitment, regulators have a poor history in this regard. For example, in the Netherlands the regulator, OPTA, announced that it would adopt a policy of staged price increases on the unbundled loops. The plan was later abandoned because the Commission mandated the price basis for local loops, so that OPTA no longer had the discretion to fine tune access prices over time. In Canada, the regulator decided in 1997 to divide the country into two regions - large urban, and small urban/rural. In the latter, unbundled loops would be available for an indefinite period as an "essential facility". In the former, they would be available for five years only, giving entrants a window of opportunity to build their own loops. Once again the plan to withdraw the mandatory

access to loops was abandoned in 2001, as take up of the loop was lower than expected.

... And those who tried failed

The "ladder of investment" theory induces memories of trade theory in the 1970's, which contained a policy proposal similar to that contained in the "ladder of investment" – referred to as "infant industry protection".[129] The idea was that industries in developing countries needed to be protected from overseas competition by trade barriers. These barriers would allow them to establish themselves and grow large enough to take advantage of economies of scale. Once established and efficient the trade barriers could be removed and local consumers would finally be able to take advantage of low prices. Many countries, especially in Latin America, based their trade policies on these theories, but the consensus is now that the policies were a failure. Protected entrants had distorted incentives and a culture of dependence. They became reliant on guaranteed margins and their incentives were to engage in wasteful lobbying, rather than becoming more efficient. Empirical studies showed that there was little evidence of faster increases in output or growth when compared to non-protected sectors,[130] leading eminent economists such as Paul Krugman to conclude that;

> *"The economic cautions about the difficulty of formulating useful interventions and the political economy concerns that intervention may go astray combine into a new case for free trade."*[131]

4.5 Second presumption: sustainable fragmentation?

The second presumption underlying the "ladder of investment" is that the fragmented form of competition that access-based entry will give rise to can be sustained in the long run without the support of perpetual regulation. In what follows, we will show that this second presumption may also be unjustified. As Breyer, we also conclude that,

> *"Increased sharing by itself does not actually mean increased competition. It is in the unshared, not in the shared portions of the enterprise that meaningful competition would be likely to emerge. Rules that force firms to share every resource or element of the business would create not competition, but pervasive regulation, for the regulators, not the market place, would set the relevant terms."*[132]

---

[129] The infant industry argument originated in the trade theories departure from the assumption of perfectly competitive markets to oligopolistic markets, see Paul Krugman & Maurice Obstfeld, Internaltional Economics: Theory and Practice 207-208 (McGrawhill, Fourth ed., 1999) for a good review of these arguments.

[130] See Anne Krueger & Baran Tuncer, *An Empirical Test of the Infant Industry Argument,* 72(5) The Am. Econ. Rev., 1142, 1142-1153 (1982).

[131] Paul Krugman, *Is Free Trade Passe?*, 1(2) J. of Econ. Persp., 131, 131-144 (1987).

[132] Quoted in Alfred E. Kahn, Whom the Gods Would Destroy or How Not To Deregulate (AEI Press, 2001).

The first effect of the application of the NRF by the NRAs should be an increase in the degree of fragmentation – i.e., a reduction in the degree of concentration – of European telecoms markets. After all, the NRF mandates the imposition of remedies, such as third-party access, in all those markets, which are found to be not "effectively competitive".[133] Since for the purposes of applying the NRF the Commission has defined effective competition as the absence of dominance,[134] and the first and most pre-eminent factor in considering whether a firm enjoys a position of dominance (alone or collectively with others) is its market share, we have that the application of the NRF will result in regulatory intervention on all those markets characterized by high concentration and that such an intervention will be lifted only if those markets become more fragmented.

Given the nature of the remedies specified in the Access Directive, it is most likely that the new entrants initially will offer services in direct competition with those of the SMP operators whose networks they share. As those entrants are dependent on the wholesale inputs offered by the SMP operators, the potential for product differentiation is arguably less than under facilities-based competition. Indeed, the access-based remedies specified in the NRF will allow entrants to purchase inputs from the incumbents, so as to either resell them to end consumers or repackage them with other products of their own. A good example of this tendency towards product homogeneity in shared networks is given by broadband access in the UK. In the UK, broadband is provided either by ADSL over BT's network or by cable networks. Cable networks were the first to introduce broadband in the UK and have also been active in innovating by developing different speed packages. The two cable networks have chosen different strategies. NTL has introduced a low speed, low price broadband product. Conversely, Telewest has chosen to target the premium end of the market and to offer high speeds. However, whilst the cable networks have been active in trying to differentiate their products making use of their own infrastructures, the same has not been true in the ADSL segment of the market, where competitors rely on purchasing wholesale inputs from BT. Until 2004, virtually all ADSL competition focused on an identical 512 Kbits product.

The proponents of the "ladder of investment" appear to believe that, despite their undifferentiated offerings, those access-based entrants will be able in the short run to acquire an installed base of customers on which to make positive rents, and that those rents will allow them to develop their own networks step by step, so that over time the access-based entrants

[133] See *Commission guidelines on market analysis and the assessment of significant market power under the Community regulatory framework for electronic communications networks and services*, 2002/C 165/03, para. 19.

become infrastructure competitors.[135] According to this view, in the long run, the markets subject to intervention will remain fragmented but populated by competitors who own their own facilities. But is it reasonable to expect that entrants offering relatively undifferentiated services succeed in the marketplace, accumulating the rents that could allow them to develop their own networks? Is it possible to reconcile market fragmentation with sustained investment and innovation? The answer to both questions appears to be a qualified no.

Without innovative services no entrant will be able to induce consumers to switch providers unless it offers significant discounts. The problem is that in the absence of product differentiation the entrant will not be able to profit from the business stolen to the incumbents, because to retain its customers it must keep time after time the same low prices that were used to attract them in the first place. It will not be able to recover even its marketing costs. After a period accumulating losses, the entrant will exit. As Crandall notes, in the US,

> *"... virtually every new carrier that tried to enter by using the unbundled local loop to offer the same services as incumbents in a marketplace with declining total revenues has, not surprisingly, failed and disappeared in just a few years."*[136]

So the presumption that access-based entrants will become facilities-based competitors appears to be unjustified unless one of the following conditions holds: (a) the entrants succeed in differentiating their product offerings in ways that prove attractive to consumers, or (b) access is limited to a few entrants so as to avoid the dilution of rents. If neither (a) nor (b) holds, then the promotion of access-based entry will not result in infrastructure competition and, what is more, it may lead to the perpetual regulation of the market and to a significant reduction in industry-wide investment. A fragmented market sustained by perpetual access regulation is incompatible with sustained investment in infrastructure. As shown by Jerry Hausman, among others, compulsory third-party access at (long-run incremental cost) may chill the incentives of telecoms operators to invest in infrastructure.[137] This is because the company investing in the new or

---

[135] This belief seems to be supported by recent experiences in the UK and France. For example, in the UK, mobile operators used to lease all their backhaul from BT, but are starting to build their own now. Also, the main ISPs in the UK appear to rely on more and more basic wholesale inputs from BT. As they get bigger they integrate backwards into network provision to benefit from potential economies of scale. Finally, in France, Illiad – a broadband internet provider – has developed a successful vertically integrated business that appears to have benefited from cheap unbundled local loops. Note, however, that (a) none of these stories is the result of an access pricing policy where access prices are fine tuned over time, and (b) in all these cases competitors offer differentiated products. The success of Illiad, for example, appears to owe more to its unmatched triple play combination than to the regulator's unbundling policy. Indeed, according to *The Economist* (The Broader Art of Deregulation, 19 August 2004), *"[u]nder half of its customers use unbundled lines: the rest were attracted to the firm by its record of innovation."*

[136] Robert W. Crandall, *The Remedy for the "Bottleneck Monopoly" in Telecom: Isolate It, Share It, or Ignore It?*, University of Chicago Law Review, (forthcoming) at page 10. For further details, see Robert W. Crandall, Competition and Chaos: U.S. Telecommunications since 1996, Brookings, 2004 (forthcoming), Chapter 4.

[137] See Jerry Hausman, *The Effects of Sunk Costs in Telecom.* Regulation, in Real Options: The New Investment Theory and Its Implications for Telecommunications Economics, (James Alleman & Eli Noam eds., Kluwer, 1999).

improved infrastructure is placed at a competitive disadvantage vis-à-vis its rivals. For the investing company, the assets are sunk. Yet its competitors, who initially enjoy access to the facility, are free to adopt any superior technology if and when it becomes available without having to contribute to funding the initial investment.

4.6 Conclusions

The approach to regulation embodied in the NRF has been described as "pre-emptive competition law";[138] a way to achieve "the appropriate balance between *ex ante* and *ex post* regulatory approaches in the specific context of the network industry";[139] and "a major step down the transition path between monopoly and normal competition, governed exclusively by generic competition law".[140] But as we have seen in this paper, there are reasons to believe that it is none of that.

