

DISCUSSION PAPER SERIES

No. 3988

MARKET DEFINITION IN THE TELECOMS INDUSTRY

Jordi Gual

INDUSTRIAL ORGANIZATION



Centre for **E**conomic **P**olicy **R**esearch

www.cepr.org

Available online at:

www.cepr.org/pubs/dps/DP3988.asp

MARKET DEFINITION IN THE TELECOMS INDUSTRY

Jordi Gual, IESE, Barcelona and CEPR

Discussion Paper No. 3988
July 2003

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **INDUSTRIAL ORGANIZATION**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Jordi Gual

ABSTRACT

Market Definition in the Telecoms Industry*

Market definition for antitrust purposes is by now firmly rooted in economic analysis, both in the US and the EU, even if the approaches are slightly different. This Paper examines the theoretical basis for the legal definitions and assesses whether the general principles need to be adapted when dealing with the telecommunications services industry.

The Paper finds that the conventional antitrust methodology for market definition can be, to a large extent, readily applied to the telecoms industry but points out some key adjustments that have to be made to ensure that the antitrust and regulatory authorities define markets that capture adequately the nature of the competitive interaction in this industry.

Three adjustments are needed. First, careful consideration should be given not only to demand substitutabilities, but also to the competitive interaction with all potential suppliers. Second, demand complementarities and the economies of joint production should be properly recognized since they imply that in this industry bundles or systems of services may become a significant unit of antitrust and regulatory analysis. Third, market definition should take into account that the telecoms service industry is characterized by fixed and continuous sunk costs of service provision, and that this will often imply the need to revise the conventional concepts of market power and substitutability based on price elasticities.

JEL Classification: L40, L51 and L96

Keywords: competition policy, market definition, telecommunications and telecoms

Jordi Gual
IESE Business School
Avenida Pearson 21
08034 Barcelona
SPAIN
Tel: (34 93) 253 4200
Fax: (34 93) 253 4343
Email: gual@iese.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=109136

*Prepared for the European Commission (DGCOMP) under contract:
COMP/2001/7050/PSE/02. Comments by Professor Patrick Rey and Mr
Geoffrey Myers are gratefully acknowledged.

Submitted 03 July 2003

1. Introduction

Market definition for antitrust purposes is by now firmly rooted in economic analysis both in the US and the EU, even if the approaches are slightly different. This paper examines the theoretical basis for the legal definitions and assesses whether the general principles need to be adapted when dealing with the telecommunications services industry.

The paper finds that the conventional antitrust methodology for market definition can be, to a large extent, readily applied to the telecoms industry but points out some key adjustments that have to be made to this methodology to ensure that the antitrust and regulatory authorities end up defining markets which capture adequately the nature of the competitive interaction in this industry.

Three adjustments are needed. First, careful consideration should be given not only to demand substitutabilities, but also to the competitive interaction with all potential suppliers. Second, demand complementarities and the economies of joint production should be properly recognized since they imply that in this industry bundles or systems of services may become a significant unit of antitrust and regulatory analysis. Third, market definition should take into account that the telecoms service industry is characterized by fixed and continuous sunk costs of service provision, and that this will often imply the need to revise the conventional concepts of market power and substitutability based on price elasticities.

The paper also argues that if some competitive problems are detected with particular services (for example mobile call termination), that should not condition the market definition exercise. The definition of markets should be based on a detailed analysis of demand (both complementarities and substitutabilities) and the consideration

of all companies which have the assets and capabilities to satisfy these consumer needs. Such an exercise should be done first, and distinguished from the subsequent analysis of the competitive conditions in the markets defined as relevant.

The paper is organized as follows. Section 2 develops the general conventional framework. Section 3 looks at the issues specific to telecoms and analyses how the question of market definition in this industry is handled in the main antitrust jurisdictions. The paper considers first (section 3.1) how the conventional market definition framework is used to define the relevant product and geographic markets in the US and the EU, and how it is applied in particular to several common problems of market definition. For example, the distinction between fixed and mobile telephony, mass and business customers or wholesale and retail services. Some specific attention is also given to other telecoms industry features such as the importance of networks and bottlenecks. Section 3.2 analyses in detail the implications of complementarities in demand for the conventional market definition model, an issue which is central to the specific discussion on mobile call termination in section 3.3. Finally, section 3.4 considers the adjustments that are needed to the traditional model due to the fact that some segments of the telecoms industry are characterized by a rapid pace of service and technological innovation. Section 4 summarizes the main conclusions of the paper.

2. The economics of market definition

The definition of the relevant market for competition policy purposes is based upon the principle of the hypothetical monopoly. This concept is already well established in

antitrust legislation, both in the European Union and in the United States¹, and provides the standard framework for market definition analysis in competition policy cases.

The principle states that a product (or geographic) market should be defined as the minimum set of products (or areas) which could be successfully monopolized. That is to say, the range of products or services whose provision, if it were in the hands of a single firm, could profitably be restricted. The principle is also known as the SSNIP test, since it refers to a small but significant non-transitory increase in price by the hypothetical monopolist.

