Bender, Christian M. et al.

Article — Published Version

Effective competition in telecommunications, rail and energy markets

Intereconomics


This Version is available at:
http://hdl.handle.net/10419/74560

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
“Effective Competition” in Telecommunications, Rail and Energy Markets

The markets for network-based products and services pose particular problems with regard to competition. The transition from monopolistic to competitive structures and the issues of infrastructure sharing, dominant players and network externalities have all been subject to intensive debates. Despite liberalisation, deregulation and privatisation – and the quasi-automatic solutions they seemed to present – a number of problems persist. Furthermore, the regulation which was introduced as a remedy for distortions of competition during the period of transition from monopoly to competition has not ended with market liberalisation. Each new generation of network technologies creates new challenges to be addressed by regulators and competition authorities, and each industry presents specific problems. The following articles discuss a number of the persisting and new questions concerning competition in selected network-based industries.

Christian M. Bender, Georg Götz and Benjamin Pakula

Effective Competition: Its Importance and Relevance for Network Industries

Establishing effective competition is a core objective of European regulatory policy for network industries. The intention to establish effective competition is written down and holds a prominent position in telecoms, railway and energy legislation. For example, Viviane Reding, former European Commissioner for Information Society and Media, stated with respect to the telecommunications sector that “effective competition is the key for current and future success.” It is all the more surprising that effective competition lacks a clear definition. This raises two problems: First, it is not possible to judge whether the policy pursued is successful, i.e. there is no benchmark for effective competition. Second, whether a network industry should (still) be regulated ex-ante or be subject to competition law depends on whether competition in the respective market is effective. Without a clear definition, a decision on sunset legislation is hard to make. This paper discusses the meaning and importance of the concept of effective competition for network industries, highlighting the dynamic aspects of competition and the importance of potential competition.


Economic Concepts and Legal Importance of Effective Competition

Effective Competition in the Legal Framework

The New Regulatory Framework for the telecoms sector states that effective competition can best be promoted through an efficient level of investment in infrastructure. Furthermore, a lack of effective competition is defined as the ability of an operator to sustain prices at an excessively high level. In the German Telecommunications Act (TKG), effective competition is defined as the absence of significant market power.

Introducing competition into the railway market is a key objective of the EU 1996 White Paper. German railway legislation states that competition should be effective

2 See COM 2009/140/EC, recital 54.
3 See ibid: Article 2, amendment 9.
4 See TKG 2004 (31). Note that the European Commission in the beginning of the telecoms liberalisation considered an operator with a market share greater than 25 per cent as exerting significant market power (see Interconnection Directive, 97/33/EC).
and undistorted. Effective competition is interpreted as having several railway companies in the market that actually compete, whereas undistorted competition is defined as non-discriminatory access to infrastructure and absence of cross-subsidisation within the incumbent company.

In the energy sector, effective competition is equally desired. It should be reached through non-discriminatory, transparent and fairly priced network access. For example, transit fees for long-distance gas pipelines are exempt from ex-ante regulation if these conditions are met.

Effective Competition in Economic Theory

In economic theory, there is no single concept that defines effective competition. It is therefore helpful to consider some ideas from oligopoly theory. With respect to competition in network industries, two aspects appear especially important: market power and potential competition.

The concept of workable competition introduced by Clark explicitly takes market power into account and constitutes the basis for what today is termed effective competition in economic theory and law. Clark found market imperfections like product heterogeneity, intransparency, time-lags etc. to be indispensable for economic progress. These market imperfections are necessary to make competition workable. There are similarities here with the “Austrian perspective”. Following this approach, the major aim of regulation is not to eliminate all excess profits, but to give competing companies the incentive to discover more efficient forms of production and to find out what customers want. This approach enables a better accounting for dynamic developments in the industry like product and process innovation than a static one could do.

Potential competition is also relevant in network industries. If a monopolist that produces a homogeneous good in a market without entry cost faces potential competition, the outcome in the market will be first or – if we consider fixed costs – second best because competitors may replace him immediately. This means a high market share does not necessarily indicate market power. Although there is criticism concerning the robustness of the assumptions, e.g. goods are rarely homogeneous and entry is generally not costless, this concept is valuable because it highlights that potential competition may serve as a disciplining device for dominant firms. This may apply to intramodal as well as intermodal competitors. For example, a rail provider might trigger entry by bus transport operators if his prices are sufficiently high.

The German Monopolies Commission, the Government’s academic advisory group for antitrust and regulation is-

6 See Allgemeines Eisenbahngesetz (AEG) §1, paragraph 1. In original terms “wirksam und unverfälscht”.
9 See Gasnetzentgeltverordnung (GasNEV) For a more thorough discussion of this aspect see F.C. Haus: Effective competition and the essential facilities doctrine, in this issue, pp. 31-35.
12 See S. Littlechild: The nature of competition and the regulatory process, in this issue, pp. 10-17.

Contributors to this Forum

Christian M. Bender, Justus-Liebig-University Giessen, Germany.

Georg Götz, Justus-Liebig-University Giessen, Germany.

Benjamin Pakula, Justus-Liebig-University Giessen, Germany.

Stephen Littlechild, Emeritus Professor, University of Birmingham, and Fellow, Judge Business School, University of Cambridge, UK.

Günter Knieps, Albert-Ludwig-Universität, Freiburg, Germany.

Paul W.J. de Bijl, CPB Netherlands Bureau for Economic Policy Analysis, Den Haag, Netherlands; TILEC, Tilburg University and ENCORE.

Anton Schwarz, Austrian Regulatory Authority for Broadcasting and Telecommunications (RTR), Economics Division, Vienna, Austria.

Florian C. Haus, Freshfields Bruckhaus Deringer, Cologne, Germany.
sues, states with respect to the interplay of sector-specific regulation and general competition law that competition is workable (“funktionsfähiger Wettbewerb”) if it is structurally established and persists even when regulation is reduced. Workable competition does not necessarily exclude having a dominant player in the respective market. Kahn argues that even a small number of competitors with a small combined market share might impose high competitive pressure, especially when their investment costs are sunk.15 Cable providers with a small market share in the broadband internet access market, for instance, might exert sufficient competitive constraints on a telecommunications incumbent if their network is already in place and consequently their marginal costs are low.16

Effective competition does not imply absence of market power. On the contrary, “the prospect of having some market power (i.e. some profit) represents a most powerful incentive for firms to innovate and invest”.17 In competition policy, it is important to mind that “[d]efending competition is not tantamount to defending competitors. [...] Protecting inefficient firms [...] would be detrimental from a welfare perspective.”18 Thus, politicians and regulators should take into account that a variety of different market settings are in line with what economic theory would call a competitive market. Furthermore, they need to consider that regulation should protect competition rather than competitors or business models based on regulation.19

**Definition of the Relevant Market in Network Industries**

As mentioned above, an evaluation of the intensity of competition in an industry requires taking account of all the forces that exert competitive pressure on the companies in that industry. This is the task of market definition. In the following, we discuss whether and how the standard instrument for market definition in antitrust cases, the SSNIP test,20 can be applied to delineate markets in regulated industries. We argue for an economic market definition, which takes the substitutability of services and products into account, rather than one based on technological considerations. For example, if we consider the broadband access market from a technological perspective, the local telephone company has significant market power as it owns 100% of the local loops.21 From an economic perspective, the situation may change if the substitution potential of alternative access technologies, e.g. cable or mobile networks, is examined.

The above discussion touches on the relation between market definition and the existence of an essential facility. Many network industries are subdivided into the network infrastructure itself (upstream market) and the markets for services (downstream markets) that need the infrastructure as an essential input. There is a long discussion about the conditions under which an input constitutes an essential facility and whether viable alternative ways to enter the downstream market exist.22 Market definition is the key to answering these questions. An input might be a bottleneck for an industry, but the relevant market might reach beyond the boundaries of the industry. Accordingly, the decision whether and how to regulate may change drastically: imagine an industry with fierce intermodal competition like the shipping industry where river boat, freight railway, and road transport compete. Considering only the railway industry qualifies tracks as an essential facility. In the broader context of the relevant market, this assessment may be different.

Market definition also has a geographic dimension. Infrastructure facilities might be an essential input in some regions with no intermodal competition whereas they are not in others. Therefore geographically differentiated markets should be considered.

Finally the time dimension enters. Technology changes over time, especially in dynamic markets, such as telecommunications.23 Based on this insight, we should focus on persistent essential facilities and distinguish them from temporary bottlenecks.

---

16 Ibid, p.162.
18 Ibid.
21 See the recent decision M 3/09 of the Austrian telecom regulator on ULL regulation.
22 See United States vs. Terminal Railroad Association of St. Louis, 224 U.S.338 (1921) and 236 U.S.194 (1914); R. Sherman: Market Regulation, Boston 2006, Pearson, p. 354. In Europe, today’s essential facility doctrine is based on the 1998 Bronner case of the European Court of Justice, cf. European Court of Justice (1998). Recent cases for the relevance of this doctrine for the network industries are provided by F. C. Haus, op.cit.
Employing the SSNIP Test in Regulated Industries

The SSNIP test evaluates whether a small but significant and non-transitory increase in prices of all products offered by the firms in a candidate market yields higher profits. If the price increase is profitable, the included products (and geographic areas) are considered as one market, otherwise one has to include the closest substitutes among the remaining products and to perform the test again.

It is of particular importance for regulated industries that the SSNIP test takes the competitive price level as its starting-point, which is not necessarily the same as the actual price level. Considering only the actual price level could induce consumers to substitute the product with products of inferior quality or with products that would not be in the same market at competitive prices. The actual price level may be higher than the competitive one which would lead to a too broadly defined market. The case that seems more relevant for regulated industries is that the price level – due to regulation – is low with respect to the competitive one. Therefore there would be a tendency to define the market too narrowly and conclude that the companies possess market power.

Applying an SSNIP test to network industries is the exception rather than the rule. Markets defined by regulatory authorities generally do not exceed the boundaries of the industry in the technological sense. Nonetheless, the SSNIP test can also be of assistance in network industries.

Measurement of Effective Competition

Competition Policy

In competition policy, a few instruments are generally used to measure competition in a market. The Lerner index is among the most important ones. However, in the context of network industries this index does not seem appropriate because fixed costs are not considered. With respect to regulated industries the notions of “dominance” and “market power” are of greater importance. Dominance is generally measured in terms of market shares whereas the appraisal of whether a company exerts market power requires a more detailed analysis. The concept of dominance goes back to the Structure-Conduct-Performance (SCP) approach developed in the 1930s by Chamberlin and Mason. The difficulty with a definition of dominance based on high market shares is that the latter might be an economic signal that a company is more efficient or more innovative than its competitors. Boone et al. argue in a similar way and state that a sector becomes more competitive if the profit distribution becomes more unequal. This leads to higher market shares for efficient firms at the expense of inefficient firms. This argument is all the more relevant in network industries where network effects or economies of scale lead to high market shares for one or two companies, but this does not necessarily mean that the companies have strong market power.

Nonetheless, concentration ratios or the Herfindahl index witness the relevance of the SCP approach in today’s competition policy. They may serve as an indicator for a deeper analysis of the industry, but are not sufficient to evaluate if an industry is competitive or not for they often fall short of a sufficient assessment of industry-specific and case-specific characteristics. Moreover, as argued above, potential competition and/or strong price competition may impose a strong competitive constraint on a dominant firm.

Hausman and Sidak have shown for the Irish mobile phone market that despite high concentration ratios of the two largest mobile operators the market outcome was competitive. Thus, a structural analysis might be the first step but further analyses are required.

Consequently, we should distinguish between dominance in terms of high market shares and (significant) market power. The term “significant market power” is defined by the European Commission as the ability “to behave to an appreciable extent independently of competitors, customers, and ultimately consumers.” This can be translated as the ability to significantly raise prices above competitive levels. Determination of these competitive levels is a particular problem in network industries exhibiting network
Effects, economies of scale and scope, and a large amount of sunk (entry) costs. The appropriate benchmark appears to be an industry that is served by few firms.

Assessment of Competition in the Telecommunications Industry: The Three Criteria Test

The telecommunications industry is a network industry where sunset legislation is intended. The procedure to evaluate whether a market is sufficiently competitive and may be released from ex ante regulation to competition law comprises two steps: First, a list of different markets, based on a recommendation from the European Commission, is analysed by the National Regulatory Authorities (NRAs) using the “three criteria test” (TCT). The NRAs examine whether (i) there exist “high and non-transitory barriers to entry”, whether (ii) the market “structure does not tend towards effective competition in a relevant time horizon”, and whether (iii) the application of “competition law alone would not adequately address the market failure(s) concerned”. These three criteria are used cumulatively so that “[a]ny market which satisfies the three criteria in the absence of ex ante regulation is susceptible to ex ante regulation”. Second, if a market passes the TCT, the market is analysed as to whether any firm has significant market power. If this is not the case, the market is said to be effectively competitive and may be deregulated.