First, while it is true that the NRF uses the tools and concepts of competition policy to define markets and identify dominance, it may mandate access remedies in circumstances that are much broader than those laid out by the European Court of Justice in *Magill*, *Bronner* and *IMS Health*.[141] Thus regulatory intervention under the NRF likely will go beyond the limits imposed by competition law. Second, in contrast to the claims of those who promote the new regime, competition law and *ex ante* regulation are competing instruments for the control of market power. Those who defend the complementarity of both approaches fail to see that *ex ante* intervention necessarily pre-empts *ex post* enforcement. And third, the new regime is unlike to result in less regulation but in a different, and possibly perpetual, form of access-based regulation.

The NRF is based on the expectation that a clever access pricing policy aimed at facilitating service competition in the short run will promote long-run infrastructure competition and investment while ensuring short-run efficiency. This expectation is based on two presumptions: that today's regulators can successfully micromanage the evolution of competition in electronic communications markets, and that the fragmented form of competition that access-based entry likely will promote can be sustained in the long run without the support of perpetual regulation. We have found no theoretical or empirical support for the first presumption and have expressed serious doubts about the second.

To summarize, although we value positively the NRF's attempt to clarify and harmonize telecoms regulation in the EU, we are sceptical about the role of competition policy in the framing of the NRF and its alleged flexibility. The NRF embodies an approach to regulating telecoms that promotes access-based competition that, in our view, cannot be derived

---

[138] Alexandre de Streel, *The Integration of Competition Law Principles in the New European Regulatory Framework for Electronic Communications*, 26(3) World Competition, 489, 490 (2003).

[139] Pierre Buiges, *The Competition Policy Approach,* in The economics of Antitrust and Regulation in Telecommunications 26 (P. Buigues & Patrick Rey eds., Edward Elgar, 2004).

[140] Martin Cave & P. Larouche, *European Communications at Crossroads*, report of a CEPS working party, Brussels, October 2001.

[141] See supra notes 101, 105, and 107.

from the principles of competition law. We are unconvinced by those who claim that there is no conflict between access-based and facilities-based competition. Regulators are unlikely to have the detailed knowledge that is required to micro-manage the investment ladder that could transform the service providers of today into the facilities-based competitors of tomorrow. Instead, we believe that regulators still face a fundamental choice between access-based entry and facilities-based entry. This choice will in turn determine the relative roles played by *ex ante* regulation and competition law in the telecoms industries of the Member states. And, most importantly, it will influence greatly the performance of European telecoms companies and the welfare of European consumers.

5. Managing unilateral market power in electricity

Frank A. Wolak

5.1 Introduction

The past two decades of international experience with wholesale electricity markets has demonstrated that significant consumer harm can result from firms simply engaging in unilateral profit-maximizing behavior given the actions of their competitors. Different from other product markets, coordinated actions among suppliers or the concentration of production capacity in the hands of small number of firms is unnecessary for electricity suppliers to raise prices substantially above competitive levels. A number of wholesale electricity markets with Hirshman-Herfindahl Indexes (HHIs) that would not raise market power concerns if they were from other industries have been subject to severe market power problems. In addition, for all of these market power episodes, the relevant competition authorities have not found evidence of coordinated actions among suppliers to raise prices in violation of the competition or antitrust law. These facts provide strong evidence that competition or antitrust policy as it is applied to other industries may be insufficient to protect electricity consumers.

This paper argues that the technology of electricity production and remnants of the former monopoly regime imply that conventional competition policy must be augmented with an industry-specific regulator endowed with a pre-specified set of responsibilities. This combination of regulatory oversight and competition law will provide consumers with the same level of market power protection they receive for other products from conventional competition law. An industry-specific regulator is necessary because: (1) unilateral market power problems can be extremely difficult to predict, and (2) they can impose significant economic harm for a sustained period of time when they do occur. Moreover, how this unilateral market power is exercised and the harm it causes arises from a number of unique features of wholesale electricity markets. Clearly specified regulatory safeguards tailored to the electricity supply industry are needed to prevent the harmful exercise of unilateral market power before it can occur and rapidly implement the necessary remedies if it does occur. The primary goal of this regulatory process should be to prevent market participant behavior that significantly degrades system reliability and market efficiency.

It is important to emphasize that it is impossible for the regulator to prevent firms from exercising unilateral market. In fact, markets function most efficiently when suppliers have high-powered incentives to exercise all available unilateral market power and there are few barriers to entry. The role of the regulatory process is to ensure that the conditions necessary for vigorous competition exist and to limit the economic harm associated with the exercise of unilateral market power when they do not exist.

Regulatory mechanisms that attempt to prevent all exercise of unilateral market power can introduce market inefficiencies that cause more economic harm than the market power they are attempting to prevent.

This regulatory process should be self-correcting in the sense that there are pre-specified regulatory responses to certain market outcomes and market participant behavior that significantly degrades system reliability and market efficiency that are known to all market participants in advance. The aftermath of the events in the California electricity market from June 2000 to June 2001 and the Enron bankruptcy have demonstrated that ex regulatory intervention is extremely costly and very unlikely to remedy the harm done. This experience and similar ones from other markets around the world argues in favor of a prospective regulatory process that anticipates possible harmful market outcomes and builds in incentives for market participants to solve these problems without the need to formulate new regulatory policy. Rather than attempt to formulate this policy under the intense scrutiny that accompanies harmful market outcomes occur, a more prudent regulatory policy is to build in mechanisms that anticipate and address as many of these potential harms as possible.

The technology of electricity supply creates an additional reason, not relevant in other markets, to limit the exercise of unilateral market power and encourage market participants to obey the wholesale market rules. Because all electricity must be delivered through a common transmission network, the actions of some market participants can impact the ability of other market participants to buy or sell electricity. A regulatory process is necessary to ensure that the unilateral profit-maximizing behavior of each market participant does not significantly degrade the reliability of the transmission network. The regulatory process should therefore strive to make the market rules transparent to all market participants and as straightforward as possible to enforce. The experience of a number of wholesale electricity markets suggests that imprecise market rules that appear to prohibit a wide range of behavior may lead to a less reliable transmission network than seemingly less restrictive but more clearly defined market rules that are substantively more straightforward to monitor and enforce. I will present examples of imprecise rules that detract from system reliability and market efficiency, along with recommendations for increasing the precision and enforceability of these rules.

Because actions that earn some market participants substantial profits can impose significant harm to market efficiency and system reliability, there must be market rules prohibiting these sorts of actions. Enforcing these market rules without causing more harm than is cured is extremely difficult. The prime example of this phenomenon in electricity market oversight is distinguishing between the unilateral exercise of market power and the poorly defined, but often invoked, concept of market manipulation. The widespread use of this term to describe the behavior of certain market participants in wholesale electricity markets occurs precisely because the unilateral exercise of market power in this industry can cause substantial consumer harm. US antitrust law contains no formal definition of market manipulation. However, the concept does have meaning in the US in the context of formal securities and commodities markets, where it is defined by the market rules. Futures exchanges define certain actions by market participants, for example, corners and squeezes, as prohibited market manipulation. A necessary condition for a finding of market manipulation is that the market participant took the prohibited action with the intent to

produce an outcome that significantly harmed other market participants. For example, a market participant may purchase a large position in the futures market and then take delivery on these contracts because of a genuine need for the commodity or because of a desire to corner the market. Building on the general framework used in financial markets, I propose a definition of market manipulation for wholesale electricity markets that the regulator should enforce, as well as an administrative process for the regulator to follow in order to make a finding of intent, the crucial step in a market manipulation determination. The goal of this administrative procedure is to limit economic harm to due to deviations from price-taking behavior by market participants, not to detect and punish all instances of alleged market manipulation.

The remainder of the paper proceeds as follows. The next section summarizes the reasons why the exercise of unilateral market power can cause so much harm in a seemingly competitive wholesale electricity market and why an industry specific regulator is essential to adequately addressing these market power problems. Section 5.3 describes the goal of the regulatory process and responsibilities of the industry-specific regulator. Section 5.4 describes guidelines for setting penalties and sanctions that ensure compliance with market rules. Section 5.5 suggests a mechanism for administering and enforcing a market rule prohibiting behavior harmful to system reliability and market efficiency. The paper closes with a discussion of the protocols for the competition authority to interact with the industry-specify regulator to provide maximal protection against the harmful market outcomes. This section also outlines a procedure for the relevant government authority to follow to phase out aspects of the industry-specific regulatory process as the wholesale market matures.

5.2 Why electricity is different

It is difficult to conceive of an industry more susceptible to the exercise of unilateral market power than electricity. It possesses virtually all of the product characteristics that enhance the ability of suppliers to exercise unilateral market power. Supply must equal demand at every instant in time and each location of the network. It is very costly to store and production is subject to extreme capacity constraints in the sense that it is impossible to get more than a pre-specified amount of energy from a given generation unit in an hour. Delivery of the product consumed must take place through a potentially congested transmission network. Historically, how it has been priced to final consumers makes the wholesale demand extremely inelastic, if not perfectly inelastic, with respect to the wholesale price. The technology of electricity production historically favored large generation facilities, and in most wholesale markets the vast majority of these facilities are owned by a relatively small number of firms. Finally, generation capacity ownership also tends to be concentrated in small geographic areas within these regional wholesale markets. All of these factors also make wholesale electricity markets substantially less competitive the shorter the time lag is between the date the sale is negotiated and the date delivery of the electricity occurs.