There are several features of this definition which are worth pointing out. The first is that if one of the key goals of antitrust analysis is to assess whether the relative position of one or more firms in a market gives them the power to raise prices, it makes sense to measure that position in the marketplace relative to an aggregate which –if fully controlled by a single entity– would provide a monopoly position and therefore full control over prices.

The second key feature is that the process of determining the relevant market is carried out through a gradual increase of the number of products or areas under consideration, starting from the smallest possible set. This again makes sense, since increasing the collection of products controlled by a firm or a group of firms acting jointly increases the potential for monopoly as alternative substitute products are eliminated by their gradual inclusion in the set under examination.

A third important feature is that the framework is explicitly designed for quantitative analysis. The profitability of higher prices refers to a five or ten percent mark-up, sustained for a period of about one year. Even if the quantitative information

needed to compute this is not available, the precise definition offers in practice a useful framework which guides the analysis of the data at hand.

A fourth remark is that the hypothetical monopolist principle requires a thought experiment which refers to the competitive benchmark. That is to say, the analyst should undertake a counterfactual experiment, trying to assess what would happen if all these goods were provided by a single firm. The goal is to determine whether the hypothetical monopolist would be able to profitably increase prices by 5% or 10% relative to those that would prevail in a situation of perfect competition. Since current prices may not correspond to those of a perfectly competitive market, the analysis will often imply a twofold counterfactual: assessing the pricing of both the hypothetical monopoly and the competitive regime.

Given the practical difficulties of undertaking a statistical analysis which follows rigorously the hypothetical monopolist framework, the concept is very often applied by considering informally, for a given group of products or areas, the forces which could restrain a firm from increasing prices if it were to control the sale of all the products under consideration. These forces are typically classified under the headings of demand and supply substitution.

Demand substitution refers to the products, services or geographical areas to which consumers could turn for a substitute of the monopolized good or collection of goods. If close substitutes are available, the ability of a potential monopoly to raise prices is going to be diminished. Note that this is a pure demand-side perspective. The goods to be considered could be delivered by a completely different production method. What matters is that they are reasonably good alternatives in consumption.

Supply substitution is meant to include all the producers which, if the hypothetical monopolist were to exercise its power by raising prices, would have the

capability –in the short run– to enter the market providing new output and thus limiting the ability of the monopoly to restrain production. Note that supply substitution means that the definition of the market should comprise all firms that can, without incurring sunk costs, be relevant competitors at short notice (i.e. firms that can easily rearrange their production facilities and serve some or all of the relevant goods). These potential entrants must be distinguished from those that could enter in the medium run (around two years) by deploying new capacity (and incurring sunk costs). In fact, in the U.S. Guidelines these producers are considered when assessing the relevance of entry barriers.

Informal assessment of the forces that would restrain the market power of a potential monopolist is a far superior method for determining the relevant market to relying on market boundaries based on production process similarities or marketing and industry practices. Nevertheless, it is an exercise fraught with difficulties, particularly because it will usually lead to an informal gathering of evidence about the substitutability of products in consumption. Data on choice patterns and reactions to changes in market conditions (prices, quality and availability) can be useful, but the lack of a proper statistical experiment will clutter any attempt to draw reliable inferences when assessing the behavior of consumers. For example, when we analyze substitution in consumer surveys, are we sure that all influences on consumer decisions other than price changes have been taken care of? The shortcomings of an informal approach have triggered the development of a more systematic framework, to which we now turn.

2.1. A formal approach to market power and the relevant market

Economic analysis provides a clear methodology that, under certain conditions, allows a rigorous determination of the degree of market power and can be used for market delineation purposes. This section summarizes the main principles for the simple case where the technology is characterized by constant returns to scale and the number of firms is given.

The measure of market power

Consider, first, a market where all output is supplied by a single firm i . Let the constant marginal cost be c_i and the market demand be expressed by $Q_i=D(p_i)$ where p_i is the price charged by the monopolist and Q_i the quantity sold. Let ϵ_M be the price elasticity of this demand function, which we will refer to as the market demand elasticity ($\epsilon_M = -(\delta D(p_i)/\delta p_i)/(Q_i/p_i)$). As is well known, the choice of the optimal price by the monopolist implies the following equilibrium relationship:

[1]

$$\frac{p_i - c_i}{p_i} = \frac{1}{\epsilon_M}$$

The left-hand-side of this equation, the relative mark-up also known as the Lerner index, provides a measure of market power to the extent that it assesses the ability of the firm to charge a price above marginal cost. The ability to increase price

will be inversely related to the elasticity of demand, with a larger mark-up when the market demand is less elastic (close to 1). For example, if the elasticity is 4, the equilibrium mark-up will be 0.25 (or 25%). If the market is perfectly competitive, the elasticity faced by the firm is extremely large, and the mark-up collapses to zero. A very low elasticity (say 1.5) leads to very large mark-ups (66%). The monopolist will never sell at a point where $\epsilon_M < 1$, since this is not consistent with profit maximization (revenues –and profits– can be increased by reducing output since the price increase more than compensates for the reduction in volume).