The TCT has been criticised for several reasons. For example, the Commission defines structural barriers to entry as cost or demand structures which yield “asymmetric conditions between incumbents and entrants preventing market entry for the latter”36, and includes economies of scale and scope within this definition. From an economic perspective, asymmetries between firms based on economies of scale and/or scope are neither a necessary nor a sufficient condition for barriers to entry and the focus should be put on the question whether there are sunk costs related to market entry.37 The European Commission argues that, in the presence of barriers to entry, competitive constraints may exist that are based on “a limited — but sufficient — number of undertakings having diverging cost structures and facing price-elastic market demand”.38 The number of firms required to create sufficient competitive constraints is thus a key question which cannot be answered exactly.39 Moreover, the Commission’s position regarding diverging cost structures between competitors is problematic. The existence of several firms in one market with different cost structures seems to be an indicator of weak rather than effective competition as firms with higher costs would have to exit the market in a competitive environment.40 The third criterion emphasises not structural but legislative problems and verifies the commensurability of ex ante regulation and opens the discussion on whether competition law will ever be appropriate for solving market failures in network industries.41

The necessity of performing the test of significant market power (SMP) in addition to the TCT is heavily disputed: Briglauer argues that a market analysis which evaluates effective competition would be in opposition to a positive evaluation of the second and third criteria.42 Moeschel adds that the two-step approach, the TCT followed by an SMP test, turns the analysis upside-down as the analysis within the TCT seems like a rough estimation whereas the detailed examination is relocated to the SMP test.43

To summarise: the two step approach starting with the TCT followed by an SMP test on the national level seems like an objective instrument to evaluate competition in markets at first sight but there are some major criticisms and problems in practice. Nevertheless, a carefully and accurately performed TCT with some specifications might be a useful economic approach to analysing whether there is effective competition in markets.

Assessment of Competition in the Rail Industry

In the rail industry, regulation is limited to the essential facilities. Competition is assumed to be possible in the downstream market, i.e. in the provision of rail services. A rough definition of the rail market generally encompasses long-haul passenger traffic, local passenger traffic and freight transport. A more detailed analysis is certainly required in many cases.

Competition is often assessed by considering market shares only. With respect to the German market, the German regulator, the Bundesnetzagentur44, and the Monopole

Effects, economies of scale and scope, and a large amount of sunk (entry) costs. The appropriate benchmark appears to be an industry that is served by few firms.

Assessment of Competition in the Telecommunications Industry: The Three Criteria Test

The telecommunications industry is a network industry where sunset legislation is intended. The procedure to evaluate whether a market is sufficiently competitive and may be released from ex ante regulation to competition law comprises two steps: First, a list of different markets, based on a recommendation from the European Commission, is analysed by the National Regulatory Authorities (NRAs) using the “three criteria test” (TCT). The NRAs examine whether (i) there exist “high and non-transitory barriers to entry”, whether (ii) the market “structure does not tend towards effective competition in a relevant time horizon”, and whether (iii) the application of “competition law alone would not adequately address the market failure(s) concerned”. These three criteria are used cumulatively so that “[a]ny market which satisfies the three criteria in the absence of ex ante regulation is susceptible to ex ante regulation”. Second, if a market passes the TCT, the market is analysed as to whether any firm has significant market power. If this is not the case, the market is said to be effectively competitive and may be deregulated.

The TCT has been criticised for several reasons. For example, the Commission defines structural barriers to entry as cost or demand structures which yield “asymmetric conditions between incumbents and entrants preventing market entry for the latter”36, and includes economies of scale and scope within this definition. From an economic perspective, asymmetries between firms based on economies of scale and/or scope are neither a necessary nor a sufficient condition for barriers to entry and the focus should be put on the question whether there are sunk costs related to market entry.37 The European Commission argues that, in the presence of barriers to entry, competitive constraints may exist that are based on “a limited — but sufficient — number of undertakings having diverging cost structures and facing price-elastic market demand”.38 The number of firms required to create sufficient competitive constraints is thus a key question which cannot be answered exactly.39 Moreover, the Commission’s position regarding diverging cost structures between competitors is problematic. The existence of several firms in one market with different cost structures seems to be an indicator of weak rather than effective competition as firms with higher costs would have to exit the market in a competitive environment.40 The third criterion emphasises not structural but legislative problems and verifies the commensurability of ex ante regulation and opens the discussion on whether competition law will ever be appropriate for solving market failures in network industries.41

The necessity of performing the test of significant market power (SMP) in addition to the TCT is heavily disputed: Briglauer argues that a market analysis which evaluates effective competition would be in opposition to a positive evaluation of the second and third criteria.42 Moeschel adds that the two-step approach, the TCT followed by an SMP test, turns the analysis upside-down as the analysis within the TCT seems like a rough estimation whereas the detailed examination is relocated to the SMP test.43

To summarise: the two step approach starting with the TCT followed by an SMP test on the national level seems like an objective instrument to evaluate competition in markets at first sight but there are some major criticisms and problems in practice. Nevertheless, a carefully and accurately performed TCT with some specifications might be a useful economic approach to analysing whether there is effective competition in markets.

Assessment of Competition in the Rail Industry

In the rail industry, regulation is limited to the essential facilities. Competition is assumed to be possible in the downstream market, i.e. in the provision of rail services. A rough definition of the rail market generally encompasses long-haul passenger traffic, local passenger traffic and freight transport. A more detailed analysis is certainly required in many cases.

Competition is often assessed by considering market shares only. With respect to the German market, the German regulator, the Bundesnetzagentur44, and the Monopole

34 COM 2009/140/EC, recital 5.
39 P.W.J. de Bijl: The need for a communications regulator: a lesson from the Netherlands, in this issue, pp. 21-26, discusses the question whether two firms in one industry can ensure sufficient competition.
44 See Bundesnetzagentur: Jahresbericht 2009.
olkommission\textsuperscript{45} consider competition to be developing in the freight and local passenger traffic with market shares of the competitors at around 20\%, whereas there is hardly any active competition in long-haul passenger traffic. In the latter case, intermodal competition is also considered but estimated not to be very intense. As explained above, market shares serve as an indicator but not as a proof of market power, so we should handle this analysis with care. Moreover, competition is only measured for the rail industry and not for the different markets, in which intermodal competitors may also constitute competitive constraints.

Ivaldi and Vibes explicitly take intermodal competition in the intercity passenger market into account and simulate different regulatory scenarios. They conclude that evaluating the effectiveness of competition in a market requires accounting for all potential travellers, all modes and all firms. Moreover, they find that a small number of competitors is enough to create a high degree of competition.\textsuperscript{46} Friebel and Niffka analyse in a case study how the entry of low cost airlines in Germany affected the traffic volume and pricing strategies of Lufthansa and Deutsche Bahn. They find that the entry put heavy pressure on both companies, which leads to the conclusion that intermodal competition has more bite than usually considered. Moreover, they argue that it was misleading to look at the regulation of railroad markets in an isolated way.\textsuperscript{47} WIK Consult analyses whether the German incumbent, Deutsche Bahn, holds a dominant position in four markets: long-haul business traffic, local traffic, bulk cargo transport and single wagon freight transport. WIK cannot substantiate the dominant position of Deutsche Bahn in these markets and estimates that intermodal competition (indirectly) contributes to restrict the market power of the network company.\textsuperscript{48}

These studies show that looking only at (the different kinds of) rail transport is too a narrow perspective for making an assessment of competition in the respective markets.

**Assessment of Effective Competition in Network Industries: A Conceptual Framework**

In this paper we have discussed different meanings of the term “effective competition” as used in the network indus-

\textsuperscript{45} See Monopolkommission: Wettbewerb erfordert Weichenstellung, Sondergutachten 55, 2010, Nomos.


**Figure 1**

The Conceptual Framework

1. **Definition of the relevant market:** Are there demand-side substitutes?
   - Yes: add products
   - No

2. **Market power:** Do the firms have significant market power?
   - Yes
   - No

3. **Essential facility:** Does market power arise from an essential facility?
   - Yes
   - No

4. **Persistence of the essential facility:** Are there supply-side substitutes if prices change?
   - Yes
   - No

5. **Safeguard price cap for the essential facility**

6. **Regulation of the essential facility**

7. **Deregulation/Competition law**

1. **The definition of the relevant market.** The first step in this analysis is the definition of the relevant market. As discussed above, the SSNIP test can be a suitable tool for this purpose. Similarly to competition policy, we focus on demand-side substitutes in this first step. The SSNIP test has to be performed carefully because the elasticity of substitution may be distorted due to the fact that prices in regulated markets do not necessarily reflect the level of competition. All products that belong to the relevant product and geographic market according to the SSNIP test have to be included in the subsequent analysis.

2. **The assessment of market power.** In order to assess the market power of the different companies in the defined market, we consider market shares as a first indicator. If the market shares are unequal, e.g. if the market share of one firm is excessively high, an SMP test should be performed. Moreover, the calculation of market shares should be based on available capacities rather than on actual quantities. The more economic approach applied in com-
petition policy provides a range of instruments to identify market power using econometric methods. If we find that no firm exhibits significant market power, the market should be assigned to competition law.

3. The identification of the essential facility. If significant market power is confirmed for one firm, we have to determine whether market power is based on an essential facility. Here, essential facilities should always be defined from an economic perspective and not from a technological one. If market power does not stem from an essential facility, the market should be assigned to competition law. Otherwise, we have to verify if the essential facility is persistent. The question is whether the facility is still considered essential even when the regulatory regime is changed. For example, given a higher access fee the replication of the essential facility might be economically reasonable or firms may offer supply-side substitutes that are not competitive at a lower access price level. If, and only if, the essential facility is persistent should it be regulated to guarantee competition in the downstream market. If not, the market should be considered for deregulation.

4. The phasing out of regulation. According to the previous analysis, there are several cases for which regulation should be phased out. In these cases, competition policy should be applied. Additionally, a regulatory safeguard may be implemented, especially if essential facilities are judged non-persistent as discussed above. For example, a safeguard price cap may be imposed slightly below the level at which prices might be considered excessive under competition law.

To summarise, given this conceptual framework and a thoroughly conducted application of the suggested methods and steps, we are confident that the suggested approach provides an unbiased framework for evaluating competition and facilitating sunset legislation. The assessment of competition should focus not only on prices but should also consider dynamic aspects. Even though competition in network industries is inherently imperfect, policymakers should not disregard the function of competition as a Schumpeterian discovery process. This may require taking the chance of deregulating to see if competition is effective. The risks appear limited, as competition policy instruments are always available as a last resort. Without taking this risk, one might never find out whether competition would be effective without regulation.

Stephen Littlechild*

The Nature of Competition and the Regulatory Process

If regulation is supposed to replicate (some would say “mimic”) or at least reflect the results of competition, then it is necessary to ask what competition is supposed to do. In order to design effective regulation we need to understand the meaning of effective competition.

There are different ways of assessing competition, and hence effective competition. One approach focuses on equilibrium, another on market process. These have different implications for policy – for example, whether to allow or prohibit competition, or whether to impose, maintain or remove a price cap.

This will be illustrated by discussion of nationalisation and privatisation, and the setting and removing of transitional retail price caps in telecommunications and electricity.

The paper suggests a means of calibrating the concept of a “safeguard” price cap.

Seeing competition in terms of market process suggests an approach to utility regulation focused on facilitating the market process rather than replacing it. Some examples are given from the energy sector worldwide. The paper concludes with an application to airport regulation, where the potential for regulation to enable effective competition presents particular opportunities today.

The Meaning of Competition

Neo-classical economics applies the tools of welfare economics to a benchmark based on perfect competition. This assumes many buyers and sellers. It is a static approach, taking cost and demand curves as given. In its simplest manifestation, it focuses on equilibrium where price equals marginal cost which equals average cost, hence there is zero profit.

* This paper is based on a keynote speech at the Workshop on Effective Competition, Justus-Liebig-University, Giessen, 27 May 2010. I am grateful to Georg Goetz for comments on an earlier draft.
Another approach as reflected in the writings of Adam Smith and the Austrians Schumpeter and Hayek stresses the concept of rivalry regardless of the number of competitors. This is a dynamic approach, based on creativity and innovation, and on the search for profit opportunities via the discovery of shifts in cost and demand curves. The economy is characterised by profits and losses as the market tends to equilibrium (without in practice ever reaching it). In Schumpeter’s words, competition is a “perennial gale of creative destruction”.