Electricity suppliers possess differing degrees of system-wide and local market power. System-wide market power arises from the capacity constraints in the production and the inelasticity of the aggregate wholesale demand for electricity, ignoring the impact of the transmission network. Local market power is the direct result of the fact that all electricity must be sold through a transmission network with finite carrying capacity. The geographic distribution of generation ownership and demand interact with the structure of the transmission network to create circumstances when a small number of suppliers or even one supplier is the only one able to meet an energy need at a given location in the transmission network. A supplier that is a monopolist or duopolist for a local energy need possesses substantial local market power.

The distinction between system-wide and local market power is often blurred by the choice of the relevant market. If electricity did not need to be delivered through a potentially congested transmission network subject to line losses, then it is difficult to imagine that any supplier could possess substantial system-wide market power in a market comprised of the entire US. Consequently, the market power that an electricity supplier possesses is a function of the size of the geographic market it competes in, which depends on the characteristics of the transmission network and location of final demand.

These two determinants of market power imply that a supplier possesses local market power regardless of the congestion management protocols used by the wholesale market. In single-price markets, zonal-pricing markets and nodal-pricing markets, local market power arises because the existing transmission network does not provide the supplier with sufficient competition to discipline its bidding behavior into the wholesale market. This is particularly the case in the US, where the rate of investment in the transmission network has persistently lagged behind the rate of investment in new generation capacity over the past 25 years.

Most of the existing transmission network in the US was designed to support a vertically-integrated utility regime that no longer exists. Particularly around large population centers and in geographically remote areas, the vertically-integrated utility used a mix of local generation units and transmission capacity to meet the annual demand for electricity in the region. Typically, the utility supplied the region's baseload energy needs from distant inexpensive units using high-voltage transmission lines. It used expensive generating units located near the load centers to meet the periodic demand peaks throughout the year. This combination of local generation and transmission capacity to deliver distant generation was the least-cost strategy for serving the utility's load in the former regime.

The transmission network that resulted from this strategy by the vertically integrated utility creates local market power problems in the new wholesale market regime because now the owner of the generating units located close to the load center may not own, and certainly does not operate, the transmission network. The owner of the local generation units is often not even the load-serving entity (LSE) for that geographic area. Consequently, during the hours of the year when system conditions require that some energy be supplied from these local generation units, it is profit-maximizing for the owner to bid whatever the market will bear for any energy these units provide. This incentive exists regardless of the locational pricing scheme used by the wholesale market operator.

This point deserves emphasis: Absent a local market power mitigation

mechanism, the bid of the unit or units with local market power must be taken before lower-priced bids from other firms. The configuration of the transmission network and location of demand makes this unit the only one physically capable of meeting the energy need. Without some form of regulatory intervention, these suppliers will be paid at least their bid price to provide the needed electricity willingly. The configuration of the existing transmission network and the geographic distribution of generation capacity ownership in all US wholesale markets and a number of wholesale markets around the world results in a frequency and magnitude of substantial local market power for certain market participants that if left unmitigated could earn these firms enormous profits and therefore cause substantial harm to consumers.

The uncertain availability of generation units and portions of the transmission network because of forced outages implies that system conditions can arise when virtually any generation unit owner in the wholesale market possesses substantial local market power. Consequently, a local market power mitigation (LMPM) mechanism that provides effective bid mitigation is a necessary component of any wholesale market design. The need for an effective LMPM mechanism is one unique feature of wholesale electricity markets that makes an industry-specific regulator essential at least during the initial stages of the transition from the vertically-integrated utility to the wholesale market regime.

A second rationale for an industry-specific regulator during the transition period is the potential for small market design flaws that cause little harm during most system conditions to lead to substantial consumer harm under certain system conditions. The experience of the California market from when it began operation to the present time illustrates this point. From its start in April 1998 until April 2000, the California market was probably the most competitive wholesale market in the US. Average wholesale prices over this period were less than $35/MWh and the average hourly magnitude of the market inefficiencies as measured by the methodology given in Borenstein, Bushnell and Wolak (2002), hereafter BBW, was less than or very close to equal to those measured by Mansur (2003) for the PJM market (serving some or all of the states Pennsylvania, New Jersey, Maryland, Delaware, and Virginia) and Bushnell and Saravia (2003) for the New England market. This level of market performance occurred in spite of the fact that virtually all of the wholesale energy purchases by the three large California retailers were made through the day-ahead or real-time market.

The amount of hydroelectric energy available from the Pacific Northwest during the summer of 2000 was significantly less than the previous two summers. Consequently, as documented in Wolak (2003a), the five largest fossil fuel electricity suppliers in California now faced significantly less elastic residual demand curves than they did during first two years of the market. As a consequence, these suppliers found it in their unilateral interest to bid less aggressively into the spot market in order to raise wholesale electricity prices in California. As discussed in Wolak (2003b), this strategy was not unilaterally profitable during the first two years of the market because the greater availability of hydroelectric energy from the Pacific Northwest and inexpensive coal-fired energy from the Desert Southwest during that time period caused these suppliers to face significantly more elastic residual demand curves.

This change in competitive conditions during the summer of 2000

enabled in-state suppliers to raise prices substantially through their unilateral actions, as demonstrated in BBW (2002). For example, during the summer months of June to September of 2000, the average difference between the actual price and the competitive benchmark price was more than $70/MWh, which is more than twice the average price of electricity during the first two years of the market. The California experience demonstrates that some market design flaws, in this case insufficient forward contracting by electricity retailers, can be relatively benign under a range of system conditions. However, when system conditions conducive to the exercise of unilateral market power occur, this market design flaw can cause enormous harm to consumers. Consequently, industry-specific regulatory oversight is necessary to intervene as quickly as possible to limit the potential damage when these system conditions arise.

While it is not possible to rule out coordinated actions among the major electricity suppliers in the Western Electricity Coordinating Council (WECC) as a potential explanation for the enormous increase in wholesale prices over the period June 2000 to June 2001, as shown in Wolak (2003a), this is not necessary. The behavior of prices during this time period relative to the first two years of operation of the market can be explained by the unilateral profit-maximizing actions of the major California fossil fuel suppliers given the residual demand curves they faced. Despite extensive multi-year investigations by almost every state-level antitrust and regulatory commission in the western US, the US Department of Justice Antitrust Division, the Federal Energy Regulatory Commission, and numerous Congressional committees, no significant evidence of coordinated actions to raise wholesale electricity prices in the WECC during the period June 2000 to June 2001 has been uncovered.

From the perspective of antitrust law, the most surprising aspect of the period June 2000 to June 2001 in the California market is that despite estimated total market inefficiencies of close to $20 billion, virtually all of which was due to the exercise of unilateral market power, US antitrust law did little to prevent this enormous wealth transfer from occurring. Moreover, following the Enron bankruptcy and disclosure by a number of energy trading firms (in very unflattering terms) that they did attempt to exercise all available unilateral market power in the California market, US antitrust law has been unable to obtain refunds of any of these market power profits. This outcome has occurred because US antitrust law does not prohibit firms from fully exploiting their unilateral market power. Because of the enormous potential harm from the exercise of unilateral market power in electricity, other regulatory safeguards are necessary.

Besides the need to intervene to correct market design flaws after they are determined to be harmful, there is also a need to engage in prospective market monitoring to find market design flaws that lead to substantial harm by less noticeable means. For example, certain aspects of the wholesale market design can increase the likelihood that coordinated actions to raise prices might occur. Aspects of the market design can also enhance the ability of suppliers to exercise their unilateral market power. This logic suggests that another important role for an industry-specific regulator is to monitor the wholesale market to determine which market rules might be enhancing the ability of suppliers to exercise unilateral market power or increasing the likelihood that the attempts of suppliers to coordinate to raise market prices will be successful. Particularly, during the initial stages of the wholesale market regime, this prospective approach to

regulatory oversight should implement market rule changes before they expose consumers to significant harm.

This role for the industry-specific regulator also has a pedagogical component. The transition to a wholesale market regime involves a dramatic change in behavior by a number of market participants. Companies that fail to adapt to the new regime are very likely to go bankrupt and exit the industry, but there are often significant external costs to consumers associated with this outcome. Consequently, an industry-specific regulator can take prospective actions to encourage to adaptation to the new regime and limit the resulting external costs if this change in market participant behavior does not occur. An example of a necessary change in behavior is the need for retailers to hedge spot price risk, something that was unnecessary in the former vertically-integrated regime because electricity retailers typically owned enough generation capacity to meet their load obligations. A significant regulatory challenge in a number of markets around the world is how to provide strong incentives for retailers to engage in the efficient amount of forward contracting, while at the same time not exposing consumers to sustained periods of very high spot prices. A purely market-based solution would be to expose retailers to risk of bankruptcy by having extremely high price caps or bid caps on the spot market. The fear of bankruptcy associated with a sustained period of extremely high spot prices when the retailer has a significant spot market exposure will cause retailers to engage in the appropriate amount of hedging of spot price risk. This strategy has worked remarkably well in Australia, where the bid cap on the spot market was $5,000/MWh, and has recently been raised to $10,000/MWh. Spot prices at or near these levels occasionally occur, but because of the high levels of forward contracting by Australian retailers these prices do not cause significant harm to consumers.