A benchmark of a 5% (or 10%) mark-up implies, in particular, that markets where the elasticity of demand is above 20 (or 10), are considered to be fairly competitive even if controlled by a single firm. Indeed, the high elasticity of demand is in fact capturing the existence of many alternative or substitute products. Note, also, that in a monopolized market, observed prices will already be above marginal costs, and by definition it will not be profitable for a firm to increase them more. The observed mark-up in equilibrium reflects already the market power which is being exercised.

Equation [1] provides a valid starting point which can be used for the formal definition of market power in situations where there is more than one provider.

Consider first the case of a market with a dominant firm which faces a fringe of competitive suppliers: a large number of firms which take prices as given and have an aggregate supply function defined by $Q=S(p)$. Let ϵ_S be the elasticity of this supply function ($\epsilon_S(p) = (\delta S(p)/\delta p)/(Q/p)$).

The dominant firm will choose its output taking into account the supply of the competitive fringe. The corresponding Lerner index can be shown to be²:

[2]

$$\frac{p_i - c_i}{p_i} = \frac{s_i}{e_M + e_S(1 - s_i)}$$

where s_i corresponds to the share of the market served by the dominant firm. It is interesting to note that equation [2] can also be derived by defining the problem of the dominant firm as one of profit maximization when facing a “residual demand function” computed as the difference between the market demand and the supply of the competitive fringe. From this point of view, the market power of the dominant firm will be measured (in the equilibrium) by the elasticity (ϵ_i^R) of its residual demand function.

That is:

[3]

$$\frac{p_i - c_i}{p_i} = \frac{1}{e_i^R}$$

where:

$$e_i^R = \frac{e_M + e_S(1 - s_i)}{s_i}$$

That is to say, from an empirical point of view, the extent of market power can be directly assessed by estimating the elasticity of the residual demand faced by the dominant firm. This elasticity captures the joint effect of the existence of substitute products (the elasticity of market demand), and the presence and importance of alternative providers (ϵ_S and s_i in [3]).

Consider next the case of a market where n oligopolists compete through the provision of slightly differentiated products. Let the (inverse) market demand system be³:

$$p_i = p_i(q_1, q_2, \dots, q_i, \dots, q_n) \quad (i=1, 2, \dots, n)$$

and define

$$m_i = -\frac{dp_i}{dq_i} \frac{q_i}{p_i} \quad m_{ij} = -\frac{dp_i}{dq_j} \frac{q_j}{p_i}$$

where $\mu_i > 0$ and $\mu_{ij} = 0$ indicating that the products are gross substitutes ($(\delta p_i / \delta q_j) = 0$, for all $j \neq i$).

With constant marginal costs ($c_i=1, \dots, n$), the equilibrium conditions of the oligopolistic competition model are defined by a set of n first-order conditions as follows⁴ (where we assume constant elasticities, following Baker and Bresnahan (1985)):

$$\frac{p_i - c_i}{p_i} = \mathbf{m}_i^R = \mathbf{m}_i + \sum_j \mathbf{m}_j \mathbf{r}_{ij}$$

The parameter ρ_{ij} captures the extent to which firms compete aggressively.

$$\mathbf{r}_{ji} = \frac{dq_j}{dq_i} \frac{q_i}{q_j}$$

That is to say, $(\delta q_j / \delta q_i)$ is the marginal change in the quantity of j when the quantity of i changes. The parameter ρ expresses this as an elasticity. A negative ρ corresponds to aggressive behavior (reacting to a quantity reduction –which would tend to raise price– with an expansion of output –which lowers price–) and a positive parameter reflects co-operation. The rivals react with output contractions when one firm reduces production.

Typically, we would expect that the elasticity of the (inverse) residual demand would be smaller⁵ than the own price elasticity: $\mu_i^R < \mu_i$. That is, the presence of strong substitutes (large μ_{ij} in absolute value) and aggressive competitors (large and negative ρ) limits market power and reduces the elasticity of the (inverse) residual demand (μ_i^R), which measures directly the mark-up.

From market power to market definition

The elasticity of the residual demand curve provides a convenient measure of market power which can be readily applied to the definition of markets for antitrust purposes. The hypothetical monopoly test with a 5% (or a 10%) mark-up can be applied with the use of an econometric estimate of the appropriate residual demand elasticity. The

procedure involves an iterative estimation process, beginning with the estimation of the residual demand at the level of one individual firm, and gradually enlarging the set of products under consideration if the null hypothesis, $\mu_i^R < 0.05$ for a 5% benchmark, cannot be rejected. The relevant market will correspond to the set of firms/products which find it profitable to sustain a mark-up larger than 5% (the group of firms/products which face a direct residual demand with elasticity below 20). If the 10% mark-up is used as a benchmark, the relevant elasticity will be 10⁶.

The framework implies, of course, that the firms/products which should be added first are the closest substitutes (large cross-price elasticity μ_{ij}) provided by very aggressive competitors (large and negative ρ). Removing this kind of firm as a competitor is most likely to enhance the market power of the merged (or co-ordinating) entity and is the appropriate way to establish the minimum set of firms which can (if acting jointly) monopolize the market.