Consider three main properties of effective competition: A: eliminating excess profits, B: discovering more efficient methods of production, C: discovering what customers want.

Both neo-classical and Austrian approaches acknowledge property A (as reflected in so-called allocative efficiency), but Austrian economics has placed greater weight than neo-classical economics on properties B (productive or X-efficiency) and C (which does not seem to have a name in neo-classical economics though Michael Beesley once coined the term “Y-efficiency”).

Property A has attracted most interest in discussions of competition policy, for example underlying the SSNIP (Small but Significant Non-transitory Increase in Price) test. But properties B and C are arguably more important over the longer term.

**Example: nationalisation without competition versus privatisation with competition**

During the 1960s and 1970s economists used static welfare economics to analyse UK nationalised industries. They posed the question: how should the nationalised industries set prices? Answer: the industries should follow optimal pricing and investment rules. These take cost and demand curves as given. There is no role for competition.

However, the more fundamental problems were dynamic. These industries were characterised by inefficiency, excessive costs, uneconomic investment, old products, too little regard for customers’ preferences and too little innovation. It was necessary to discover better ways of doing things – to change cost and demand curves, not to take them as given.

Privatisation and competition sought to do this. Private ownership provided better incentives to find more efficient production methods, and to discover and deliver the products and services that customers preferred. Competition provided opportunities for others to challenge the incumbents, and for rivalry. This could be expected to lead to lower costs and prices, to new products and to innovation.

### Incentive Regulation with RPI-X Price Cap

The concept of incentive regulation using an RPI-X price cap reinforced this approach. It enhanced the incentive to efficiency and innovation. It did not assume that the regulator would specify the outcomes. Rather, it was for companies to discover these opportunities.

Appraised against the three properties of competition noted above, price cap regulation has had many successes.

- **Property A**: eliminating excess profits has become a central regulatory focus. The building block approach is built on estimates of efficient operating and capital cost and a stringent weighted average cost of capital (WACC). This has led to tough price caps and typically lower prices in real terms.

- **Property B**: privatisation, competition and regulation have worked well in terms of discovering more efficient production techniques. There has been significantly increased efficiency – in round terms a greater output than before is now produced with about one third of the previous workforce.

- **Property C**: there has been greater investment to provide better quality of service, which it is assumed that customers prefer.

But price cap regulation as practised in the UK also has its limitations.

- **Property A**: tight price caps can discourage effective competition even where such competition is feasible. This limits the scope for deregulation.

- **Property B**: how can regulators discover the efficient levels of operating and capital cost that they plug into their calculations of X? In practice, price control processes have become rather burdensome.\(^1\)

- **Property C**: how is the regulator to discover customer preferences to inform the decisions about tradeoffs between quality and price? Costs and preferences can vary from one area to another, but regulatory centralisation limits the ability to tailor price controls to the particular circumstances of each area. Moreover, the uniform-

---

\(^1\) Regulatory papers by the Office of Electricity Regulation (Offer)/Office of Gas and Electricity Markets (Ofgem) increased about eight-fold over the first three distribution price control reviews. See S.C. Littlechild: Regulation, over-regulation and deregulation, CRI Occasional Lecture, 22, given at the Royal Society, London, on 24 November 2008, University of Bath.
ity of approach across companies reduces the scope for innovation and learning.

**Transitional Price Caps**

Regulators often set a transitional price cap where competition is not yet effective. The effectiveness of competition is often judged by market shares or the SSNIP test. Regulators tend to set the transitional price cap equal to their estimate of efficient cost, as they do for networks, and then wait for competition to arrive before removing the cap. However, this assumes that the growth of competition is independent of the price caps. This is not the case.

Price caps set in this way underestimate the costs and prices that characterise an actual competitive process. Setting them equal to cost removes the element of monopoly profit that generally characterises actual competitive markets. It removes or reduces the price disparities due to different parties having different efficiencies, and different information and expectations. It reflects a projected greater efficiency in future, which actual market prices do not. It may reflect a low assumed cost of capital appropriate to the conventionally regulated monopoly networks, rather than the higher risk obtaining in competitive markets, which may be exacerbated by the cost of regulatory risk.

Potential competitors are less interested in entering a market if the regulator’s policy is to reduce the incumbent’s prices to the level that could be offered by the entrant. Customers are less interested in switching if the regulator sees it as its task to ensure that the prices charged by all companies reflect what the best competitors can offer. For these reasons, setting transitional price caps may in practice make new entry more difficult and deter the development of competition.

**Example: retail telecommunications**

In 1983 an RPI-X retail price cap was proposed for British Telecommunications (BT), “to hold the fort until competition arrives”. Over the following years, Oftel repeatedly found that BT’s profitability was higher than would be expected in a competitive market; and concluded that competition was not yet fully effective. It retained and repeatedly tightened the retail price cap. It claimed that “Retail Price Caps have brought about a steady reduction in prices to the point that the UK has some of the lowest prices for residential telephony among developed countries.”

This was no doubt true. But was that policy conducive to competition and choice? Not until 2006 did Ofcom end the retail price cap – after 22 years! Is it not possible that retaining and repeatedly tightening the price cap discouraged the development of competition and prolonged the time when the price cap could be abandoned?

**Example: retail electricity competition**

In 1998, as Director General of Electricity Supply, I introduced transitional retail price restraints when opening the residential electricity retail market to competition. I said that “the restraints should not seek to do the job of competition, or discourage its development. … The aim is to … leave scope for competitors to purchase and operate more efficiently than the incumbent PESs [Public Electricity Suppliers]. It is then for the competitive process to bring these further benefits to customers”.4

In 2002 my successor at Ofgem removed the price cap. This was a courageous move given the political opposition to doing so. However, Great Britain subsequently led the world in retail electricity competition. As of 2006, competitors had managed to take 52% of the UK residential market. Markets in other countries with no or very light retail price control had competitor market shares in the range 47% to 28%. Markets with moderate price controls had competitor market shares in the range 30% to 16%. Finally, markets with heavy price controls or other barriers had competitor market shares at 8% or below, often little more than 1%.5

**Competition Is More than Prices**

Decisions to impose and maintain price caps often implicitly assume that price is more important to customers than other aspects of competition, such as choice and innovation. This is not necessarily the case, as a former chairman of the Competition Commission (CC) pointed out: “Lower prices are by no means sufficient if the process of rivalry is weakened.” With such weakening, “several dimensions of rivalry will often still be diminished, including the choices available to consumers concerning the number of independent sources of new ideas,”

---


new strategies, innovative products or processes and the like." Competition is not only about price: "Competition is, to an important extent, a mechanism by which new ideas emerge and the best ones survive, only to be superseded by other still better ones."

This was well illustrated in Germany: “When the Berlin Wall came down, West Germans were not amazed at how high prices were in the East; they were amazed at the extraordinary lack of choice and poor quality of the products which were available, suggesting that this had been the real, enduring benefit of a competitive market economy.”

Such innovation can also be found in competitive retail electricity markets. Whereas the previous monopolies used to set a single variable tariff, competitive suppliers now offer a choice of contracts. Examples include price guarantee tariffs (1-3 yrs) in the UK, chosen by 4.6 million customers a few years ago. Fixed prices up to 10 years have been offered. In Norway spot price contracts have been chosen by about 25% of customers. In Sweden one supplier has offered contracts with fixed prices in winter and spot prices in summer. Market contracts have been chosen by over 50% of customers in Sweden and by 69% of customers in South Australia.

Example: price caps at UK airports

The UK Civil Aviation Authority (CAA) recommended de-designation of Stansted Airport. (A “designated” airport is subject to economic regulation including price control, so “de-designating” an airport is tantamount to removing its price control.) The Secretary of State for Transport rejected this, saying “On balance the evidence suggests that it is more likely than not that Stansted airport alone will acquire significant market power in the future, although this conclusion is finely balanced.”

It seems hard enough to judge whether a market is competitive now. It is even harder to assess the likely course of competition in future. It would seem easier to assess the strength of competition by removing the price control to see whether any market power does exist and is exerted. If necessary a price control or some other restriction could be reimposed.

Elsewhere, medium-term and long-term contracts between airports and airlines have helped to facilitate the coordination of investment and reduce market risks. Price controls prevent or discourage the development of such contracts. They thereby distort the competitive market process.

A Safeguard Price Cap

Unable to remove the price control at Stansted, the CAA considered other alternatives. It noted the problems of the conventional building block approach. This introduced a risk of distorting airport investment, which the CAA considered was greater than the risk of market power at Stansted. It therefore envisaged a “safeguard” price cap, set just below the level at which prices might be excessive under general competition law rather than set equal to cost.

The Competition Commission rejected this approach. It held that the risk of investment distortion was less than the risk of market power at Stansted. It considered that a building block approach (using a WACC of 7.1%) was more certain.

The issues impinge on the debate between ex ante and ex post regulation. Removing a price control is essentially a move from ex ante to ex post regulation. Ex ante regulation provides certainty: prices set below this level are OK, prices above this level are NOT OK. In contrast, ex post regulation provides certainty: prices set below this level are OK, prices above this level are NOT OK. In contrast, ex post regulation means uncertainty: the price a company sets MAY BE OK – or it may NOT be. Ex post regulation may therefore be more risky for customers and companies – and for the regulator, who could be accused of opening the door to possibly excessive prices.

Could a safeguard price cap reduce these risks? Is it possible to combine ex ante and ex post regulation by indicating zones of OK, MAYBE and NOT OK? Figure 1 illustrates in terms of rates of return.

Calibrating a Safeguard Price Cap

We shall attempt to calibrate the diagram by asking: What is an acceptable competitive return? Regulators would say:

---

8 Secretary of State for Transport: Decision on the regulatory status of Stansted airport, Department for Transport, 15 January 2008.
a return about equal to the cost of capital. They regularly agonise about WACC, and tend to settle on numbers of about 7%, or at least in the range 6 – 8% (all figures in this section pretax real).

Competition authorities might accept higher returns. The same former chairman of the Competition Commission argued that “profits are the key signal and incentive for the proper functioning of a market economy … There is no per se reason why profits in excess of the cost of capital represent anything other than the effective working of a competitive market.”9 Consistent with this, in 2000 the Competition Commission found that grocery companies had a WACC of about 10% and found them not guilty when they were earning a return of about 14%: about double the return that regulators find reasonable for network monopolies.

During the earlier period 1973–1998, the Office of Fair Trading (OFT) tended to make referrals for companies earning returns of about 20%, in other words about three times the typical regulatory WACC. In judging these cases the Monopoly and Mergers Commission (MMC) seems to have found companies guilty when they were earning returns of about 35%, in other words about 5 times regulatory WACC.10

This suggests that transitional price caps set on the same basis as monopoly network price caps – that is, about 7% return on capital – may be too severe. Regulatory safeguard caps might be set on the basis of a return in the range 10 - 20%. Whether regulators would find that persuasive remains to be seen.

**Network Regulation Where Competition Is Not Effective: Alternative Approaches**

The standard approach to network regulation focuses on competition property A: how to set the appropriate price cap so as to prevent excessive profit? But effective competition also poses questions about properties B and C: How to discover efficient production and investment? How to discover what customers want?

This is an increasing dilemma for Ofgem. Its RPI-X@20 review considered the strengths and limitations of its approach to date. It documented great success over the last 20 years, but concluded that the approach will not be appropriate for future conditions. In effect, it posed a question that is familiar from the Austrian perspective on competition: how to set price controls when future needs are unknown? Ofgem concluded that greater incentives on companies were required in order to discover these future needs. But will greater incentives alone be sufficient? Would it not be helpful to involve the customers whose future needs are in question?

Some alternative and newer approaches to regulation better replicate effective competition as seen from the Austrian perspective. They do so by greater involvement of companies and users/customers in decision-making. The regulator facilitates the market discovery process, instead

---

of replacing it. We illustrate with some examples from Argentina, the USA, Canada, the UK, Australia, Germany and the EU.

**Example: the public contest method in Argentina**

When Argentina privatised its electricity sector in 1992, the government sought a method of regulation that did not put an undue strain on the abilities and independence of companies and the regulator. For the existing transmission grid it applied a conventional RPI-X price cap. But new investment proposals had to be made, voted for and paid for by users. Then they were put out to tender to determine the minimum cost of provision. Initially there were a few problems but generally the approach worked well. Users worked together to decide on future investments and the future investment schedule.