In the US and New Zealand there have been a number of bankruptcies of large retailers as a result of a sustained period of very high spot prices that occurred when the retailer had a significant exposure to the spot market. Over-reliance on the spot market by retailers in the US is not surprising given the decision of the Federal Energy Regulatory Commission (FERC), the US wholesale market regulator, to impose hour-by-hour market power mitigation in the form of relatively low bid or price caps on the spot market and automatic mitigation procedures (AMP) that set very low bid caps for specific generation units. This mitigation limits the volatility and level of spot prices, which dulls the incentive for retailers to sign forward contracts or engage in other hedging arrangements to limit their exposure to the spot market. Consequently, regulatory invention to limit the volatility of spot prices creates the need for regulatory intervention to mandate that retailers purchase a portfolio of financial contracts to hedge their spot price risk. Particularly for the US, where retail market regulation is the domain of the state regulatory commissions and wholesale market regulation the domain of FERC, this strategy by FERC may be necessary to gain the consent of the state regulator to relax its control over retail electricity prices. Forward contracting requirements on electricity retailers are a part of virtually all Latin American markets. The industry-specific regulator typically mandates forward contracting requirements for retailers at various time horizons in advance of delivery to ensure there is adequate generation capacity to serve demand and that no retailers are overly exposed to the spot market.

Regulating forward contracting levels has an additional spot market competitiveness benefit. If these purchases are structured as fixed-price forward contracts for fixed amount of energy in the hour, they have very beneficial impacts on the competitiveness of shorter-term energy markets. As discussed in Wolak (2000a), forward contract obligations by a supplier make it unilateral profit-maximizing to bid more aggressively in the spot market. Moreover, Wolak (2000a) also emphasizes that forward contract obligations by one supplier can make it unilaterally profit-maximizing for other suppliers to bid more aggressively, regardless of their own forward contract holdings. The most successful wholesale electricity markets, as judged by the competitiveness of their spot markets, are those with where only a very small fraction of the total amount of electricity consumed is actually purchased in the spot market.

The final rationale for an industry-specific regulator arises because the technology for supplying electricity implies the need for a single transmission and distribution network for a given geographic area. Even after accounting for the distortions from least-cost supply in setting regulated prices for these services, regulated monopoly provision is generally acknowledged to have a lower overall cost than competitive provision of transmission and distribution services for a given geographic area. Regulated monopoly provision implies the need for a regulator to set the price that wholesale and retail market participants pay for access to the transmission and distribution network. These prices must also allow the network owners to recover the cost of past investments, the ongoing cost of operating the network and the cost of future investments necessary to serve a growing demand for electricity. This task requires the regulator to have significant industry-specific expertise.

There are also important market competitiveness benefits from regulatory oversight of the transmission network. They are determining the terms of conditions for new generation units to interconnect to the transmission network and whether transmission upgrades should take place and where they should take place. As discussed in Wolak (2003c), in the wholesale market regime transmission capacity has an additional role as a facilitator of commerce. Expansion of the transmission network typically increases the number of independent wholesale electricity suppliers that are able to compete to supply electricity at locations in the transmission network served by the upgrade. The industry-specific regulator is best-suited to develop the expertise necessary to determine the transmission network that maximizes the competitiveness of the wholesale electricity market. With the exception of the US, most countries re-structured at a time when they had significant excess transmission capacity, so the issue of how to expand the transmission network to serve the best interests of wholesale market participants has not yet become significant. In the US, determining how to expand the transmission network to serve the needs of wholesale market participants has been a major stumbling block to realizing the expected benefits of electricity industry re-structuring.

5.3 Responsibilities of industry-specific regulator

Goal of regulatory process

This section first states the primary goal of the regulatory process and then sets out the major responsibilities of the industry-specific regulator. Because electricity markets are so susceptible to the exercise of unilateral market power, the primary goal of the industry-specific regulatory process is to limit the ability of market participants to engage in behavior that degrades system reliability and market efficiency, the two major adverse consequences of the exercise of market power.

As noted earlier, it is impossible to prevent firms from exercising all available unilateral market power. This would imply the existence of a perfect regulatory process. The market or system operator would need to know each supplier's minimum cost of producing power. It could then dispatch suppliers based on their minimum cost of producing power. However, if such a regulatory process existed there would be little need to introduce a competitive market because, by assumption, a lower average cost of supplying power to consumers could be achieved by paying suppliers only their minimum cost of production, rather than the market-clearing price set through a process where all suppliers bid to maximize their expected profits for all of the energy they produce. Consequently, any mechanism used to mitigate market power is necessarily imperfect in the sense of being unable to protect consumers from the exercise of all market power.

By the same logic, there are no perfectly competitive markets. However, there are many markets that yield outcomes very close to those predicted by the perfectly competitive ideal. Unfortunately, electricity is not always one of these markets. Consequently, the market designer is typically faced with the choice between an imperfectly competitive market and an imperfect regulatory mechanism to set the compensation paid to a firm. A social welfare maximizing market designer would make this choice based on which mechanism entails the smallest loss in social welfare.

This logic implies a regulatory process that provides incentives for efficient market outcomes, instead of focusing on preventing firms from exercising all unilateral market power. The regulatory process should provide the strongest possible incentives for least-cost provision of electricity to final consumers consistent with the long-term financial viability of the industry. Running a regulatory process is costly and regulatory invention even more so. Therefore, it is important to account for these costs in the design and operation of the regulatory process. Specifically, the regulatory process should first focus on actions that have a very high market efficiency benefits relative to their implementation costs. The regulator should also periodically review the costs and benefits all of aspects of the regulatory process.

Three major responsibilities of regulatory process

I will now describe the three major responsibilities of the industry-specific regulator and how these should be carried out. They are: (1) disseminating information to existing and prospective market participants, (2) ensuring

compliance with all the market rules, and (3) protecting against behaviour that degrades market efficiency and system reliability. Successfully fulfilling each role requires much greater regulatory authority and sophistication on the part of the regulatory process than the previous one.

Smart "sunshine regulation"

A minimal requirement of any industry-specific regulatory process is to provide "intelligent sunshine" regulation. The regulator must have access to all information needed to operate the market and be able to perform analyses of this data and release the results to the public. At the most basic level, the regulator should be able to replicate market-clearing prices and quantities given the bids submitted by market participants, total demand, and other information about system conditions. This is necessary for the regulator to verify that the market is operated in a manner consistent with what is written in the market rules. A second aspect of "smart sunshine regulation" is public data release. Specifically, all data submitted to real-time market and produced by the system operator should be immediately released to the public. Little trading of energy should take place through the real-time market, because it is operated primarily for reliability reasons and all market participants have a common interest in the reliability of the transmission network. Immediate data release best serves these reliability needs.

There should be no limitation on the regulator's access to data either submitted to the system operator by market participants or produced by the system operator. Besides all of the information needed to operate the energy and ancillary services markets and the transmission network, the regulator should also have the ability to request information from market participants on a confidential basis to perform further analyses. Rather than have an ex ante limitation on the type of data it can request, the regulator should have open-ended authority to request information subject to an economic cost-benefit test. To enforce this authority, the regulator should also have the ability to impose financial penalties on market participants that fail to provide the requested data in a reasonable period of time.

Wholesale markets that currently exist around the world differ considerably in terms of amount of data they make publicly available and the lag between the date the data is created and the date it is released to the public. Nevertheless, among the industrialized countries there appears to be a positive correlation between the extent to which data submitted or produced by the system operator is made publicly available and how well the wholesale market operates. For example, the Australian electricity market makes all data on bids and unit-level dispatch publicly available the next day. Australia's National Electricity Market Management Company (NEMMCO) posts this information by market participant name on its website. The Australian electricity market is generally acknowledged to be one of the best performing re-structured electricity markets in the world (Wolak, 1999). On the other hand, the former England and Wales electricity pool kept all of the unit-level bid and production data confidential. Only members of the pool could gain access to this data. It was generally acknowledged to be subject to the exercise of substantial unilateral market power by the larger suppliers (Wolak and Patrick 1997 and Wolak, 1999). The UK government's displeasure with pool prices eventually led to the New Electricity Trading Arrangement (NETA) which began operation on March 27, 2001. Although these facts do not provide definitive proof that

rapid and complete data release enhances market efficiency, the best available information on this issue provides no evidence that withholding this data from the public scrutiny enhances market efficiency.

The public data release should identify the market participant and specific generation unit associated with each bid, generation schedule, or output level. Masking the identity of the market participants, as is done in all US wholesale markets, limits the disciplining value of public data release on market participant behavior. Under a system of masked data release, market participants can always deny that their bids or energy schedules are the ones exhibiting the unusual behavior. The primary value of public data release is putting all market participants at risk for explaining their behavior to the public. In all US markets, the very long lag between the date the data is produced and the date it is released to the public, at least six months, and the fact that the data is released without identifying the specific market participants, virtually eliminates much of the potential benefit of public data release.