2.2. Fixed costs and the relevant market

This formal approach to market definition runs into difficulties when we consider markets characterized by the presence of fixed costs. The previous model is based upon the definition of market power as pricing over marginal cost, because marginal cost pricing is taken to be the competitive benchmark. Mark-ups above 5% or 10% would correspond to a non-competitive market and result in abnormally high profitability. Yet, under conditions of increasing returns to scale, a positive mark-up (possibly above 10%) may in fact be necessary if a firm is to be viable, and the same may hold at the industry

level. What this implies is that under increasing returns the SSNIP benchmark could lead to excessively narrow markets. As the set of products is enlarged, the potential mark-up grows, but it does not make sense to stop at 5% or 10%, since under increasing returns to scale these percentages could correspond to mark-ups which barely allow the industry to recover its fixed costs.

Markets with fixed costs are pervasive and, unfortunately, economic analysis provides much less guidance as to what the proper benchmark is in these situations. Ideally, the benchmark should correspond to the mark-up which generates sufficient revenues to cover the fixed costs without leading to excess profitability. Since this endogenous determination of the benchmark is likely to be difficult, one way to deal with the problem is to consider changes in the mark-up (say, 5% increases) above the pre-existing level. This boils down to assuming that current prices correspond to “competitive” levels, where “competitive” is defined now as the mark-up over marginal cost which allows the recovery of industry fixed costs. Indeed, both the European Commission (EC) Notice on the definition of the relevant market (paragraph 19) and the US Merger Guidelines (section 1.1) admit that current prices can be used as benchmarks, in the absence of indications of insufficient competition (EC) or coordinated interaction (US).

If the SSNIP test is applied to a situation where current prices are above marginal costs, two potential sources of bias should be borne in mind when assessing the increase in the mark-up. Assume, for the moment, that the own and cross-price elasticities are constant. Since prices are kept above marginal cost, possibly thanks to a certain degree of co-ordination among firms, the assessment of the impact of the joint price-setting for a group of products results in a small increase in the mark-up. This is so because co-ordination had already led to a higher price level to begin with (in terms

of the formal framework, ρ_{ij} is positive). In this case, in general, adding the product will not increase the mark-up a lot, and the analyst will be led to consider incorrectly a wider range of products.

If the elasticities increase with prices (decrease with quantities), which is the common case, there will be an additional effect. Due to the high prices, the starting point will be a higher direct elasticity and this will imply that even small increases in the mark-up will generate significant substitution. This is the idea that at a higher price the degree of substitutability increases (the Cellophane fallacy) and so does the overall own-price elasticity⁷. This effect leads also to the incorrect establishment of excessively large markets.

The problems of applying the hypothetical monopoly test in industries where increasing returns to scale are significant do not end with the adjustments mentioned so far. Not only do we have to ensure that we choose the appropriate competitive benchmark, check the absence of excess profitability and take into account the potential upward biases in the definition of the market. It is also important to consider what is the exact source of the increasing returns and its expected impact on market structure.

The previous qualifications may be appropriate for industries where the increasing returns are the result of large non-sunk fixed costs. These are industries which can be well approximated by models of monopolistic competition or contestable markets. These qualifications could also be appropriate in industries that correspond to what John Sutton (1991) classifies as industries with **exogenous sunk costs**. That is, industries with technical scale economies where firms invest in sunk (non-recoverable) assets but where the magnitude of these investments is determined by the nature of the technology and not by the strategic interaction between competitors. In these industries,

one can think of the positive mark-up as the margin that is needed to pay (over the life of the assets) for the up-front investment costs.

However, these qualifications and the overall framework prove less useful in industries with **endogenous sunk costs**, where competition does not take place through prices or production capacities, but rather through continuous expenditures on non-recoverable advertising and R&D. These are dynamic industries where companies compete through product choice, product development or the introduction of new network products. William Baumol⁸ has recently used the term “continuing sunk costs” for these industries and argued that “you need new guidelines for even old-line industries with heavy but continuing sunk costs”, since these industries “cannot be expected to live up to marginal cost pricing”.

In these industries we not only have problems of identifying the appropriate “competitive” mark-up. More importantly, companies compete along new dimensions which are often much more important than price. In markets where competition is driven by innovation, where firms do not compete *in* the market but rather *for* the market, the use of methodological instruments such as the SSNIP test can be misleading and the concepts of market power and substitutability may have to be reinterpreted.

Indeed, there is a long tradition of analysis relating the extent of product market competition (the Lerner index which we have used so far as the measure of market power) to the rate of innovation. Recent research in this area points to an inverted U relationship (Aghion et al., 2002), showing that there is an optimal level of product market competition in terms of its impact on industry innovation and, therefore, dynamic efficiency. In other words, if there is too much product market competition, too little innovation takes place; but the same happens if product market competition is

too soft. For our purposes what this implies is that the mark-up (the static measure which assesses the degree of product market competition and is also used by Aghion et al. (op. cit.)) is not correlated with “dynamic efficiency”. It may be an adequate measure of static efficiency but it is a very poor guide in emerging industries where dynamic efficiency gains should be the key welfare criterion.