**Example: US energy regulation at FERC**

The US Federal Energy Regulatory Commission (FERC) has long encouraged parties to settle (initially, in order to cope with a great backlog of cases). During 1994-2000 there were 41 gas pipeline cases, of which 34 settled in full, 5 in part, and only 2 were litigated. The main gain from settlements was that the different process led to innovative rate freezes, which the regulator could not legally impose. These were more certain, and had better efficiency incentives than the litigated approach. In practice, FERC staff play a significant role in facilitating the settlements.

**Example: the consumer advocate in Florida**

The Florida Public Service Commission (PSC) is the state regulatory body, but in practice the Public Counsel (the consumer advocate body) has negotiated settlements with utilities. In the electricity sector it has negotiated over three-quarters of the total rate reductions, worth $4 billion. It argues that customers have preferred rate reductions to building up company reserves. For their part, the utilities got greater accounting flexibility, plus revenue-sharing efficiency price freezes instead of rate of return controls.

---

Example: UK

In the UK it has been possible to remove price control from the smaller airports, and those where competition is effective. This leaves the London airports still subject to control. The Civil Aviation Authority (CAA) had concerns about the previous price control process, which had proved antagonistic. It proposed a process of constructive engagement, whereby it asked airlines and airports to try to agree certain inputs into the price control review: traffic forecasts, quality of performance standards and future investment programmes. The CAA retained responsibility for assessing future operating costs, the cost of capital, financing and the final price control. These inputs were largely agreed at Heathrow and Gatwick airports. There was also some improvement in relationships and understanding in the industry.

Example: Australia

In 2000 Australia privatised its major airports and imposed five year price caps. In 2002 the Government removed the price caps, encouraged contractual agreements between airports and airlines, extended the concept of monitoring and threatened to reintroduce price control if necessary. A Productivity Commission Review in 2006 found that there were still some airline concerns e.g. as to the definition of service quality, and acceptance of appropriate terms and conditions by airports. However, investment was better, prices were not excessive, more information was being exchanged, and industry relationships were better. Other assessments too are favourable. In 2007 the Government decided to continue the policy, and addressed some weaknesses of the original framework, for example by clarifying the valuation of initial assets. It agreed to clarify the threat of re-regulation, though the successor Government abandoned this plan. There is an ongoing debate whether binding dispute resolution would undermine this approach, as the Productivity Commission feared, or would be a useful supplement to it, consistent with the present paper.

Example: Germany

In Germany it is widely held that competition between airports would not be effective. There is traditional cost-of-service (cost-plus) regulation of airport landing charges, implemented by the federal states. This provides little incentive to efficient operating and capital costs. Most airports are content with this because they are largely owned and regulated by the federal states. Airlines argue that this approach is not transparent, certain or effective. Niemeier and other airport economists have argued for independent regulation to implement incentive price controls, along the lines of RPI-X price caps in the UK.

In parallel, however, airlines have brought several civil law cases since 2000. They have argued that the landing charges approved by the regulatory bodies are not equitable, transparent and cost-related. In certain respects the courts eventually found against the airports. To avoid such disputes, Hamburg, Frankfurt, Hannover and Düsseldorf Airports entered into so-called framework agreements with their airlines, which provided stability of pricing for fixed periods of time (often 4 years). There was some sharing of the benefits and risks of changes in traffic volumes. The agreements often provided for quality monitoring, consultation and cooperation, which the traditional regulatory approach did not. They also provided for some flexibility of response – for example, at Hamburg the parties agreed to a suspension of the agreement after the events of 9/11. This suggests that, although the present airport regulatory framework in Germany is inadequate, RPI-X price cap regulation may not be the best solution, and that a framework for facilitating negotiated agreements would be preferable.

Example: the EU Airport Charges Directive 2009

This Directive does not require price controls on airport charges, but instead specifies a consultation procedure between airports and users to cover the structure (system) and level of charges, and quality of service. There is emphasis on transparency and the exchange of information with respect to cost structure, traffic forecasts and the impact of airport investments. Where possible changes in charges should be by agreement. There is provision for an independent dispute resolution procedure.

The Directive is not without its problems. It is presumably designed to apply where competition is not effective, and

---

to that end is compulsory for all airports with over 5 million passengers/year. However, in the UK there is effective competition up to the level of about 30 million passengers/year. The specification of the cut-off point for effective competition therefore needs further consideration.

A particular concern in Germany is that the Directive provides that appeal to an existing regulatory body would constitute an acceptable independent dispute resolution procedure. Yet in Germany the existing regulatory bodies are the federal states that also own the airports. This conflict of interest would compromise the independence of the dispute resolution procedure. In other respects, the Directive would represent a useful development.22

Conclusions

If regulation is to seek to replicate competition, then effective regulation necessitates an appropriate definition of effective competition. This should reflect competition as a dynamic market discovery process. The criteria for assessing competition refer not only to price in relation to cost (property A) but also to finding more efficient methods of production (property B) and to the discovery and meeting of customer preferences (property C).

An analysis of all three aspects of effective competition has informed previous policy decisions, notably with respect to privatisation, competition and regulation. It should similarly inform present and future policy decisions.

Where there are prospects of effective competition emerging, the misapplication of price cap regulation on a transitional basis may not replicate all these aspects of effective competition, and may indeed deter it. In these circumstances a safeguard price cap is worth considering.

Where there is little prospect of effective competition emerging, newer regulatory approaches better replicate the market discovery process than do the conventional cost-of-service approach or the RPI-X price cap approach. There are now numerous examples of these newer approaches from which to learn.

In both cases, there is scope for competition and utility regulators to be more innovative in future – and innovation is, after all, a central feature of effective competition.

Günter Knieps

The Three Criteria Test, the Essential Facilities Doctrine and the Theory of Monopolistic Bottlenecks

Since the abolishment of legal entry barriers in network industries effective competition has become a key topic in regulatory economics. When the debate on the possibilities of privatisation and deregulation started in the late 1970s and early 1980s, the primary focus was on active and potential competition in the markets for network services. Whereas the focus of the theory of contestable markets concentrates on the role of potential competition with identical cost functions for both active and potential competitors1, competition in the markets for network services not only means potential competition. As soon as competition takes place, the behaviour of players in markets for network services becomes more complex than is assumed in the model of the theory of contestable markets. Examples may be strategies of product differentiation, price differentiation, creation of goodwill etc. However, even strategic behaviour in competitive markets should not lead to the conclusion that these markets should be regulated. On the contrary, the goal of the disaggregated regulatory approach is the development of the preconditions for effective competition in the markets for network services.

The focus of this paper is on the division of labour between competition policy and regulatory interventions. Whereas competition policy aims to intervene once anti-competitive conduct has been identified, the focus of sector-specific regulation is on ex ante regulatory provisions before an abuse of market power has taken place. Ex ante regulation is only justified in those kinds of network areas in which a systematic abuse of market power is likely in the absence of regulation.

The potentials for phasing out sector-specific regulation are of particular relevance in the telecommunications sector.2 In 1999, an EU review was started with the aim of...


maximising the application of general European competition law, the minimisation of sector-specific regulation, a rigorous phasing out of unnecessary regulation and the introduction of sunset clauses. Nevertheless, the unspecific regulatory obligations based on the EU directives in the 1999 review package – in particular the Framework Directive\(^3\) and the Access Directive\(^4\) – resulted in a tangle of contradictory decisions and statements.\(^5\) The Commission’s guidelines\(^6\) do not present a clear and economically well-founded concept for localising network-specific market power. Criteria like relative market share, financial strength, access to input and service markets and so forth can only serve as a starting-point for evaluating the existence of market power, but the development of an ex ante regulatory criterion creates a need for a more clear-cut definition of market power. Nevertheless, in the meantime the process of phasing out sector-specific regulation gains increasing momentum.\(^7\)

In the following the question is considered, whether and to what extent sector-specific regulation is essentially a transitional phenomenon, or whether important differences can be found between the individual sectors.

**Regulatory Reforms Towards Rule-based Regulation**

**The Theory of Monopolistic Bottlenecks**

The theory of monopolistic bottlenecks is central to the disaggregated regulatory approach in terms of locating network-specific market power to determine the minimum basis for regulation.\(^8\) The aim is to come up with a coherent basis for access regulation consistent with network economics which can be applied to all network sectors and which regardless of historical or institutional quirks provides justification for ex ante regulatory measures. For the remaining competitive network areas the application of general competition law is sufficient.

The conditions governing a monopolistic bottleneck are met when:

- A facility is necessary for reaching customers, i.e. if no second or third such facility exists, in other words if there is no active substitute. This is the case when due to economies of scale and economies of scope a natural monopoly exists and a single provider is able to make the facility available more cheaply than several providers.

- At the same time the facility cannot reasonably be duplicated as a way of controlling the active provider, in other words when there is no potential substitute. This is the case when the costs of the facility are irreversible.

Consequently, network-specific market power is only to be expected in those parts of networks which are characterised by a natural monopoly and irreversible costs. Although irreversible costs are no longer relevant for the decision-making of the established enterprises, potential competitors must decide whether or not to invest in such irreversible costs. This means that established firms have room for strategic behaviour, with the result that inefficient production or profits no longer necessarily stimulate newcomers to enter the market. The market power of the firm that enjoys such a monopolistic bottleneck is therefore stable, even if all market players are fully informed, all users are prepared to switch to another provider, and small price adjustments have an effect on demand.

**Monopolistic Bottlenecks and the Concept of Essential Facilities**

When applying rule-based regulation in order to discipline network-specific market power, the concept of essential facilities is of crucial importance. This concept suggests the connection to the essential facilities doctrine, derived from US antitrust law, which is now increasingly being applied in European competition law. The doctrine states that a facility is only to be regarded as essential if the following conditions are fulfilled: entry to the complementary market is not effectively possible without access to this facility; it is not possible for a supplier on a complementary market to duplicate this facility at a reasonable expense; and there are also no substitutes.\(^9\)

---


6 Commission guidelines on market analysis and the assessment of significant market power under the Community regulatory framework for electronic communications network and services, OJ 2002, C 165, pp. 6-31.


8 G. Knieps: Phasing out... op. cit., here pp. 327-331.

In the context of the disaggregated regulatory approach the essential facilities doctrine is no longer applied case by case – as is common in US antitrust law – but to an entire class of cases, namely, monopolistic bottleneck facilities characterised by a combination of natural monopoly and irreversible costs in the relevant range of demand. The design of non-discriminatory conditions of access to essential facilities must be specified in the context of the disaggregated regulatory approach. It is important in this context to view the application of the essential facilities doctrine in a dynamic context. Therefore, an objective for the formulation of access conditions must be not to obstruct infrastructure competition by regulatory micro-management, but rather create incentives for the symmetric development of infrastructure and service competition by rule-based regulation.

**Monopolistic Bottlenecks and the Three Criteria Test**

In the context of European telecommunications policy, in February 2003 the European Commission recommended the so-called three criteria test. This test seems to substantiate the requirements for regulatory intervention. The Commission summarises the three criteria as follows: “The first criterion is the presence of high and non-transitory entry barriers whether of structural, legal or regulatory nature. … [T]he second criterion admits only those markets, the structure of which does not tend towards effective competition within the relevant time horizon … The third criterion is that application of competition law alone would not adequately address the market failure(s) concerned”.10

Thus, the intention to avoid over-regulation with respect to new markets can be observed in the EU telecommunications regulatory framework. However, an economic approach to the remaining need for sector-specific regulation is still missing. In order to provide a consistent regulatory framework, the three criteria in the Commission’s Recommendation of February 2003 have to be rewritten in economic terms, applying the theory of monopolistic bottlenecks.11 After entry liberalisation of network industries, high and non-transitory entry barriers are only present if a monopolistic bottleneck infrastructure exists. Markets do not tend towards effective competition within the relevant time horizon as long as a natural monopoly in combination with sunk costs is stable over the foreseeable future without phasing out potential. The question whether the application of competition law alone would adequately address the market failure(s) concerned raises the question whether ex ante or ex post intervention is more efficient.

**Disaggregated Monopolistic Bottleneck Regulation**

**The Need to Regulate Third Party Access**

It is important to differentiate between efficient private bargaining of access conditions among competitive networks and regulated third party access to monopolistic bottlenecks. Competition fulfills the function of mitigating market power. It can be expected that private bargaining of access conditions between the different network owners under competition will lead to economically efficient solutions. Strategic behavior can be excluded, because every bargaining partner can easily be substituted by an alternative (potential) network carrier. Private bargaining solutions on access conditions among network carriers under competition are not only beneficial for the carriers themselves but in particular improve the market performance of the network services provided to the customers. Independent of the market size of the carriers involved, inefficient suppliers of access services are rapidly confronted with strongly decreasing market shares due to the strong pressure of alternative (potential) network service providers.