Putting market participants at risk for explaining their behavior to the public is different from requiring them to behave in a manner that it is inconsistent with their unilateral profit-maximizing interests. A number of markets have considered implementing "good behavior conditions" on market participants. The most well-known attempt was the United Kingdom's (UK) consideration of a Market Abuse License Condition (MALC) as a pre-condition for participating in its wholesale electricity market. The fundamental problem with these "good behavior" clauses is that they can prohibit behavior that is in the unilateral profit-maximizing interests of a supplier that is also in the interests of consumers. These "good behavior" clauses do not correct the underlying market design flaw or implement a change in the market structure to address the underlying cause of the harm from the unilateral exercise of market power. They simply ask that the firm be a "good citizen" and not maximize profits.

For the case of the UK, the MALC anticipated punishing those market participants that exercised significant amount unilateral market power. However, one difficulty with this approach is that the major beneficiaries of the unilateral exercise of market power are the firms that exercised little if any unilateral market power. One could therefore imagine some firms finding ways to compensate larger firms for exercising their unilateral market power so that all firms can reap benefits. A second difficulty is distinguishing the exercise of significant market power worthy of punishment from expected profit-maximizing behavior. In testimony to the United Kingdom Competition Commission, Wolak (2000b) made these and a number of other arguments against the MALC, which the Commission eventually decided not to implement.

Another potential benefit associated with public data release is that it enables third-parties to undertake analyses of market performance. The US policies on data release severely limit the benefits from this aspect of a public data release policy. Releasing data with the identities of the market participant masked makes it impossible to definitively match data from other sources to specific market participants. For example, some market performance measures require matching data on unit-level heat rates or input fuel prices obtained from other sources to specific generation units. Strictly speaking, this is impossible to do if the unit name or market participant name is not matched with the generation unit.

The long time lag between the date the data is produced and the date it

is released also greatly limits the range of questions that can be addressed with this data. Taking the example of the California electricity crisis, by January 1, 2001, the date that masked data from June of 2000 was first made available to the public, the exercise of unilateral market power in California had already resulted in more than $5 billion in overpayments to suppliers in the California electricity market as measured by BBW (2002). Consequently, a long time lag between the date the data is produced and the date it is released to the public has an enormous potential cost to consumers that should be balanced against the benefits of delaying the data release.

The usual argument against immediate data release is that suppliers could use this information to coordinate their actions to raise market prices through sophisticated tacit collusion schemes. Although the immediate availability of information on bids, schedules and actual unit-level production could allow suppliers to design more complex state-dependent strategies for enforcing collusive market outcomes, it is important to bear in mind that coordinated actions to raise market prices are illegal under US anti-trust law and under the competition law in virtually all countries around the world. The immediate availability of this data means that the public also has access to this information and can undertake studies examining the extent to which market prices difference competitive benchmark levels as described in BBW (2002). Keeping this real-time data confidential prevents this potentially important form of public scrutiny of market performance from occurring.

This data can also be used to undertake third-party studies of whether coordinated actions, explicit or tacit, are occurring. Although these third parties would find it difficult to produce evidence of explicit or tacit collusion that would lead to a conviction in a court of law, they only need to present evidence that has a high likelihood of yielding a conviction in the court of public opinion. The prospect of such adverse publicly is very likely to increase the perceived cost to market participants of engaging in explicit or tacit co-ordinated behavior to raise market prices. Despite a prohibition against conscious parallelism in under US antitrust law, I am not aware of a successfully prosecuted lawsuit against such behavior in any industry in the US. Consequently, the fear of a conviction in the court of public opinion appears to be best way to prevent such behaviour.

Economic theory provides no clear predictions about the relationship between the information made available to market participants and their ability to implement less competitive market outcomes. A number of theoretical papers have identified circumstances when asymmetric information between market participants can be a very effective device for implementing less competitive market outcomes, through either unilateral or co-ordinated actions. I am also not aware of any systematic empirical evidence industries demonstrating that making more information available to market participants leads to less competitive market outcomes.

Coherent arguments in favor of masking the identity of market participants in the publicly released bid, schedule and production data are more difficult to find. Assuming that the concerns with public data release enhancing the ability of market participants to coordinate actions had been addressed, it is difficult to determine what market efficiency-enhancing benefit results from masking the identity market participants. Masking the identity of the market participant only limits the "sunshine regulation" value of public data release.

An important aspect of the public data release question is the

distinction between data that the regulator can request and receive from market participants and data that must be released to the public. There is a natural boundary between these two types of data. Any data that the system operator must request from market participants or must produce in order to operate the real-time market should be released to the public. As noted earlier, the real-time market is operated primarily for reliability reasons and this data release policy is consistent with the goal of preventing harm to system reliability and market efficiency.

Public release of any information on forward market positions or transactions prices, where the large volumes of energy are typically traded, does not serve this goal. This is information about a market participant that is unnecessary to operate the real-time market, although it does impact the bidding, scheduling or production behavior of that market participant, as discussed in Wolak (2000a). Knowledge of these financial positions is not needed by the system operator to run the spot market or the transmission network.

Because of the fundamentally financial nature of forward market transactions sold by electricity suppliers, it is very difficult to get accurate information on the true forward market positions of electricity suppliers. They can re-trade forward market obligations among themselves to yield forward market positions far above or below their expected production of electricity. A number of studies of electricity trading in the US before by Enron bankruptcy in late 2001 estimated that each electron ultimately delivered through a US wholesale electricity market was bought and sold in forward markets more than five times. For this reason, even if the regulator attempted to collect this forward market data from suppliers on a regular basis it would not be very useful. For example, if the regulator specified a minimum quantity of forward contract sales for each supplier it regulated, these suppliers could undertake forward contract transactions with affiliates not subject to regulatory oversight to meet these minimums. Moreover, those affiliates not subject to oversight by the regulator could then re-construct their holding company's desired forward contract position. Consequently, routinely collecting the forward contract positions of suppliers could cause them to render this information of little or no use to the regulator through affiliate transactions.

There is a strong argument for keeping any forward contract positions the regulator might collect confidential. As noted in Wolak (2000a), the financial forward contract holdings of a supplier are major determinants of the aggressiveness of its bids into the spot market. Only if a supplier is confident that it will produce more than its forward contract obligations will it have an incentive to bid or schedule its units to raise the market price. Suppliers recognize this incentive created by forward contracts when they bid against competitors with forward contract holdings. Consequently, public disclosure of the forward contract holdings of market participants can convey useful information about the incentives of individual suppliers to raise market prices, with no countervailing reliability or market-efficiency enhancing benefits.

A final aspect of the data collection portion of the regulatory process is concerned with scheduled outage coordination and forced outage declarations. A major lesson from wholesale electricity markets around the world is the impossibility of determining whether a unit that is declared out-of-service can actually operate. Different from the former vertically integrated regime, declaring a "sick day" for a generation unit – saying that

it is unable to operate when in reality it could safely operate – can be a very profitable way for a supplier to withhold capacity from the market in order to raise the wholesale price. To limit the ability of suppliers to use their planned and unplanned outage declarations in this manner, the market operator and regulator must specify clear rules for determining a unit's planned outage schedule and for determining when a unit is forced out.

Before the start of each year, suppliers should submit to the system operator a schedule of planned outages for each of their units. The system operator would then compile the planned outage schedules submitted by all suppliers and verify that they do not compromise system reliability. If they do, then the system operator will suggest modifications to achieve a schedule of planned outages for all units consistent with reliable network operation on annual basis. Although the system operator should attempt to accommodate the wishes of each supplier, it must have the ultimate authority to set the final schedule for all planned outages. Once this planned outage schedule is set, it should be released to the public. Modifications of these unit-level planned outages schedules during the year are subject to the approval of the system operator. These modifications should be released to the public once they are approved.

A similar process should be followed for scheduling planned transmission line outages. The system operator should coordinate the planned transmission outage process with all of the transmission owners and the generation unit owners. It should also make the final decision on when both generation units and transmission lines can be taken out for maintenance.

To limit the incentive for "sick day" unplanned generation outages, the system operator should specify the following scheme for outage reporting. Unless a unit is declared available to operate up to its full capacity, the unit is declared fully out or partially out depending on the amount capacity from the unit bid into the market at any price at or below the current price cap. This definition of a forced outage eliminates the problem of determining whether a unit that does not bid into the market is actually able to operate. Such a unit should be assumed to be forced out, because the owner is not offering this capacity to the market. The system operator should therefore only count capacity from a unit bid in at a price at or below the price cap as available capacity. Information on unit-level forced outages according to this definition should be publicly disclosed each day on the system operator's web-site.

This disclosure process cannot prevent a supplier from declaring a "sick day" to raise the price it receives for other energy or ancillary services that it sells. However, the process can make it more costly for the market participant to engage in this behaviour by registering all hours when capacity from a unit is not bid into the market as forced outage hours. For example, if a 100 MW generation unit is neither bid nor scheduled in the spot market during an hour, then it is deemed to be forced out for that hour. If this unit only bids 40 MW of the 100 MW at or below the price or bid cap during an hour, then the remaining 60 MW is deemed to be forced out for than hour. The regulator can then periodically report forced outage rates based on this methodology and compare these outage rates to historical figures from these units before re-structuring or from comparable units from different wholesale markets.