Since in these markets competition takes place through numerous dimensions other than price and the Lerner index is a poor proxy of efficiency, some authors⁹ have argued that the conventional antitrust tools of market definition and market power should be adapted to the specific features of these rapidly changing industries.

Much of the discussion has centred on the case of high-tech industries where firms compete through product innovation. Teece and Coleman (1998) argue that in these markets the SSNIP will tend to define markets which are too narrow. Since in these industries consumer demand is driven by a set of performance variables rather than price, without changes in performance there might not be substitutability across products even for price increases as large as 20%. What this literature suggests is that the analysis should focus on these new dimensions of competition (what Pleatsikas and Teece call product performance) and not on price. Some of the indicators that are suggested include: the depth of technology competition, the effect of product innovation on customer response, the extent of market share shifts over time, the extent of spending on R&D and of actual innovation, even if the same company out-innovates its rivals and does not lose the leadership position, etc. We will review later (section 3.4. below) the implications of these criticisms of the standard antitrust approach for market definition in the telecoms industry.

3. Defining markets in the telecoms industry

The formal model outlined in section 2.1 provides a rigorous set-up for the empirical determination of the relevant market in antitrust cases. It is certainly quite demanding in terms of the information available to the analysts, but nevertheless it has been applied to several specific industries by academic and professional economists¹⁰. The framework provides a robust technique, rooted in standard economic analysis, which can be deployed if sufficient data are available. If information is poor, the formal model constitutes a good conceptual apparatus to evaluate scattered data such as information on cross-price elasticities and consumer behavior (survey data).

When considering the market definition issue in an industry such as telecommunications, however, the formal framework can only shed partial light on the analysis. This is due to three facts. First, the telecommunications industry is characterized by the pervasive presence of fixed (sunk) costs. Second, the telecommunications industry comprises a wide array of very different services, ranging from conventional local or long distance voice telephony to high-speed internet access. For some of these services, in particular those which are provided with mature technologies, the framework will prove fairly adequate, once the fixed costs problem is properly taken into account. For others, however, the rapid pace of technological change means that quite often competition takes place through dimensions other than price (for example the introduction of new services with improved performance). In those instances, the framework of static oligopoly, which is the basis of the model sketched in section 2.1, may not be appropriate and we will have to adopt a broader perspective. This question is considered in section 3.4.

The third relevant fact is that telecommunication services are typically, although not necessarily, consumed in bundles. Moreover, they are often provided by multi-service firms in a joint production process where costs of stand-alone services may be hard to ascertain. We will consider the implications of these facts later (see sections 3.2 and 3.3). Before that, we will review in some detail how the standard framework can be applied in practice to the more conventional aspects of the telecoms industry.

3.1. Conventional market definition problems in telecoms

Many of the market definition questions raised in the telecoms sector can, in principle, be handled with a straightforward application of the conventional framework developed in section 2.1 of this paper. This approach could be used to assess the antitrust relevance of the telecommunications submarkets which are usually distinguished in the industry. For example, local and long distance calls, mass and business markets, fixed versus mobile calls, pre-paid versus contract mobile markets, etc.

Indeed, both the US and the EU recognize the concept of the SSNIP test as an appropriate instrument for the delineation of the relevant product and geographic markets. Nevertheless, the lack of sufficiently good data implies that in practice the establishment of antitrust markets is undertaken focusing on the concept of substitutability, with an informal use of the hypothetical monopoly idea. As we will see, the methodology used by the regulators in the two jurisdictions is substantially different, even if in practice the actual market delimitations are not that different.

Let us consider first the question of defining the **relevant product/service market**. In the US, the FCC introduced in 1996 (in the LEC In-Region Interexchange

Order¹¹, henceforth the LEC Order) a new methodology which purported to apply the 1992 Merger Guidelines. The goal was to define markets explicitly on the basis of demand substitutability, discarding previous procedures which took into consideration the substitutability in supply. The new principle was later on used in a set of landmark cases involving several mergers between major carriers and local operating companies (NYNEX/Bell Atlantic, MCI/WorldCom, Ameritech/SBC and Bell Atlantic/GTE).

As stated in the LEC Order, the methodology bases market definition explicitly on the hypothetical monopoly principle, focusing on the determination of substitute products from the demand side, and enlarging the product group until a large enough set of products is included so that market power could hypothetically be exercised. The LEC Order does not refer to a specific percentage, but it certainly discusses the absence of viable substitutes which could prevent a potential monopolist from increasing price¹². Following the Merger Guidelines, the LEC Order uses supply substitutability as an instrument to assess which companies have the incentive and the ability to enter the market if the price increase takes place¹³.

The FCC recognizes, however, that in practice the use of substitutability in demand could lead to the impractical definition of a very large number of markets. As a consequence, the agency argues that it is not necessary to consider all of these markets unless there is an indication of lack of competition in a particular individual service.