In order to allow active and potential competition on service markets, non-discriminatory access to monopolistic bottleneck infrastructures is necessary. To the extent that a monopolistic bottleneck is observable, ex ante regulation should be in place; otherwise the evolution of service markets will be hampered. Innovative ways of access to existing bottlenecks should be guaranteed in order to allow the evolution of new service markets.

Illustrative examples of monopolistic bottleneck infrastructures are railway systems, where competitive suppliers of rail services need access to the tracks and railway stations. In contrast to rail services, railway tracks must be regarded as a natural monopoly with sunk costs, which cannot be shifted to another market. Therefore, if a potential competitor planned to enter with a parallel track, the incumbent railway owner could reasonably claim to reduce his tariffs to short-run variable costs. As soon as a railway network is completed, one therefore cannot expect further entries with additional tracks. The decision-relevant costs of entry include the costs of tracks, which could not be covered by tariffs based on short-run vari-

---

able costs, and thus the established track owner has obtained market power. A similar situation holds for other network infrastructures, such as airports, electricity and gas networks.

In contrast to access to competitive networks, the market power involved in monopolistic bottlenecks fundamentally disturbs bargaining processes. One extreme alternative could be (vertical) foreclosure of competitors on a complementary service market. Such a tying can be used as a method of price discrimination, enabling a monopolist to earn higher profits. Another way of abusing market power within the bargaining process on access conditions is to provide insufficient network access quality or require excessive access charges.

Until now the role of regulation from the perspective of intramodal competition has been considered. The question arises to what extent effective intermodal competition, e.g. the entry of low cost airlines, could restrict the market power of the Deutsche Bahn AG. There is no doubt that intermodal competition by trucks, cars and airlines can limit to some extent the profit potential of a railroad infrastructure provider. However, the goal of the new railway regulations in Germany is to stimulate active competition on the railroad service markets. Opportunities for new entrants include the detection and exploitation of new train service networks, such as the development of a Europe-wide express service for passengers and goods based on a high-performance, computer-assisted logistics system. Mandatory access requirements to tracks are based on the intramodal perspective of train service companies; the decisive factor is the need for complementary service providers to have non-discriminatory access to the monopolistic railway infrastructure.

Limiting Regulation to Monopolistic Bottlenecks

Access to monopolistic bottleneck facilities can also be excluded by providing access only at prohibitively high tariffs. This shows that an effective application of the essential facilities doctrine must be combined with a suitable regulation of access conditions to monopolistic bottlenecks with regard to price, technical quality and timeframe. However, the fundamental principle of such a regulatory policy should be to strictly limit regulatory measures to those network areas where market power potential does indeed exist. There are two further issues that have to be taken into account. On the one hand, the existence of competition on the service level should not lead to the conclusion that there is no market power potential on the upstream network level, as long as the latter fulfils the criteria of a monopolistic bottleneck. On the other hand, there is the question of the minimum regulatory depth necessary to guarantee non-discriminatory access to essential facilities without, however, disproportionately interfering with the property rights of the regulated firm.

Price-cap Regulation of Access Charges

The reference point for regulatory rules concerning access charges should be the coverage of the full costs of the monopolistic bottleneck (in order to guarantee the viability of the facility). In particular, when alternatives to bypass essential facilities are absent, the cost-covering constraint may not be sufficient to forestall excessive profits. In this case, price-cap regulation should be introduced. Its major purpose is to regulate the level of prices, taking into account the inflation rate (consumer price index) minus a percentage for expected productivity increase. It seems important to restrict such price-cap regulation to the bottleneck components of networks, where market power due to monopolistic bottlenecks is really creating a regulatory problem. In other subparts of networks price-setting should be left to the competitive markets.

Regulation of infrastructure access charges should be limited exclusively to price-capping. The basic principle underlying price-capping regulation is that price levels should be regulated in areas where there is network-specific market power. The benefits of price-capping in terms of efficiency improvements and future investment activities can only unfold if price-capping is not combined with input-based profit regulation. Individual pricing agreements lead to over-regulation, which is harmful to competition.

The question remains whether regulators should also be allowed to prescribe pricing rules for tariff structures within monopolistic bottlenecks. There are serious arguments for regulators to refrain from detailed tariff regulation. In the first place, firms should have the flexibility to design (Pareto superior) optional tariff schemes. Pricing rules prescribed by the regulator could induce inefficient bypass activities. For example, a first pricing rule could be access tariffs according to the long-run average costs of the essential facility. Since in such a case differentiation among different user groups according to different price-elasticities is impossible, incentives for an inefficient bypass of the bottleneck facility may be created for certain

12 C.M. Bender, G. Götz, B. Pakula: Effective competition: what it means and why it is relevant for network industries, in: Intereconomics, pp. 4-10.


user groups. A second pricing rule would be access pricing according to the Ramsey pricing principle. Mark-ups on the marginal costs of access to the monopolistic bottlenecks are chosen according to the elasticity of demand for network access in order to maximise social welfare given the cost-covering constraint. However, Ramsey prices could become unsustainable, even if applied strictly to monopolistic bottlenecks. The technological trend towards the unbundling of monopolistic bottleneck components increases the possibilities for inefficient bypass. Secondly, the danger arises that regulators extend the regulatory basis to include the competitive subparts of networks. From the point of view of increasing static (short-run) efficiency such behaviour could even be justified by welfare theory. It is well known that efficiency distortions caused by applying Ramsey pricing can be reduced by extending the regulatory basis. Nevertheless, such an endeavour would in fact mean a return to fully regulated networks, including price and entry regulation of the competitive subparts. As such, this would not be a suitable response to problems of dominance after deregulation. Hence, the design of pricing rules should be part of the decision-making process of the firms.

Towards a Disaggregated Regulatory Mandate

When transferring regulatory competencies from a legislative body to a regulatory authority, the regulatory authority’s future scope of responsibilities is also defined. This involves a regulatory mandate between the legislator (principal) and the regulatory authority (agent). The regulatory authority can be granted varying competencies in this respect.16

European telecommunications regulation is a clear example of how a vague regulatory mandate can systematically lead to overregulation. As has been pointed out, the phasing out process of sector-specific regulation gains increasing momentum. Nevertheless, due to the vague regulatory mandate of the EU regulatory framework the EU Commission gained a large potential of regulatory discretion. The markets which the Commission classified as potentially in need of regulation include service markets such as international and domestic telephone calls, leased lines and transit services that are undoubtedly competitive. The EU Directives are an ideal breeding ground for varied forms of discretionary intervention, depending on the particular influences of the interest groups involved.17

In order to avoid the problem of extensive discretionary behaviour by regulatory agencies the legislator should limit the regulatory authority’s discretionary behaviour by means of the following disaggregated regulatory mandate. Firstly, regulation should be limited to areas with network-specific market power. End-to-end regulation, which also includes competitive areas, is incompatible with this principle. Secondly, when the network-specific market power disappears in a network area, say as a result of technical progress, regulation of this sub-area must also be ended. Thirdly, non-discriminatory access to the monopolistic bottleneck facilities must be ensured. Incentive regulation must be restricted to monopolistic bottleneck components. The disaggregated regulatory mandate should also provide a binding restriction on the regulatory authority’s competencies and consequently reduce its possibilities for opportunistic behaviour.

Paul W.J. de Bijl*

The Need for a Communications Regulator: a Lesson from the Netherlands

In July 2009, Morgan Stanley1 published a thought-provoking report that, based on a conversation with a fifteen-year old summer intern, reported on media consumption patterns of teenagers. They do not want to pay for content on the Internet, and they do not want to be bothered with intrusive advertisements. They enjoy searching for music and other content with their mobile handsets, preferably

Wi-Fi-compatible devices, since 3G networks are expensive. They are hardly interested in e-mail and printed media but do not hesitate to pay for concerts and movies. The world has changed. Telecoms have been transformed by the Internet. The sector is now closely connected to the media sector, a phenomenon known as

* The author is grateful to Georg Götz and Anton Schwarz for helpful comments and suggestions.

---


---

17 G. Knieps: Telecommunications markets..., op. cit., p. 78.
To discuss whether the current regulatory framework can deal with the future, this paper builds on Noam. For more extensive discussions that go beyond the design of regulatory institutions, see De Bijl. In that paper, I develop the argument that policymakers have overestimated the viability of competition in infrastructure, although in the public debate, some commentators claim that “the regulator has finished its job” – a claim that I challenge. Complementary work includes Littlechild, who discusses the meaning of effective competition in relation to regulatory goals. Also related is Bender et al. on effective competition in network industries. They propose a notion of effective competition based on the innovative process. The authors discuss that it may be necessary to take a chance on deregulation in order to discover if the market is competitive from a dynamic perspective.

Market Failures in Electronic Communications

Liberalisation and privatisation in the telecommunications market have had a multifaceted background. First, there was a desire for improved efficiency and investments. The era of centrally managed utilities had ended and governments were unable to continue to fund these investments themselves. For government bodies, the pursuit of cost efficiency, modernisation and innovation is of little concern and has low priority. Second, technological advances had brought down the costs of rolling out networks. This was expressed in lower infrastructure costs and the emergence of mobile telephony. This argument...
eroded the prominence of the natural monopolistic nature of the market and it was believed to pave the way for the viability of several providers of fixed-line networks. Third, governments were unable to keep up as well as the market could with the dynamic innovations that presented themselves. Finally, monopolies were at odds with the internal market principles in the European Union (EU).

The fact that competition was introduced does not mean that the market failures from the past have been "resolved", or that there are no more market failures in the telecom sector. The principal market failures in telecommunications markets are [examples of government intervention are given in square brackets]:

- market power and distortion of competition, with underlying causes:
  - economies of scale (natural monopoly) and bottlenecks [regulation of access];
  - costs of switching to another provider [imposition of number portability];

- externalities, in particular:
  - network externalities [enforcement of interconnection and interoperability];
  - importance of communication for the economy as a whole [imposition of duty to provide universal service and enforcement of interconnection and interoperability];
  - failures due to signal interference within the radio spectrum [central control of spectrum allocation];

- coordination problems [enforcement and coordination of technological standards and radio spectrum policy];

- consumer protection in the field of:
  - asymmetrical information [combating the switching of a consumer’s telephone service to another telephony service provider without the customer’s consent, also known as slamming];
  - risks of costly mistakes [regulation of rates for mobile calls made while abroad];
  - other [e.g. combating spam].

In addition to these market failures, social or political preferences may also constitute an argument for government intervention:

- social or political preferences with respect to:
  - redistribution [regulation of end user rates; universal service obligations];
  - social aims [universal service obligations].

This overview shows that electronic communications is a sector in which virtually all forms of market failure occur. The principal market failures, in light of recent and future developments, are market power and externalities. Market power failure is due to the increased importance of economies of scale (due to next-generation networks (NGN) investments, as will be discussed below) and also to the vertical integration between networks and content services, which can reduce competition. In addition, the small number of players with networks means that imperfect competition is a recurring fact of life in telecoms. Scale effects still exist in the form of the significant fixed costs associated with building fixed-line, finely meshed networks. Bottlenecks or essential facilities are present in some parts of networks, resulting in market power. A bottleneck is a part of a network of fixed proportions that (i) constitutes a necessary input in order to be able to reach end users, while (ii) this facility cannot be reproduced in an economically efficient manner because there are substantial embedded costs.10 This latter condition means that a bottleneck constitutes a barrier to entry. An example is the connection to individual end-users; without access to them, a competitor is unable to deliver a call request to another operator’s customers. Another market failure that leads to market power is the existence of costs for consumers wanting to switch to another provider. By imposing number portability, regulators have been able to reduce these costs to some extent.

In the Netherlands, the regulatory authority OPTA’s interpretation of the Telecommunications Act is mainly based on curbing anti-competitive market power (this touches on the first point in the overview above) and consumer protection (point 4). The scope of the Telecommunications Act is greater, which expands OPTA’s scope of activities to include imposing an obligation on incumbent KPN to provide universal service.

A Market in Transition

Noam11 distinguishes between three phases in the telecoms market. Telecom 1.0 was the age of public monopolies with a utility function in voice telephony over fixed, analogue networks. In the EU, this age ended with the privatisation of the monopolists and the liberalisation of the market. In 1998, these operations resulted in the creation of the European regulatory framework for the liberalisation of the markets for networks and services in the telecoms sector. This moment marked the start of Telecom 2.0.