If the wholesale market makes capacity payments to generation units, then the amount of capacity a unit owner is allowed to sell should be based

on the capacity of the unit multiplied by the 12-month rolling average of availability factor of the unit computed based on these outage rates. For example, if the 100 MW unit only bids into the market during half of the hours of the year, then it should only be allowed to sell 50 MW of capacity, because this is the average amount of capacity the unit provides to the market.

A similar process should be followed for unplanned transmission line outages. The regulator should compile information on the hourly amount of available transmission capacity. As soon as outages or de-ratings occur, this information should be made publicly available. The regulator should also compile the annual distribution of hourly transmission capacity availability and make this information publicly available. This information can also be used by the regulator and system operator to implement penalty and sanctions schemes for transmission owners that fail to maintain their transmission facilities in a manner consistent with good utility practice.

Ensuring compliance with market rules

Many market outcomes that are harmful to system reliability and market efficiency could be prevented if all market participations fulfilled all of their contractual obligations. Because the actions of each market participant impact system reliability and market efficiency, this implies that all parties have a common interest in honoring their contractual obligations. If the cost of violating a contractual commitment or market rule is less than the unilateral benefit from this action, the market participant will find it profitable to violate, which also adversely impacts system reliability and market efficiency. This logic implies that the second responsibility of the regulatory process is to: (1) design market rules to resemble publicly verifiable contractual obligations, and (2) determine the appropriate penalties and sanctions to deter violations of these rules without adversely impacting market efficiency or system reliability.

Not all market rules are defined to resemble publicly verifiable contractual obligations. Prohibitions against market manipulation or the abuse of market power are prime examples. These prohibitions have done very little to prevent harmful market outcomes. The California market rules contained these prohibitions, but they did little to prevent the events of June 2000 to June 2001. The third responsibility of the regulatory process is to prevent harmful market outcomes that cannot be prevented by market participants obeying market rules that resemble publicly verifiable contractual obligations.

A large fraction of harmful market outcomes can be prevented and the costs of operating the market and the costs of participating in the market will be lower if all market participants are confident that all contractual commitments will be honored regardless of system conditions. Contract enforcement costs stem from ambiguous or overly broad market rules or market rules that are not, or cannot be, enforced. A transparent rule that can be rigorously enforced is superior to an overly broad rule that is difficult to enforce. Irregular enforcement, either because of imprecise rules or inconsistent effort, increases the cost of market participation. This can also lead to increased market rule violations as more market participants push the boundaries of acceptable behavior.

This logic implies that regulator should divide market rules into two categories: (1) those that resemble publicly verifiable contractual

obligations with little subjective judgement to determine compliance, and (2) those that require a formal administrative process to determine compliance. Rules in first category should be written to limit ambiguity and simplify enforcement. Those in the second category should have pre-specified administrative processes that deter behavior harmful to system reliability and market efficiency because of the large amount of judgement associated with determining a violation has occurred.

The first type of market rule should be written so that a violation resembles the process of issuing a speeding ticket as closely as possible. If the regulator measures the speed of the car using a publicly verifiable measuring device and finds that the car's speed exceeds the posted limit, then the regulator should assess a pre-specified penalty. The penalties and sanctions process should not involve a finding of intent in order for the regulator to assess a penalty. An example of a market rule violation covered by this mechanism is a failure to comply with terms implied by a bid into the wholesale market. One example is a market participant submitting a bid to supply a fixed quantity of energy within a given response time and then failing to meet this commitment. Suppose the supplier bids to provide 50 MWh of energy in 10 minutes from the time the bid is accepted. If the supplier fails to provide any of the purchased energy when it is called upon, the unit owner should be penalized for failing to meet this contractual commitment.

Both types of market rules require penalty and sanction mechanisms, but for slightly different purposes. In both cases, penalties and sanctions are imposed to deter market rule violations. For the market rules where determining compliance is straightforward, the penalties and sanctions are the primary mechanism for deterring violations. For case where subjective judgement is required to determine a violation, the penalties and sanctions are the ultimate backstop, but the administrative process is the primary mechanism for preventing harmful market outcomes.

Protecting against behaviour harmful to market efficiency and system reliability

The final responsibility for the regulator is to deter behavior that is harmful to system reliability and market efficiency that occurs despite public disclosure of data and market participant behavior and penalties for publicly-observed, objective market rule violations. This is the most complex aspect of the regulatory process to implement, but it also has the potential to yield the greatest benefit. It involves a number of inter-related tasks. In a bid-based market, the regulator must design and implement a local market power mitigation mechanism. The regulator must also determine when a market rule detracts from system reliability and market efficiency and suggest and implement the necessary changes in this market rule. The regulator must determine when market outcomes cause enough harm to some market participants to merit explicit regulatory intervention. Finally, if the market outcomes become too harmful, the regulator must have the ability to temporarily suspend market operations. All of these tasks require a substantial amount of subjective judgement on the part of regulator.

Local market power mitigation (LMPM) mechanism. In all bid-based electricity markets a local market power mitigation mechanism is necessary

to limit the bids a supplier submits when it faces is insufficient competition to serve a local energy need. An LMPM mechanism is a pre-specified administrative procedure (usually written into the market rules) that determines: (1) when a supplier has local market power worthy of mitigation, (2) what the mitigated supplier will be paid, and (3) how the amount the supplier is paid will impact the payments received by other market participants. It is increasingly clear to regulators around the world, particularly those that operate markets using Locational Marginal Pricing (LMP), that formal regulatory mechanisms are necessary to deal with the problem of insufficient competition to serve certain local energy needs.

Formulate and implement efficiency-enhancing market rule changes. The regulator must determine which market rules detract from market efficiency or system reliability and formulate and implement the appropriate market rule changes. Because the level and geographic distribution of demand, the mix of input fuels used and ownership shares for generation capacity in the control area, and the configuration of the transmission network can all change over time, market rules must also change. The regulator must continually analyze and assess the market efficiency impacts of all market rules. Once it has identified a deficient market rule, the regulator must then work with the system and market operators to devise the necessary remedy. This duty underscores the need for the regulator to analyze market performance using the data it has compiled.

Penalize behavior harmful to system reliability and market efficiency. The regulator is the first line of defense against harmful market outcomes. Persistent behavior by a market participant that is harmful to market efficiency or system reliability should be subject to penalties and sanctions. In order to assess these penalties, the regulator must determine first intent on the part of the market participant. The market rules should contain a general provision prohibiting persistent behavior detrimental to system reliability and market efficiency. The goal of this provision is to establish a process for the regulator to intervene to prevent a market meltdown. As shown in Wolak (2003a), there are instances when actions very profitable to one or a small number of market participants can be extremely harmful to system reliability and market efficiency. A well-defined process must exist for the regulator to intervene to protect market participants and correct the market design flaw facilitating this harm. This provision protects against the harmful exercise of unilateral market power, which is distinct from the exercise of unilateral market power, which is equivalent to maximizing profits. The discussion in Section 5 states that besides a finding of "intent", the regulator must also find "significant harm" for violation of this market rule to occur.

Determine when market activities can be temporarily suspended. The regulator must have the ability to suspend market operations on a temporary basis when system conditions warrant it. The suspension of market operations should only occur after a pre-specified administrative procedure has been followed and it has been determined that it is the only option available to the regulator to prevent significant harm to market efficiency and system reliability. As has been demonstrated in various countries around the world, electricity markets can sometimes become

wildly dysfunctional and impose enormous harm over a very short period time. For example, during the early stages of the New England market, there were short-lived, but severe distortions in the Installed Capacity and Operating Capacity markets that eventually lead to a suspension of market activities. During the California market's first summer, one of the reserve capacity markets experienced extremely high prices for a short period. During the initial stages of the wholesale market in the state of Queensland in the Australia, unilateral market power problems became so severe that it was necessary to suspend market operations under sufficient interconnection capacity with neighboring states and generation capacity within the state could be built. Under these sorts of circumstances, the regulator should have the ability to suspend market operations temporarily until the problem can be dealt with through a longer-term regulatory intervention or market rule change.

5.4 Guidelines for setting penalties and sanctions

This section presents three guidelines for designing penalty and sanctions mechanisms for ensuring market participant compliance with market rules.

The first guideline is that the penalty for a market rule violation should be sufficiently high to make it unilaterally unprofitable for a market participant to violate the rule. The strategy used by FERC until very recently of limiting the magnitude of the penalty to ordering the firm that violated the market rule to return the profits gained from their violation will not deter violations. Under this scheme, firms would have little to lose from violating rules because their violation may not be detected and, even if it is detected, they are not made any worse off than if they had followed the rules in the first place.

The second guideline is that the mechanisms for imposing penalties and sanctions should be set in advance and the relationship between a specific market rule violation and the amount of the penalty assessed should be as transparent as possible. Returning to the above example of failing to comply with a dispatch instruction, the system operator could require that the supplier either find a like replacement for the power the unit is unable to provide or require the owner to make the payments necessary to hold harmless all market participants for its failure to meet its contractual obligations. Making the relationship between a specific market rule violation and the penalties assessed as transparent as possible achieves two goals. First, it limits the opportunities for the system operator and regulator to exercise discretion in setting penalties. Second, it allows market participants to formulate the best possible cost-benefit assessment associated with a specific market rule violation.