This means that, in practice, the agency resorts to the “similarity of competitive conditions” faced by different consumers as an aggregating device to define markets. For example, in the case of the definition of service markets, all long distance services (including, for example, individual services such as directory assistance, foreign

exchange service, discount plans, etc.¹⁴) are aggregated into one single service. The FCC asserts that “we need not delineate particular product markets (...) unless there is credible evidence suggesting that there is or there could be a lack of competitive performance with respect to a particular service or group of services”¹⁵.

However, a more robust aggregation procedure would be achieved with the explicit use of the extent of supply substitutability. Indeed, in its comments to the LEC Order, ATT pointed out¹⁶ that the FCC was using a rather narrow interpretation of the Merger Guidelines. According to ATT the Guidelines support an aggregate product market where “production substitution among a group of products is nearly universal among the firms selling one or more of the products” (footnote 14 of the Merger Guidelines). The same company remarks that there “is no difference between the facilities used to provide different services”.

In the same Order, in fact, the criterion of similarity of competitive conditions is used as the basis for the aggregation of services which had been previously defined as separate markets (IMTS and non-IMTS service¹⁷), but as Sprint points out¹⁸, the true reason one can aggregate those services is the fact that players in one market can very easily be suppliers in the other. That is to say, supply substitution.

The approach introduced in the LEC Order is used in subsequent decisions. For example, in the NYNEX/Bell Atlantic decision, markets for local exchange and exchange access and long distance are distinguished on the basis of the aggregation of product markets for which customers face the same competitive conditions¹⁹. It is interesting to note that this methodology allows the FCC to recognize as a relevant

market the joint provision, as a bundle, of local and long distance calls, on the grounds of consumers facing comparable competitive alternatives. We will return to bundling in section 3.2.

In the EU, the approach put forward in the market definition notice (European Commission, 1997) and used in many cases is based from a methodological standpoint on the SSNIP test. In practical applications it implies the simultaneous assessment of demand and supply substitutability as recently reasserted by the Guidelines on market analysis of the new regulatory framework (henceforth, the Guidelines)²⁰. For most of the cases, however, what drives the delineation of markets is demand substitutability, and no detailed market breakdown is attempted when the investigations of the Commission determine that the decision would not change by a more refined market definition. That is, when the competitive conditions would not be altered by a narrower definition of markets²¹. This is not very different from the practice of the FCC.

Consider next the case of **fixed versus mobile** calls and whether or not they belong to the same relevant market. The analysis of this question has increasingly focused on the degree of substitutability between the two services. This is exemplified clearly by the debate in the UK market, where the regulator has repeatedly argued, on the basis of survey data, that the two services are not close substitutes²². What is remarkable is that the quantitative information is reported without a clear sense of what degree of substitution is considered to be sufficient. Econometric evidence on substitutability has been presented by the operators to Ofcom, but has been rejected as

technically unreliable by the agency²³. In the end, Oftel bases its conclusions on its periodic surveys of consumer behavior.

Even if it is true that the econometric studies may be the subject of substantial technical controversy, the use of survey data is even more open to debate since that information can hardly constitute evidence about demand substitutability. Oftel has based its separation of the two markets on this type of evidence. It remarks, for example, that survey data show that only 11% of customers are willing to make less use of their mobiles in response to a substantial reduction of fixed line prices²⁴.

This type of analysis raises at least two issues. First, has the change in conditions been undertaken maintaining everything else constant? Second, as pointed out by several scholars, under certain conditions the number of marginal consumers needed to discipline pricing can be rather small. As argued by Hausman and others²⁵, in industries where prices have to be above marginal cost due to the existence of fixed costs, the elasticity of the direct residual demand will tend to be comparatively lower (the mark-up is higher) and a price increase will become non-profitable even if only a few clients change provider. This caveat shows very clearly that this type of survey evidence should not be a substitute for the analysis of substitutability based on a fully specified econometric model.

The standard approach therefore has a very important role to play in the determination of telecom markets. In fact, it has been used systematically in several prominent cases, even if not exactly in the way it was presented in section 2.1. For example, Jerry Hausman has estimated, for the case of the Sprint/MCIWorldCom²⁶ merger and the effects on the long distance market, the own-price and cross-price

elasticities corresponding to the main providers (ATT, MCIWorldcom, Sprint and an aggregate of non-branded suppliers) as well as the expected increase in the price of services provided by the new firm (which is equivalent to the computation of the residual demand of the merged firm)²⁷. The analysis of the cross-price elasticities allows a determination of how close substitutes firms are, and the computation of the elasticities of the residual demand function takes into account the pre-existing degree of rivalry between the competitors involved in the merger (the parameter which we denoted as ρ in section 2.1.).

Another recurrent market definition problem is the distinction between **mass and business markets** in conventional voice telephony. The broad definition of the market, including both segments, would make sense if the key characteristics of the services (quality, coverage, capacity, etc.) were similar and/or if the providers present in one market segment, but not in the other, could easily deploy capacity and commercial presence in the neighbouring market segment. In practice, these two markets have very often been considered as separate (e.g. the MCIWorldcom cases), but not because the framework of the hypothetical monopolist has been systematically applied.