---


11 E. Noam, op. cit.
Policymakers and regulators, such as OPTA in the Netherlands, drew up their designs for the market and the necessary transition. Former monopolists were exposed to competition and new regulations and transformed themselves into commercial enterprises. Newcomers, who did not have their own networks from day one, would initially be entitled to make use of regulated access to existing infrastructure. Telecom 3.0 began around 2008 with the emergence of next generation networks (NGNs), which make very high speed data traffic possible, for example via fiber-optic links to end-users ("FTTH"). The next five to ten years will be dominated by questions as to (i) whether, how and when the mass migration to optical fiber networks will take place, and (ii) what will be provided over those networks. As will be discussed, the assumption that the natural monopoly nature of telecommunications was over – which was an important basis at the time for the transition to Telecom 2.0 – has become less realistic in this third phase, since NGNs demonstrate the increased importance of investments and scale effects.

De Bijl extensively discusses how the Dutch market developed during Telecom 2.0 and into 3.0.\textsuperscript{12} I will recapitulate the main points. When the market was opened up in 1998, it was expected that facilities-based competition (competition between operators with their own networks) would soon arise. At least two new entrants with the potential to provide nationwide coverage were waiting in the starting blocks\textsuperscript{13}: Enertel, a consortium of various power and cable companies in the Netherlands, and Telfort, an alliance between NS (Dutch national railways) and British Telecom. In addition to these two players, who were expected to develop into competitors of KPN with their own networks offering nationwide coverage, there was also the prospect of local entrants in market segments. It was believed that after making the necessary adjustments to their networks, regional cable companies would be in a position to add telephony to their product range.

Approximately five years later the picture had changed. Enertel and Telfort had failed to develop into nationwide competitors to KPN. Cable companies initially showed little interest in the market for voice telephony. Nonetheless, various entrants rolled out networks for the business market in a very targeted manner at well-chosen locations. However, the construction of local loops (connections to end-users) failed to materialise in most parts of the country and particularly for residential customers. In spite of the absence of competition in infrastructure for residential end-users, consumers nonetheless benefited from price competition, since OPTA ensured that KPN opened up its network to competitors who did not have their own network. The former monopolist was also obliged to unbundle the local loop\textsuperscript{14}, although competitors made little use of this facility in the first few years. It was hoped and expected that, with a little more time, entrants would also roll out networks in the segment not yet targeted by facilities-based entrants (in particular the residential segment). Economists labelled this approach the ladder of investment: newcomers are encouraged to gradually make investments that reduce their dependence on established network operators.\textsuperscript{15} Economists had warned against the risk that network access regulation could disrupt the roll-out of competing networks.\textsuperscript{16} Low access tariffs stimulate competition in services but discourage new entrants from investing in a network, resulting in the ongoing necessity of network access regulation.\textsuperscript{17} In response to these concerns, OPTA planned to implement the ladder of investment by allowing the regulated price for the unbundled local loop to rise gradually over a five-year period so that entrants were stimulated to roll out their own networks over time. OPTA’s plan did not work out as intended, however: the regulated price actually decreased over time.\textsuperscript{18}

At present it is unclear whether we have already arrived at a situation of effective competition in infrastructure. OPTA considered the question of whether two networks are sufficient to guarantee competition and came to the conclusion that this is not the case.\textsuperscript{19} Nonetheless, a duopoly could, in theory, result in sufficient competition since voice telephony and broadband Internet are basically commodities with vertical quality aspects, such as speed of connection, as the principal distinguishing features. Bijlsma and Van Dijk\textsuperscript{20} show that in theory, the current situation

\textsuperscript{12} P.W.J. de Bijl, op. cit.
\textsuperscript{13} The description of the developments following the liberalisation of the telecoms market in 1998 is partly taken from: P.W.J. de Bijl, M. Peitz: Regulation and Entry into Telecommunications Markets, 2002, Cambridge, UK, Cambridge University Press.
\textsuperscript{14} Local loop unbundling: a competitor rents local lines from the incumbent in order to offer its own services to end-users over those lines. Calls are therefore not completed per minute, as is the case with Caller Select.
\textsuperscript{16} See e.g. CPB: Competition in Communication and Information Services: opportunities and obstacles, Bijzondere Publicatie 12, 1997; and E.E.C. van Damme: Competition in the local loop, in: Visions, study for VECAI, 1999, Den Haag.
\textsuperscript{17} M. Grajek, L.-H. Röller: Regulation and Investment in Network Industries: Evidence from European Telecoms, Working Paper No. 09-004, 2009, European School of Management and Technology (ESMT), Berlin, empirically confirms this worry, based on empirical research using data from the EU.
\textsuperscript{20} M. Bijlsma, M. van Dijk: Nieuwe generatie netwerken, nieuwe generatie reguleren?, CPB Document No. 145, 2007, CPB.
of two networks can ensure sufficient competition, since they will compete to provide access to third parties, if the networks are not too differentiated. In time, when networks find and implement methods of horizontal differentiation, this will possibly no longer be the case.\textsuperscript{21} Note that Schwarz is relatively optimistic about the competitive pressure offered by mobile broadband.\textsuperscript{22}

It is likely that the business propositions of many new entrants without plans to make substantial investments were based on the possibilities of administrative arbitrage created by network access regulation. On the other hand, continued growth towards facilities-based competition was part of the market development foreseen within the European framework. This outcome has yet to crystallise. Arguably, effective, (i.e. sustainable) competition still appears to be as far away as it was at the start of liberalisation in 1998. In the Netherlands, differences that have developed historically, such as the special position (i.e. monopoly rights) of the cable networks with regard to the transmission of radio and TV programmes, should also be set right so as to create a level playing-field. Competition would then probably concentrate more on the future in order to gain a competitive advantage by upgrading the networks.\textsuperscript{23}

Recent developments (Telecom 3.0) underline the fact that scale effects have effectively never disappeared, since they have a considerable impact on the development of network upgrading to NGNs and network roll-out. After all, upgrading networks to optical fiber calls for significant investments. Besides the fact that the telecoms sector is not immune to the cyclical dynamism of the economy\textsuperscript{24} one can even wonder whether competition is ever going to work. Noam compares the current situation to Telecom 1.0 and asks whether we have arrived at a high-speed version of the original system, albeit cross-border with converged markets and “supplemented by oligopoly, at best”.\textsuperscript{25} Thus, the question for policymakers is as follows: is the original policy approach going to work in the environment in which we now find ourselves? While the European framework still considers competition in infrastructure as the final objective, it is sensible to take into account the possibility that it will not happen. This makes it possible to include a broader range of interventions, such as structural separation and public support for network investments.\textsuperscript{26} The new generation of optical fibre networks requires high investments, perhaps too high for viable competition between several networks that are prevented from sharing the costs of their investment. Huigen and Cave argue that it is likely that there will be no major investments in fixed local loops in most member states of the EU.\textsuperscript{27}

The Need for a Converged Communications Regulator

As argued above, it is still uncertain whether the telecoms sector, which has converged with media and Internet, can support effective competition. In spite of the current reality, in the Netherlands a frequently asked question is whether a sector-specific regulator will remain necessary or whether the competition authority can assume the “last remaining”, sometimes even referred to as “modest”, duties. The continuing uncertainty about network investments makes clear, however, that we still have a long way to go. In addition, market failures are persistent and are likely to become more prominent in the near future, for instance related to privacy, reliability of critical infrastructures and (national) cyber-security. Overall, the market is not moving towards a need for less regulation.

A relevant question is whether the current forms of supervision and regulation of telecommunications and media should converge, parallel to the disappearance of the dividing lines between telephony, Internet and media. There are significant precedents internationally for bundling supervision. In the UK, the Broadcasting Standards Commission, the Independent Television Commission, the Office of Telecommunications (Oftel), the Radio Authority and the Radiocommunications Agency merged at the end of 2004 to form Ofcom (Office of Communications). The USA has had a similar form of regulation for much longer. In 1934, the Communications Act was passed with the abolition of the Federal Radio Commission, and the Federal Communications Commission (FCC) was established. This regulatory authority is responsible for policy with regard to media and spectrum and for implementing the Telecommunications Act.

---


\textsuperscript{22} A. Schwarz, op. cit.

\textsuperscript{23} See P.W.J. de Bijl, M. Peitz: Innovation, convergence and the role of regulation in the Netherlands and beyond, op. cit.; and C.M. Bender, G. Götz, B. Pakula, op. cit.


\textsuperscript{25} E. Noam: Regulation 3.0 for Telecom 3.0, op. cit.


The Netherlands remains far from this type of institutional set-up. Redesigning broadcasting policy is a politically sensitive issue, due, amongst other things, to the historical origins of the broadcasters and their members. Nevertheless, De Bijl, Dommering and Van Damme et al. call for the existing regulatory authorities for telecommunications and media (OPTA, the Radiocommunications Agency Netherlands and the Dutch Media Authority) to be merged into a new regulatory authority, separate from the competition authority NMa. One reason to have a “converged” communications authority is convergence. A merger along the sectoral dimension enables the better alignment and consistency of regulation of electronic communication and electronic services, including broadcasting distribution and substantive aspects of content. It is likely that there are large synergies to be gained between the legal, economic and technical domains of regulatory design and market oversight. Also, a communications regulator could modernise part of media policy and integrate it with communications policy. Another consideration is that market failures in the communications and media sectors are substantial and will remain so. They affect all ranks and levels of society and involve more than just the risk of abuse of market power – recall the overview of market failures discussed above.

The above authors also call for maintaining the dividing line between ex ante regulation and ex post competition regulation, for example with the possibility of referral to the competition authority, and with a possible curtailment of sector-specific regulation (in particular network access regulation) where there are signs that competition in infrastructure is developing favorably. Another option that is often referred to, namely incorporating the communications regulator into the competition authority, could lead to too much emphasis on the competition perspective of ex post intervention, which is less concerned with other market failures than market power.

In spite of the existence of arguments for scaling back sector-specific regulation, it is crucial to design smart forms of ex ante regulation: simple, light-handed rules that place easy demands on the availability of information on cost structure or demand parameters. A regulatory authority must nonetheless have the scope to develop such instruments and be able to deploy them, regardless of whether they are contrary to the principle of ex post regulation or demand detailed knowledge of the sector and underlying technology.

**Conclusion**

Since the advent of liberalisation, innovation in electronic communications has brought us tremendous benefits. The question that gradually presents itself in a converging world with substantial scale economies is whether the approach of the European regulatory framework, based on allowing competition in infrastructure to grow under an extensive collection of preconditions and constraints, is still effective.

Although the telecom sector has not developed without some surprises along the way, the market has proven to be stubborn. This is not to say that the regulator provided wrong incentives to market players, but we now need a broader perspective. Policymakers and regulators were well prepared for the task ahead and carried out sound work which has yielded significant benefits to consumers. Nevertheless, market developments are difficult to predict or control in a market such as this with rapid technological developments. For example, a dynamic market pays little heed to detailed transition paths leading towards regulatory blueprints. Also, when push came to shove, the energy and drive that had been put into phasing out network access regulation simply evaporated. As a consequence, we still do not know whether the roll-out of competing networks is viable in practice, which makes one wonder if effective competition is actually possible in telecoms. Perhaps we were too optimistic with the idea that scale effects were no longer relevant in telecommunications – though for a different perspective, note Littlechild’s remarks on dynamic competition. We cannot sit back though. The time has come to create an integrated, sector-specific regulator for communications and media.


32 If the communications authority is asked to deal with a competition problem which is beyond the scope of the sector-specific law, then the problem can be referred to the competition authority. This is how things are arranged in the UK: Ofcom can make a referral to the Competition Commission.

33 See: P.W.J. de Bijl, M. Peitz: Innovation, convergence and the role of regulation in the Netherlands and beyond, op. cit.


35 S. Littlechild: The nature of competition and the regulatory process, op. cit.
European national regulatory authorities (NRAs) for telecommunication have to periodically review a number of markets recommended by the European Commission.1 One of these markets is the wholesale broadband access market, where alternative operators can buy a wholesale product (“bitstream services”) from the (regulated) incumbent operator and offer broadband access to retail consumers.

Whether or which ex ante regulation is needed on this wholesale broadband access market is analysed by NRAs in a three-step process: (i) market definition, (ii) market analysis and (iii) the determination of appropriate remedies, if significant market power (SMP) is found in the second step.2 While market definition is “only” the first step in this process, setting the scene for the competition analysis, the scope of the market can have a significant impact on the final result. The “case study” presented in this paper is a good example of the importance of market definition for determining whether competition is effective.