The third guideline is that the penalty associated with a market rule violation should not exceed the harm this market rule violation causes to all market participants. This guideline addresses the tendency regulators often have to set penalties sufficiently high to deter market participants from engaging in behavior that has any likelihood of violating the market rules. Excessive penalty levels have a cost. They cause market participants to focus on avoiding being penalized for a market rule violation rather than

on producing electricity in a least-cost manner or purchasing wholesale electricity in a least-cost manner. For example, setting the penalty for failing to respond to a dispatch instruction too high could cause suppliers to avoid participating in the wholesale market or to downgrade the maximum amount of energy they are willing to sell from each of their units.

5.5 Behavior detrimental to system reliability and market efficiency

This section describes a general mechanism for determining if a supplier engages in persistent behavior detrimental to system reliability and market efficiency and what the appropriate standards are for determining when market operations should be suspended. Real-time electricity markets are, by definition, centralized market mechanisms where the actions of some market participants can impact the ability of other market participants to sell their output or buy the energy necessary to serve their retail customers. For many of the same reasons that there can only be one air traffic controller for each airport, there can only be a single real-time market operator for each transmission network. The need to deliver power through a common transmission and distribution network suggests that all market participants have a common interest in preventing behaviour that significantly degrades system reliability and market efficiency because it reduces their expected profits from participating in the wholesale market.

This aspect of the regulatory process addresses the concerns about harmful market outcomes typically voiced by parties claiming market manipulation. However, it avoids what I believe to be the impossible task of demonstrating that a market participant manipulated the market. As noted earlier, whether actions constitute market manipulation depends on one's perspective. Viewed from one perspective, all suppliers that attempt to impact the price they are paid through their own unilateral actions are engaging in market manipulation. The extent of unilateral market power possessed by a supplier is typically measured by its ability to move market prices through its unilateral actions. Consequently, a blanket prohibition of market manipulation written into the market rules seems to prohibit suppliers from maximizing profits given the actions of their competitors. These actions can lead to market outcomes that benefit consumers when all suppliers face sufficient competition. This logic is why there is no explicit prohibition against market manipulation under US antitrust law-it amounts to prohibiting behavior that is a major driver of the benefits in competitive markets.

The prohibition of behavior that is detrimental to system reliability and market efficiency focuses on identifying and eliminating detrimental behavior by market participants, rather than on punishing this behavior. Penalties and sanctions are a last resort, when all other options for eliminating the behavior have been tried; including asking the market participant to stop because of the significant harm this behavior is imposing on other market participants.

There is a potential downside to giving the regulator the ability to make such a finding. To the extent that the regulator is influenced by the political environment, it may be tempted to intervene to pursue political ends rather

than allow politically favored electricity retailers to pay higher prices for electricity or politically favored suppliers to receive lower prices for the electricity they produce. That is why the regulator must follow a well-defined process before it is allowed to make a finding of persistent behavior harmful to system reliability and market efficiency and to suspend market operations temporarily.

The major difficulty associated with implementing this market rule is that the regulator would have to infer from a market participant's behavior whether its bidding, scheduling, or operating behavior *intended* to harm system reliability or market efficiency. If the regulator identifies behavior that is detrimental to system reliability, and has clear evidence (for example, a whistleblower or internal correspondence) that the market participant engaged in this behavior with full knowledge that it significantly harmed system reliability or market efficiency, penalties may be imposed without first going through the administrative process described below.

However, it seems very unlikely that the regulator would have direct evidence of intent, particularly if there is a market rule that imposes significant penalties on the market participants that have been shown to have engaged in this type of behavior. Enforcing a "behavior detrimental to system reliability and market efficiency" provision is more difficult if this market rule also imposed the very reasonable requirement that this detrimental behavior must also have a significant impact on market outcomes. This would require the regulator to make the often very subjective determination of what constitutes a "significant" market impact. Despite these difficulties with determining "intent" and "significant market impacts," an administrative procedure along the lines discussed below can adequately address these complications in making the finding of "intent to impose significant harm."

A necessary first step in any process for determining intent is the ability to demand and receive information from market participants. This reinforces the need for a pre-condition for participation in the wholesale market that each entity agree to provide, in a timely manner, all information necessary for the regulator to undertake an investigation of intent to impose significant harm to system reliability or market efficiency. As discussed above, this agreement to provide information should be subject to the constraints that the information request is necessary to undertake the current investigation and does not impose costs on the market participant that are out of line with the alleged harm that the market participant is imposing.

The regulator should implement the following multi-stage process for determining intent and imposing penalties commensurate with the harm caused by these actions. It is counterproductive for the regulator to prohibit actions that are difficult to define and even more difficult to determine if they occur. For the same reasons that I believe it is impossible to prove market manipulation, it equally difficult prove that a market participant is gaming the market or engaging in false scheduling, meaning that it is buying energy that it does not intent to consume or selling energy it does not intent to produce. Market rule prohibiting behavior harmful to system reliability and market efficiency should prevent the harm that many observers believe is the result of market manipulation, gaming or false scheduling.

Prohibiting behaviour without first finding "intent" and "significant harm" will cause market participants to avoid actions that often enhance

market efficiency and system reliability that might be interpreted as one of those prohibited. Instead, the regulatory process for determining intent should recognize that it is extremely difficult to distinguish legitimate profit-maximizing behavior from actions that intend to harm system reliability and market efficiency without some exchange of information between market participants and regulator. In addition, behavior that might be interpreted by some observers as gaming or market manipulation is often rendered unprofitable by the actions of other market participants. Consequently, these sorts of market efficiency or system reliability problems can often be solved through information provision to the market at large, thereby eliminating the need for further action.

A key feature of this market rule is a transparent process for identifying intentional behavior detrimental to system reliability or market efficiency. This should include a process for taking the actions necessary to stop this behavior or the harm that it causes. The focus of this process should be on stopping as quickly as possible intentional behavior that the regulator determines causes significant harm to market efficiency and system reliability.

The first step in this process is therefore to identify behavior that is likely to harm to market efficiency and system reliability. Two findings are necessary for the process to continue to the next step. The regulator must first determine if this behavior is persistent, and if it has the potential to impose significant harm either because it is very persistent or extremely harmful when it does occur. The next stage of the process involves alerting all market participants to the existence of this behavior and publicly disclosing the identity of the market participant engaging it. The goals of this stage of the process are to subject this market participant to public scrutiny and to provide all market participants with information that they can use to take actions that attempt to render this behavior unprofitable.

Public disclosure is a very important step in the process of determining intent because all market participants, including the market participant engaging in the behavior, know that the regulator has publicly stated that this behavior is harmful to system reliability or market efficiency. Consequently, continued behavior by this market participant that imposes significant harm provides strong evidence in favor of a finding of intent.

In most cases, this stage of the process should put an end to the behavior or the harm it causes. However, in those instances when the actions are sufficiently profitable to the market participant or group of market participants that they continue to cause significant harm, the regulator should initiate a formal investigation of intent. To do this the regulator needs the ability to request and receive in a timely manner the information from the offending market participant necessary to make a credible determination of intent to impose harm. An important goal of this information gathering effort is for the market participant to provide information to the regulator demonstrating that there is no direct causal link between market participant's behavior and harm to system reliability or market efficiency.

If the regulator's information gathering efforts reveal substantial evidence of a direct causal link between this market participant's behavior and the presumed harm, then the regulator should find that this market participant did intend to harm system reliability or market efficiency. If there is an affirmative finding of intent, the regulator may need to collect additional information to determine the appropriate magnitude of

penalties. If the regulator makes an affirmative finding of intent it would then be required to set the appropriate level of penalties. The results of the investigation and the regulator's rationale for its recommended level of penalties should be subject to judicial review.

As should be clear from the above discussion, the major focus of this process is on eliminating the harmful behavior as soon as possible, not on assigning blame or imposing penalties. Only when public disclosure of the actions and the regulator's own investigation fails to stop or eliminate the harm associated with this behavior should the regulator attempt to determine intent and assign penalties for this behavior.

To guard against the possibility that there may be circumstances when the unilateral profit-maximizing actions of market participants can lead to enormous harm to consumers, the regulator should have the ability to suspend market operations temporarily. An example of such a mechanism is the "guardrails-to-competition" approach discussed in Wolak (2003d). This mechanism relies on the competitive benchmark analysis discussed in BBW (2002). It sets a prospective measure of the extent to which electricity prices over a rolling 12-month horizon can exceed the competitive benchmark level. If the difference between the quantity-weighted average market price over the previous 12 months exceeds the quantity-weighted average competitive benchmark price computed using the methodology outlined in BBW (2002) over the previous 12 months by more than $5/MWh, then an automatic regulatory intervention would be triggered. In Wolak (2003d), I argue that this intervention should be a 12-month period of cost-of-service prices for all suppliers. The idea behind this intervention is that it is viewed as sufficiently Draconian, yet credible to implement, so that suppliers never allow this guardrail on competition to be violated. For example, rather than exercise substantial market power in the spot market, a supplier will sign forward contracts or other spot price hedging arrangement to improve spot market performance before these guardrails are exceeded.