In the US, the two markets can in principle be distinguished on the basis of aggregating users which face similar competitive conditions²⁸, but the imperfect substitutability between the different services offered to households and companies was already highlighted as important in the NYNEX/Bell Atlantic case²⁹. This is, in fact, fully consistent with the FCC approach, based on disregarding supply substitutability. However, in a later decision, the MCI/Worldcom case, the FCC Order highlights as relevant the fact that the assets owned by producers have to be different if they want to

tackle the two different sorts of users. In the Ameritech/SBC case, the FCC goes on to add not only different assets but also different firm capabilities. It is clear, therefore, that this implies a non-explicit consideration of supply substitutability. This is all the more contradictory if one notes the contents of more recent decisions such as BA/GTE³⁰. In that Order the FCC asserts that “defining relevant markets involves identifying and aggregating consumers with similar demand patterns”, and there is no mention of how important it is whether those different consumer groups can be served by the same group of companies or not.

In the EU, the distinction between the two markets is based upon the existence of significant differences in demand “and supply”³¹. The EC decisions stress, however, that the two markets are characterized by substantial differences in the demand of both types of customers, although some reference is also made to differences in supply.

Finally, the recent Guidelines consider also the ability to discriminate as reason enough to have two markets³². That is, even if the two markets are served by the same producers, they could be different markets on account of price discrimination. Of course, the ability to price-discriminate requires that the hypothetical monopolist be able to distinguish the two market segments and avoid arbitrage, and this has to be shown for the two markets to be separated³³. As the European Commission rightly stresses³⁴, price differences are not enough, since they can reflect quality differences and need not lead to the consideration of separate markets.

With regard to the **geographic** determination of the market, both the US and the EU apply in principle a methodological approach based on the SSNIP test. However, as

with the definition of the relevant product market, the FCC focuses on demand substitutability. This leads to markets that should be narrowly defined (on a point-to-point basis) on the grounds of the limited substitutability between calls with different originations and destinations. Since this approach could yield a very large number of markets, aggregation proceeds by considering jointly all consumers that face a comparatively similar competitive situation. Even if long distance calls out of Miami and Los Angeles may be considered as different services, the fact that the supply conditions are similar allows the agency to consider an aggregate US market for long distance calls³⁵: “We conclude that when a group of point-to-point markets exhibit sufficiently similar competitive characteristics (i.e. market structure), we will examine the group of markets using aggregate data”³⁶. As argued before, this is in fact an indirect way to consider supply substitutability. This is exemplified when considering the markets for long distance calls, where a distinction is made between in-region and out-of-region calls. These two are considered to be different markets³⁷ on the grounds that, for any particular region, the control of the local loop by the local Baby Bell leads to potential competition problems for in-region calls. It is clear that the same conclusion would be achieved if supply substitutability was explicitly considered. Given the ‘de facto’ monopoly of the local loop, there is no supply alternative to the incumbent for in-region calls.

The EC Guidelines approach the issue of geographic determination by grouping all areas where “competitive conditions are similar”³⁸. Consider, however, the decision

on the joint venture between BT and ATT³⁹. In this decision the Commission recognizes that country pairs could be the relevant markets for international carrier services, but sees no need to define any route in detail since it does not find any evidence of the “creation or strengthening of any dominant position in the area of international carrier services resulting directly from the proposed operation” (...) “the question of whether the provision of carrier services should be looked at in terms of country pairs routes or on a more global basis may be left open”⁴⁰. This is not what happens with international voice telephony services, where the US-UK route is defined as a relevant market on the basis of the non-existence of good substitutes (rerouting, calling cards, call-back) and the share of capacity owned by the players in the proposed operation⁴¹. This market is not defined using demand and supply substitutability but rather using an evaluation of market power even before the market has been defined. In practice, though, both approaches yield the same result: if prices were to go up, given the capacities installed, no other players would have the incentive and ability to provide service in that particular route.

The general approach discussed so far, which is in theory based on the SSNIP test but less so in practice, does not change significantly when we look at some of the more specific features of the telecoms industry, such as 1) the existence of vertical relations with a small number of upstream firms, 2) the presence of networks with varying degrees of capilarity and 3) the existence of bottlenecks.

A recurrent problem in telecommunications markets is whether a distinction should be drawn between **wholesale and retail** services. This applies to market

segments such as the origination of mobile calls, access to fixed voice telephony or internet access.

Wholesale markets comprise network services provided to service providers, and retail markets include the services aimed at end users. As usual, the extent to which these markets should be distinguished for competition policy purposes depends on the degree of substitutability. Demand-side substitutability is in principle not possible, since the two types of services are aimed at different group of users. However, the extent to which the wholesale and retail service are technically different is sometimes not clear (think, for example, of access to the public switched telephone network, PSTN). Even if we assume that from the demand side the two services are not substitutes, they could still be in the same market if there were ease of entry from one market into the other. This is usually the case for the retail market, where wholesalers can easily enter if there is a price rise. But not the other way around, because typically the provision of wholesale services requires building costly infrastructure and/or owning expensive operator's licences (entry, if possible at all, involves substantial sunk costs).