This article deals with the market definition of broadband markets in Austria as laid down by the Austrian Regulatory Authority for Broadcasting and Telecommunications (RTR GmbH). The focus will be on substitution among DSL, cable and mobile broadband in the residential segment. Austria is a good subject for such a case study since mobile broadband (broadband provided by mobile network operators based on UMTS/HSPA-Technology) is particularly advanced compared to almost all other European countries.

The Broadband Market in Austria

Broadband Internet access, i.e. Internet access with a maximum download speed in excess of 144 kbit/s, is offered in Austria based on the following technologies:

- DSL: broadband transmission based on copper wires, coverage >95% of households;
- cable: broadband transmission based on coaxial cable-TV networks, coverage ~50% of households;
- mobile broadband: broadband transmission based on UMTS/HSPA, coverage ~95% of households;
- other technologies: W-LAN/WiFi/WiMax, Fibre-optic (FTTH – fibre to the home), power line (power cable), satellite.

Figure 1 shows the evolution over time of the number of DSL, cable, mobile broadband and other access lines. Throughout Austria, the total number of broadband connections as of the end of December 2009 amounted to approximately 3.17 million. Since the beginning of 2007 a strong increase in the number of mobile broadband connections can be observed. If we compare the access technologies in December 2009, mobile broadband had already taken the lead with 41%, followed by DSL with 40% and cable with 18%. Austria is thus one of the leading countries with regard to mobile broadband adoption. In January 2010, Austria ranked third within the EU, with a mobile broadband penetration of 15.1% of the population (dedicated data service cards/modems/keys), well above the EU average of 5.2% and Germany with 4.0%.3

Mobile broadband connections through UMTS or HSPA/HSPA+ are offered by all four mobile network operators. The download rates that can be realised depend on the end device, the number of users in a mobile cell, and the bandwidth of the connection to the base station. While maximum bandwidths currently are 7.2 Mbit/s (HSPA) and 21 Mbit/s (HSPA+) the actual average bandwidths are mainly between one and three Mbit/s.4

Literature on Fixed-Mobile Broadband Substitution

While there is some academic research with regard to substitution between fixed narrowband and fixed broadband access or between different types of fixed broadband access, in particular DSL and cable\(^5\), there is — to our knowledge — hardly any empirical analysis which includes mobile broadband. This is likely due to the fact that the mobile broadband boom is a quite recent phenomenon which also varies in importance across different countries. Cardona et al.\(^6\) include mobile broadband in their analysis of the Austrian broadband market; however, as their analysis relates to 2006/2007 — just before the mobile broadband boom in Austria — they do not find sufficient evidence for fixed-mobile substitution. The present article can be viewed as an update of Cardona et al. since it looks at the situation in Austria over the years 2007-2009.

With regard to other countries it should be noted that no NRA — except the Austrian — has so far defined a common broadband market for fixed and mobile services.\(^7\) In many countries, mobile broadband has a much lower penetration compared to Austria or is used complementarily rather than as a substitute. Cave and Hatta\(^8\) argue, for example, that mobile broadband is unlikely to be a substitute for fixed broadband since there is an “order-of-magnitude gap” between fixed and mobile broadband speeds.

Analysis of Fixed-Mobile Broadband Substitution

Clearly, the definition of the relevant market needs to be “technologically neutral”\(^9\) in the sense that substitution and competitive constraints between products delivered by all potentially relevant technologies have to be considered.

The analysis of the Austrian Regulatory Authority for Broadcasting and Telecommunications\(^10\) is based on a number of qualitative and quantitative criteria. This article will focus on the following elements:

- prices and product characteristics of fixed and mobile broadband
- evidence from a consumer survey (3000 households, 1000 businesses) conducted in January 2009
- price-quantity developments / price reactions.

The focus will be on the residential segment although the business segment will also briefly be touched.

Prices and Product Characteristics

While fixed and mobile broadband services can clearly be used to access the internet at higher speeds compared to

---


\(^9\) See also the contribution by C.M. Bender, G. Götz and B. Pakula in this issue, pp. 4-10.

The adoption of fixed and mobile broadband already reveals significant differences between business and residential users. The share of mobile broadband users among private households (27%) is almost twice as high as the share of mobile broadband users among businesses (15.5%). In addition, while about 75% of residential mobile broadband users use only mobile broadband, the complementary use of fixed and mobile broadband is prevalent in businesses: only about 25% of mobile broadband users use “mobile only” – exactly the opposite ratio as in the residential segment.

In the case of residential users, switching from DSL/cable to mobile broadband has occurred to a similar extent as switching between DSL and cable (see Figure 2). Based on these figures it can be estimated that roughly 10% of DSL and cable users have already switched to mobile broadband (“mobile only”) in the past. This percentage is much smaller for businesses.

The adoption of fixed and mobile broadband already reveals significant differences between business and residential users. The share of mobile broadband users among private households (27%) is almost twice as high as the share of mobile broadband users among businesses (15.5%). In addition, while about 75% of residential mobile broadband users use only mobile broadband, the complementary use of fixed and mobile broadband is prevalent in businesses: only about 25% of mobile broadband users use “mobile only” – exactly the opposite ratio as in the residential segment.

In the case of residential users, switching from DSL/cable to mobile broadband has occurred to a similar extent as switching between DSL and cable (see Figure 2). Based on these figures it can be estimated that roughly 10% of DSL and cable users have already switched to mobile broadband (“mobile only”) in the past. This percentage is much smaller for businesses.

The hypothetical (stated) reactions to a 10% price increase for DSL and cable connections are depicted in Figure 3. Since stated preferences / stated switching may overestimate actual switching (some households may not be aware of the transaction costs or may not even notice the relative price change in practice), the results have to be interpreted with caution. The categories “do not know” and “get offer, am not sure” were therefore counted as “would not switch”. This leaves us with 25% of households who say that they would switch or give up their access. Even this figure can only be regarded as a maximum value, however, since some of the households who say that they would switch do not know to which technology. Subtracting these households results in a lower bound value of 15%. If this range is related to the price increase of 10%, we get an (arch) elasticity of -2.5 to -1.5.

Results from the Consumer Survey

3000 residential users (those who decide about the internet connection in their households) and 1000 business users (IT/telecom decisionmakers) were interviewed by a market research institute in January 2009. They were asked the following:

- What broadband technologies are used?
- Has there already been a switch between different broadband technologies?
- How would the household/business respond to a permanent 10% price increase of his fixed broadband connection (switch or no switch, if yes, to which technology) (“stated preferences SSNIP-Test”)?
- What applications are used via the broadband connection (residential users only)?

11 SSNIP stands for small but significant non-transitory increase in price. The SSNIP-Test has become the standard instrument for market definition in competition economics / competition policy as well as in the field of ex ante regulation.
However, there is no application which is not used by at least 10% of mobile broadband users. This indicates that mobile broadband can be used (and is used) for all applications considered particularly important.

Summing up, the results from the consumer survey indicate that there are strong differences between residential and business users. While business users predominantly use mobile broadband complementary to their fixed connection, it seems to be a substitute for residential users. The estimated elasticity suggests that it is part of the same market as DSL and cable in the residential segment.

Price Reactions and Price-Quantity Developments

Price reactions and price-quantity developments can also be used to make inferences about competitive interaction between products.

After mobile operators lowered prices for mobile broadband significantly at the beginning of 2007, the growth of fixed broadband lines slowed down significantly and even became close to zero in Q2 and Q3 2007. Only after fixed network operators reduced prices substantially at the end of 2007 (“Christmas promotion”) did fixed broadband start to grow again. A similar pattern could be observed in 2008: fixed broadband could only grow significantly during promotion periods. This suggests that changes in the price of mobile broadband have a direct effect on the

In the case of linear demand, the critical elasticity $\varepsilon_c$ is calculated as $\varepsilon_c=1/(m+t)$, where $m$ stands for the “price-cost margin” and $t$ for the extent of increase in price. The share of the variable costs in the total price is estimated at 20-40% by RTR based on data of the incumbent operator. Thus, $m$ lies between 0.6 and 0.8. The increase in price $t$ as per the question in the survey is 10%. The critical elasticity, therefore, is in a range of -1.1 to -1.4.

The entire range of the estimated elasticity is larger in absolute values than that of the critical elasticity. This indicates that the next best substitute, mobile broadband, has to be included in the market.

The same approach was also applied to business users. Here, the estimated elasticity even for DSL services only is much lower, namely -0.7 to -1.8, and unlikely to exceed the critical elasticity.

Finally, residential fixed and mobile broadband users were asked which applications they use frequently with their connections. The results are depicted in Figure 4. Not surprisingly, emails and surfing are the most widely used applications and – as they do not require high bandwidths or a particular stability of the connection – there are only relatively small differences between fixed and mobile broadband users. For applications which require higher download speeds or a particularly stable connection like downloads, online gaming, and internet telephony, mobile broadband somewhat lags behind fixed broadband. However, there is no application which is not used by at least 10% of mobile broadband users. This indicates that mobile broadband can be used (and is used) for all applications considered particularly important.

Summing up, the results from the consumer survey indicate that there are strong differences between residential and business users. While business users predominantly use mobile broadband complementary to their fixed connection, it seems to be a substitute for residential users. The estimated elasticity suggests that it is part of the same market as DSL and cable in the residential segment.
quantity, and finally the prices, of fixed broadband connections.

Conclusions

All in all, the evidence from the consumer survey and past price-quantity reactions is quite conclusive and suggests that DSL, cable and mobile broadband are part of the same market for residential users. RTR was the first NRA in Europe which arrived at such a conclusion. The evidence regarding businesses, on the other hand, indicates that neither cable nor mobile broadband is considered a sufficiently close substitute to include it in the same market as DSL services. Based on this evidence, the wholesale broadband access market in Austria was defined to include only access to bitstream services which are used to supply business customers. Bitstream services used for residential customers were de-regulated. So far, this outcome is unique within the EU.

The market for local loop unbundling is not affected by this substitution in a comparable way. Local loop unbundling is not only an important input for residential broadband products but also for fixed network voice telephony, the provision of broadband, voice and other services to business users or the provision of leased lines to other operators or business users. For these services, mobile services cannot be viewed as substitutes and infrastructure-based competition in the fixed network is limited. Also, local loop unbundling still has its roll in competition with regard to broadband services for residential users. Since the largest mobile operator is part of the fixed network incumbent, it is questionable whether the residential broadband market – despite the inclusion of mobile broadband – could be viewed as effectively competitive without local loop unbundling.

The coming years will be marked by a roll-out of fibre which will partly (FTTC, FTTB) or wholly (FTTH) replace the copper access network. This will substantially increase maximum bandwidth from 20 Mbit/s today to 50 or 100 Mbit/s or even above (in the case of FTTH). Also, some cable networks have already been upgraded to DOCSIS 3.0 and are offering bandwidths up to 100 Mbit/s.

At the same time, mobile network operators are also investing in higher bandwidths. With HSPA+ bandwidths of up to 21 Mbit/s (over the long term up to 42 Mbit/s) can be achieved. After 2011/2012 the successor technology LTE should allow for even higher bandwidths. The download rates that can be achieved practically, however, will always depend on the number of users in a mobile cell and on the bandwidth of the connection to the base station.

If and how these developments will change the substitution between fixed and mobile broadband is a subject for future research.


13 In the Netherlands, "low quality" bitstream services with a contention ratio (“overbooking”) of more than 1:20 were de-regulated in 2005 due to competition from cable networks but “re-regulated” in 2008; see Cullen International, op.cit.

14 See for example the contribution by Paul de Bijl in this issue, pp. 21-26.

Florian C. Haus

Effective Competition and the Essential Facilities Doctrine*

This article focuses on the legal aspects of the essential facilities concept and its relation to effective competition. The author’s take on the essential facilities doctrine is that of a competition lawyer rather than that of an expert in regulatory matters. Nevertheless, it is submitted that competition law and regulatory law should have a similar view on the scope and application of the essential facilities doctrine since, in an ideal world, competition in an unregulated sector and in a regulated sector should eventually be at a similar level.

At the same time, it should be noted that the essential facilities doctrine is one of the most intrusive instruments of modern competition policy. Therefore, competition policy enforcement may not be able to justify each and every sharing of property, which is what it comes down to?
to, with the noble aim of effective competition. This article outlines the current legal framework of the essential facilities doctrine and the role of effective competition in the practical application of the doctrine and then offers a brief conclusion.