5.6 Coordinating antitrust and regulatory policy

This section concludes with a brief discussion of how the industry-specific regulatory process should interact with the antitrust authority. The primary concern of the regulatory process is protecting against the economic harm associated with unilateral exercise of market power. Antitrust policy is concerned with detecting coordinated actions to raise prices and combinations (typically mergers) that result in a substantial lessening of competition. Antitrust law also prohibits attempts to monopolize, but this is unlikely to be relevant to the electricity industry beyond its implications for merger analysis. As discussed in Section 5.2, from the perspective of other industries, the concentration levels in the wholesale and retailing segments of the electricity industries in most countries appears is fairly competitive. The transmission and distribution segments are price-regulated by the industry-specific regulator, so it is difficult to see how a firm could monopolize this industry except through mergers, or because of a poorly designing industry-specific regulatory process. Given these roles, the industry-specific regulator should primarily serve a support function to

the antitrust process, with the antitrust policy functioning independent of the industry-specific regulatory process.

This hierarchical relationship implies that the industry-specific regulator is the first line of defence for consumers against harmful market outcomes. While the industry-specific regulator may wish to approve mergers, this seems redundant if the antitrust authority does a thorough review. Given the expertise of the industry-specific regulator, a thorough review would require that the antitrust authority to solicit extensive input from the industry-specific regulator, including the provision of industry-specific data that is part of the ongoing regulatory process. The industry-specific regulatory would also be a key source of information for making a determination of "coordinated behaviour in restraint of trade" by industry participants.

It is possible that actions by the regulatory process could have the unintended consequence of assisting market participants in coordinating their actions to raise prices or in facilitating the attempts of a firm to monopolize a market. For example, bringing market participants together to discuss their costs of production and how they operate their generation facilities, as FERC often did during the former vertically-integrated regime, can improve the efficacy of the regulatory process. However, in a wholesale market regime, these sorts of meetings have the potential to substantially harm competition. This potential for the regulatory process to facilitate violations of antitrust law implies that the antitrust authority must clearly specify protocols governing multi-lateral meetings between market participants and the regulatory authority. However, two outstanding antitrust issues are whether the industry-specific regulator can be held liable for inadvertently facilitating antitrust violations and whether market participants should be held harmless for antitrust violations that are the direct result of this behavior by the antitrust authority. Particularly, in the US, this issue has become increasingly relevant as the industry-specific regulator has more actively intervened in the day-to-day operation of the wholesale market.

Despite this potential for the industry-specific regulator to facilitate antitrust harm, I am extremely sceptical that the antitrust authority can eventually replace the industry-specific regulator, because antitrust law is not concerned with the unilateral exercise of market power that causes significant consumer harm, which has historically been a major problem in wholesale electricity markets. However, as market participants become more sophisticated, the need for a number of regulatory safeguards is likely to diminish.

Once the price cap on the spot market is increased to a level where it is only occasionally hit and a liquid forward market for energy develops, there is less of a need to monitor the level of forward contracting and mandate minimal levels of coverage for electricity retailers. Although it is unclear at what level of the price cap this change in regulatory oversight should occur, the necessary level is clearly above the highest price cap that currently exists in all of the US markets of $1000/MWh. To the extent that local market power mitigation mechanisms are incorporated into the market rules in an easily enforceable manner, there is less need for the industry-specific regulator to be involved in this process.

It is difficult to see how the need for industry-specific regulator to price transmission and distribution services can be eliminated. Although an administrative process for allowing new generation units to interconnect to

the transmission network could be written into the market rules, determining when and where competition-enhancing transmission upgrades should take place and whether they have expected benefits to market participants that exceed their cost is a sufficiently complex and subjective process to require regulatory oversight far into future. I would expect that as policymakers obtain a better understanding of the competition-enhancing benefits of transmission upgrades, the industry-specific regulator will be given more authority to manage the transmission network. Because so many markets outside of the US began operation with significant over-capacity in their transmission network, this has not been major issue. However, a forward-looking pro-active transmission expansion policy can limit the frequency that LMPM mechanisms that distort market outcomes must be relied upon. The industry-specific regulator has the expertise and legal mandate to manage the transmission network to maximize the economic benefits from a competitive wholesale market.

The other circumstances under which I would expect a significantly reduced role for the industry-specific regulator is if there was a substantial reduction in both the system-wide and local concentration in generation ownership. For example, divestiture of generation capacity down to the unit level (there are typically multiple generation units at a power plant) would significantly reduce virtually all of the unilateral market power concerns expressed above. Given the size of the financial returns that existing suppliers earn from the current concentration and geographic distribution of generation unit ownership and the likely economies to scope and scale associated with operating multiple generation units at a given location, it is very unlikely that such an outcome could be implemented in any existing wholesale market. Consequently, industry-specific regulators need not fear for their jobs, because there is much for them to do in the future.

References

Borenstein, Severin, Bushnell, James and Wolak, Frank A. (2002) *Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market,* American Economic Review, December, 1367-1405.(*)

Bushnell, James and Saravia, Celeste (2002) *An Empirical Assessment of the Competitiveness of the New England Electricity Market,* May, Center for the Study of Energy Markets Working Paper Number CSEMWP -101, available from http://www.ucei.berkeley.edu/pubs-csemwp.html

Mansur, Erin T. (2003) *Vertical Integration in Restructured Electricity Markets: Measuring Market Efficiency and Firm Conduct,* October, Center for the Study of Energy Markets Working Paper Number CSEMWP-117, available from http://www.ucei.berkeley.edu/pubs-csemwp.html.

Wolak, Frank A. and Patrick, Robert H. (1997) *The Impact of Market Rules and Market Structure on the Price Determination Process in the England and Wales Electricity Market,* February.(*)

Wolak, Frank A. (1999) *Market Design and Price Behavior in Restructured Electricity Markets: An International Comparison,* in Competition Policy in the Asia Pacific Region, EASE Volume 8, Takatoshi Ito and Anne Krueger (editors) University of Chicago Press, 79-134.(*)

Wolak, Frank A. (2000a) *An Empirical Analysis of the Impact of Hedge Contracts on Bidding Behavior in a Competitive Electricity Market,* International Economic Journal, Summer, 1-40.(*)

Wolak, Frank A. (2000b) *Comments on the Office of Gas and Electricity Markets (Ofgem) License Condition Prohibiting Abuse of Substantial Market Power,* submission to United Kingdom Competition Commission, July.(*)

Wolak, Frank A. (2003a) *Measuring Unilateral Market Power in Wholesale Electricity Markets: The California Market 1998 to 2000,* American Economic Review, May 2003, 425-430.(*)

Wolak, Frank A. (2003b) *Diagnosing the California Electricity Crisis,* The Electricity Journal, August, 11-37.(*)

Wolak, Frank A. (2003c) *The Benefits of an Electron Superhighway,* Stanford Institute for Economic Policy Research Policy Brief. November.(*)

Wolak, Frank A. (2003d) *Sorry, Mr. Falk: It's Too Late to Implement Your Recommendations Now: Regulating Wholesale Markets in the Aftermath of the California Crisis,* The Electricity Journal, August, 50-55.

(All references marked with (*) are available at http://www.stanford.edu/~wolak)

Other books in the same series

2003: The Pros and Cons of Low Prices
The book is about predatory pricing; an issue that has intrigued and bewildered the competition policy community for a long time and where conflicting views are held. The problem and the challenge for competition policy are to draw the fine line between pro-competitive pricing behaviour on the one hand and predatory pricing as an instrument of abuse on the other.

The purpose of this book is to assess predatory practices from a competition policy perspective and the implications of recent theoretical and empirical developments for a consistent treatment of such practices in competition policy. We have solicited contributions from experts in the field, covering the main streams of development and discussing policy issues related to predation in the light of these developments.

2002: The Pros and Cons of Merger Control

The book is intended to serve as a contribution to the debate on merger control and consists of four individual contributions from independent scholars and professionals with an expertise in economics. Naturally, the opinions expressed are those of the authors alone.

The pros and cons of merger control are high on the agenda of policy makers, competition authorities, academics, representatives of industry and labour organizations, and others. The need for merger control is widely supported - but the specific principles and tools by which it should be exercised are subject to discussion and debate, and also revision. The review of the Merger Regulation in the Green Paper by the European Commission has raised several fundamental questions.

The pros and cons of changing the "substantive test" from the dominance standard to the SLC-test ("Substantial Lessening of Competitoin") is an issue that needs careful scrutiny. The concept of collective dominance and other issues such as jurisdiction, efficiencies, and procedures are also of great importance.

2000: Fighting Cartels – Why and How?

The book takes up legal as well as economic aspects on why we should be concerned with cartels, how to detect and find sufficient evidence of cartel behaviour, and how to accomplish an effective prevention of cartel behaviour.

The theme reflects the growing concern in Sweden as well as internationally for the detrimental effects of cartels on society. The book takes up various aspects of anti-cartel enforcement, and in particular, how competition authorities of today could be successful in the prevention of cartels.

The books can be ordered from our website www.konkurrensverket.se