Current Legal Framework of the Essential Facilities Doctrine

In April 2010, the so-called Unilateral Conduct Working Group of the International Competition Network issued a 100-page report entitled “Analysis of Refusal to Deal with a Rival Under Unilateral Conduct Law”. This was presented at the 9th Annual Conference of the ICN in Istanbul in April 2010. The essential facilities concept features prominently in this report.1

According to the report, seven jurisdictions specifically define essential facilities: the Czech Republic, Estonia, Germany, Korea, Romania, the Slovak Republic and South Africa. In many more jurisdictions, no explicit reference is made to the essential facilities concept in the statutory law. Nevertheless, the agencies and/or courts have applied the concept in their case law. The European Union is among these jurisdictions. Since the early nineties, the European Commission has built a number of cases on the doctrine; relevant decisions include the cases B&I Line v. Sealink Harbours and Stena2, Sea Containers v. Sealink Stena3, Port of Rodby4, Frankfurt Airport5, British Midland v. Aer Lingus6 and IMS Health7.

Only a few cases were subsequently tested in court. The most prominent case outside intellectual property law is the Bronner case, where the ECJ took a rather conservative approach.8 Bronner is actually the only ruling to date in which the Court of Justice had the opportunity to comment on the scope of the concept of access to essential facilities in the form of a physical infrastructure. The Court of Justice’s decision in Bronner was considered by many commentators as construing the concept of abuse narrowly.9 In Bronner, the publishing house Mediaprint had refused to include the daily newspaper of the publisher Bronner, a competitor of its own products, in its newspaper distribution system, which was the only nationwide distribution system. The Court held

“45. It should be emphasised in that respect that, in order to demonstrate that the creation of such a system is not a realistic potential alternative and that access to the existing system is therefore indispensable, it is not enough to argue that it is not economically viable by reason of the small circulation of the daily newspaper or newspapers to be distributed.

46. For such access to be capable of being regarded as indispensable, it would be necessary at the very least to establish […] that it is not economically viable to create a second home-delivery scheme for the distribution of daily newspapers with a circulation comparable to that of the daily newspapers distributed by the existing scheme.”

Moreover, the Court of Justice explicitly left open whether its own case law in the Magill case10 that relates to intellectual property law could be applied at all to the issue of the access to physical infrastructures. In the most recent decision in this respect, Microsoft, the General Court (formerly the Court of First Instance, CFI) stated that the refusal of a dominant undertaking to grant third parties a licence for the use of a product protected by an intellectual property right is not abusive as such. Only in “extraordinary circumstances” may the exercise of exclusive rights by its owner be regarded as an abusive practice. The following circumstances — cumulatively — are deemed to be extraordinary for these purposes and thus are the necessary conditions for establishing abuse regarding the granting of access:

- the access to the right in question is indispensable to the exercise of a particular activity in a neighbouring market;
- the refusal is of such a kind as to exclude any effective competition in that neighbouring market;

---

the refusal of access prevents the appearance of a new product for which there is potential consumer demand.12

The General Court based its judgment on the precedents set in Magill and IMS Health. In both cases, the Court of Justice decided that the refusal of access to intangible property rights could be assumed to be abusive only if there were extraordinary circumstances in terms of the aforementioned items.

The Commission, however, applies the doctrine as if fully endorsed by the courts. This could be observed recently in the energy sector.13

By contrast, the US antitrust cases, though often referred to as example and model cases for the essential facilities concept, do not serve as a good witness: the 2004 Verizon vs. Trinko, confirmed in AT&T v. Linkline14, implies that the essential facilities doctrine – in the words of the court: “crafted by some lower courts” – cannot be considered as established law.15 In that decision, the Supreme Court also stresses that the famous Aspen case, quoted by many as an authority for the doctrine, “is at or near the outer boundary of §2 liability”, and thus may not easily be replicated in other factual circumstances. The Supreme Court recalls that false-positives, i.e. ordering access where denying access would be all right, are most dangerous because “mistaken inferences and the resulting false condemnations are especially costly, because they chill the very conduct the antitrust laws are designed to protect”.16

Leaving this aside, one may identify as the common ground of the concept under competition law and according to the Report that, firstly, access to the facility must be essential to reach customers, and secondly, replication of the facility must be impossible or not reasonably feasible. Although this may not be surprising to us, it is noteworthy that dozens of agencies worldwide adhere to a similar notion of essential facilities. One does not, however, find a lot of reasoning on “effective competition” in the Report.

In the light of years of converging competition policy, it may appear a logical consequence that German legislation, promoted by the German Federal Cartel Office’s lobbying, introduced a generic access regime in 1999 as a further legislative example of the abuse of a dominant position pursuant to Sec. 19 of the German Act Against Restraints of Competition. This in a way complements the 1996 Telecoms Act, the 1998 Energy Industry Act and the 1998 Act on Postal Services, all of which introduced rules providing for access regimes to the benefit of competitors. The legislator believed he had closed a gap in the law and was eager to provide for a catch-all access regime in network industries that were yet to be identified.

Over the past ten years, we have seen a considerable number of court cases where plaintiffs in the traditional sectors rested on the doctrine as laid down in Sec. 19 of the Act. In many of the early cases, access to gas or electricity grids was sought after. We may recall that before 2005, outright denial of access was more of an issue than it is today, and no comprehensive regulatory regime was in place. Furthermore, in the telecoms sector, companies have sought access to parts of the infrastructure, such as ducts or via bitstream access or GSM gateways, before the courts or by way of complaints to the competition authorities. Outside the regulated industries, access to port infrastructure has called competition authorities into action. Only recently, the Federal Cartel Office issued its second decision since 1999 ordering Scandlines to open up its facilities in Puttgarden to competitors, in order to enable them to provide sea transport to Rödby in Denmark. Scandlines, however, was granted interim relief by the Düsseldorf Court of Appeals on the ground that it appeared unrealistic that the competitor would eventually launch ferry services from Puttgarden since a bridge crossing the sea to Rödby was to be built in the near future.16

One of the leading cases brought by the Federal Cartel Office concerned the allegedly abusive calculation of grid fees by Stadtwerke Mainz. The German Federal Supreme Court finally upheld the Office’s approach in controlling the grid fees. In so doing, the Court in a way also addressed the issue of effective competition under

---

13 RWE gas foreclosure (Case COMP/39.402); GDF foreclosure (Case COMP 39.316); ENI foreclosure (Case COMP 39.315); EON gas grid (Case COMP/39.317); Svenska Kraftnät (Case COMP 39.351).
16 OLG Düsseldorf, Fährhafen Puttgarden II, Decision of 10 June 2010, VI-Kart 1/10 (V), WuW/E DE-R 2941.
the essential facilities doctrine. Under Sec. 19 para. 4 No. 4 of the Act, but also more widely under the doctrine as discussed in US and EU antitrust law, it was disputed in which relevant market dominance had to be established as a prerequisite for the doctrine to come into play. In other words: how and where to look for effective competition. The Federal Supreme Court took the view that it would suffice to establish dominance in the “market” of the essential facility. If one looks for dominance at the infrastructure level, however, enforcement runs the risk of jumping to conclusions, because the analysis will often be narrowed down to the facility at issue from the outset.17

Recent Cases from the Network Industries

In June 2010, the German Federal Supreme Court ended a long dispute in the mobile telecommunication sector, stating that there is no abuse where a GSM operator denies terminating calls which originate from the fixed-line network but are transmitted to its mobile network via so-called GSM gateways. The Federal Supreme Court took the view that where regulated access was in place, also with regard to the essential facilities doctrine, imposing additional access modalities would disregard the operator’s (remaining) legitimate interest to choose if and how to deal with its competitors. Another recent essential facilities case is certainly the E.ON gas grid case. E.ON’s offer to release capacity is the outcome of the Commission’s negotiation process with a number of European players like Gaz de France and Eni during which the Commission took rather far-reaching legal positions, such as a duty to cut back one’s own needs and instead offer third party access and a duty to invest in infrastructure. However, the case may be less attractive for an analysis relating to the issue of effective competition.

This article therefore aims to shed more light on the competition test applied by the Düsseldorf Court of Appeals in its recent cases relating to the access fees for gas transportation.20 Under the current regime for long distance gas transportation – the so-called transmission level – the law provides for an exception to the general rule of ex-ante regulation of access fees. Where effective pipe-to-pipe competition could be established – with the burden of proof resting with the grid operators – no ex-ante regulation would apply. Grid operators stated that they would charge market prices instead of regulated prices. The Federal Network Agency, instead, found that sufficient competition did not exist, and requested the operators to submit their application for regulated fees. The Düsseldorf Court of Appeals confirmed these findings. Leaving aside the question of whether the decision was, on the merits, correct or not, it is difficult to be convinced by the main arguments of the Court:

Firstly, it is conceded that competition, to be “effective”, needs to limit the scope of action of market participants. The Court then goes on to state that secondly, even if there was alternative transportation, i.e. pipe-to-pipe competition, this may not speak in favour of effective competition where free capacity on the network is scarce. It is submitted, contrary to the Court’s reasoning, that the level of free capacity as such cannot be regarded as a benchmark for effectiveness of competition. In particular, a high ratio of booked versus free capacity is no indication of lack of competition, but rather evidence of efficient use of the network. Ample unused capacity would, rather, imply that supply is not adjusted to actual demand and that the grid operator likely overcharges transportation customers, since fixed costs need to be borne by the existing customer base. A high degree of network capacity utilisation may thus even speak in favour of efficient pricing in itself.

Thirdly, the Court stresses that existing capacities were heavily booked by the incumbent’s sales branch, i.e. the grid operator’s sister company. The Court confirms the Agency’s view that the vertical integration of the grid operator speaks against effective competition because the sales branch would refrain from switching its demand for transportation to a competing pipeline, since it was not economically reasonable to pay grid fees to a third party provider instead of paying to the incumbent’s grid company. At the same time, the amount of fees to be paid “internally” would not matter. The Court goes on to say that, therefore, even free capacity in a competing pipeline does not change the picture, because the incumbent’s sales branch cannot be expected to step out of line.

17 See C. M. Bender, G. Götz, B. Pakula: Effective Competition: What it means why it is relevant for network industries, in: Intereconomics, Vol. 46, No. 1, 2011, pp. 4-10, argue that the relevant market from an economic point of view is often incongruent with the technologically driven industry definition.
20 OLG Düsseldorf, Ontras-VNG./Bundesnetzagentur, Decision of 13 January 2010, VI-3 Kart 74/08 (V). Further cases had been brought by Statoil Deutschland Transport, Thyssengas, DONG, ENI Gas Transport, Erdgas Münster Transport, Gasunie Deutschland, GRTgaz, WINGAS, vs. Federal Network Agency.
It is somewhat difficult to follow this reasoning: there is no cogent assessment that the incumbent’s switching was necessary to support the competitiveness of a competing pipeline. The alleged lack of free capacity does not imply such necessity. If there is sufficient third party demand, this may create competitive pressure on the prices of competing infrastructures. If there is no sufficient third party demand, pricing should not be an issue in the first place. Why then, in the competition test, ask the incumbent to switch?

In the same case, the Court highlighted two further points which are relevant to the assessment of effective competition in a market. The first aspect concerns churn rates: in the court’s view, the grid operator failed to provide churn rates as an indicator for competitive pressure. Furthermore, the court stated that churn rates, in any event, would likely be minor given the large chunk of bookings made by the incumbent’s sales branch. Here, it is questionable that low churn rates indicate a lack of effective competition, for low churn rates may also result from fierce competition forcing companies to offer very competitive products. The second aspect concerns the concept of contestable markets. The Court threw out a contestable-markets defence, stating that only an infrastructure in an advanced planning stage or which is actually erected may create competitive pressure; it further stated that the grid operator must show the specific impact of specific infrastructure in planning on the actual access pricing and must do so for the larger part of the network. This may be difficult or even impossible to prove, because pricing below a level that triggers entry may not necessarily relate to specific projects of competitors. Both aspects mentioned in this paragraph show that law-making and law-interpretation require economic input, especially in regulated industries.

Conclusion

The essential facilities doctrine is solidly integrated in competition law; however, doubts remain whether competition law really always provides what competition authorities claim it does.

Where courts are asked to decide whether the level of competition is satisfactory, recent cases show the limits of judicial review. Nevertheless, in the EU we do not see much of the reluctance prevailing in the USA, where the application of competition law is understood not to chill the conduct the law is designed to protect.

21 See C. M. Bender, G. Götz, B. Pakula, op. cit., pp. 4-10 on why potential competition is a relevant concept in this context.