This methodological approach is the one adopted by Oftel in the case of the mobile market⁴². The UK agency distinguishes the two markets on the basis of the absence of demand substitutability and the asymmetry in supply substitutability. The key issue is, of course, the existence of very significant barriers to entry in the provision of wholesale services due to the limited number of mobile operators. Yet, if mobile virtual network operators are allowed a "regulated" access to the market, the distinction between wholesale and retail may cease to be relevant since the two types of services would have (potentially) the same competitors.

The EC has distinguished some wholesale services such as the lease of transmission capacity and the provision of related services⁴³, or the pan-European market for access (SMS) to mobile infrastructure⁴⁴. The EC argues that in fixed services there is a difference between subscriber (retail) access to infrastructure and operator (wholesale) access⁴⁵, but in the case of mobile services this is to be decided on a case by case basis⁴⁶. As argued above, regulation plays a key role in determining the conditions of entry in the provision of infrastructure (in fixed networks with measures such as the unbundling of the local loop and carrier preselection) and therefore it will be the key determinant if a distinction is to be drawn between wholesale and retail markets. Note finally that the wholesale/retail distinction in telecoms is different from its equivalent in other industries to the extent that in telecoms the question overlaps with the distinction between network and final services, and the number of networks is limited by technology and/or regulation.

As for **network** effects, which are so pervasive in the telecoms industry, they appear to be easily incorporated into the standard framework of analysis. Indeed, a network which becomes very large will tend to have fewer substitutes and may end up constituting a market in itself. In fact, this is the type of reasoning underlying the concerns expressed by the antitrust authorities with regard to the effect on the internet backbone market of the failed merger between MCI/Worldcom and Sprint. The US decision⁴⁷ refers to tier 1 internet backbone providers (IBPs), while the EU⁴⁸ talks about top-level or universal connectivity, but in both cases the discussion focuses on the insufficient pricing constraint that second and third-tier providers impose on the leading

IBPs. Insufficient demand and supply substitutability due to network effects leads to the definition of two separate markets.

Another distinctive feature of the telecommunications industry is the existence of **bottlenecks** that arise due to technology or regulation. For example, access to residential customers has been for many years a bottleneck in many countries, since the home was reached only by the PSTN and, given the existence of substantial (literally) sunk costs, it was not clear that it was economical to have more than one access to the home. Therefore, the local loop constituted a bottleneck.

By definition, a bottleneck is akin to an essential facility, a key resource for which there is no alternative. Access to this resource is usually regulated in order to prevent the potential abuse of market power by its owner⁴⁹. From the market definition viewpoint, of course, a bottleneck is a market on its own, since it has no good substitutes in use, and cannot be reproduced easily.

Changes in regulation and technology may, however, modify the boundaries of this market. For example, in the local loop the widespread availability of upgraded cable networks, fixed wireless access and satellite connectivity can lead to the disappearance of the bottleneck and to a wider definition of the market, embracing several technologies. This possibility has been explicitly recognized by the EC in its decisions in this area⁵⁰.

An interesting and controversial bottleneck arises in the mobile segment. In mobile telephony, as opposed to what happens in fixed telephony, there are usually several alternative networks. Access to the network does not, therefore, constitute a

bottleneck, since the user can choose between several alternatives and the fixed (and sunk) costs of access for the user are low.

Yet, in many mobile markets the way service payments are structured leads to the appearance of “a bottleneck problem” in call termination. In countries where calls are paid only by the calling party (the “calling party pays”, or CPP, principle) the receiving party will have little incentive to change providers when the price of finalizing calls goes up. The only way⁵¹ to reach a particular end user is to call and finalize the call on the network to which that user is subscribed. If the end user were paying for receiving the call, it would react to a potential abuse of monopoly power by changing providers, but the fact that it does not pay for that service generates an inelastic demand and a result akin to a bottleneck.

From the point of view of market definition, it is clear that under these conditions call termination on each network can easily become a market on its own and of course, each network operator will have a 100% market share of its own network. It is obvious, however, that these bottlenecks –which lead to very narrow definitions of markets– are the result of the CPP principle and the subsequent price insensitivity. As bottlenecks they may require regulation, but it is less clear that they can be properly defined as relevant markets. As we will see later, the termination of calls is a service which is usually acquired together with a set of other mobile services (call origination, SMS, etc.). The relevant market for antitrust purposes could well be a bundle of mobile services (see section 3.2 below) and this need not be incompatible with a regulated termination rate.

The EU Guidelines allow the potential definition of two separate mobile markets, one for call origination and one for call termination⁵², and recognize the problems of low incentives to compete on prices for terminating traffic⁵³.

Oftel has gone further than that, and in fact concluded that call termination on each mobile network is a relevant market where price regulation is needed due to the existence of insufficient competitive pressures⁵⁴. We will take up this argument in detail later (section 3.3).

3.2. Market definition and bundles of services

Telecommunication services, like many other goods and services, can be defined in a very narrow way. This involves decomposing the overall, broader service which is more directly perceived by the user into a collection of components, considering individually each service which could conceivably be consumed or supplied independently. Some complex goods such as automobiles can also be assessed in this fashion. It is certainly the case that one can think of a car as a bundle of components, including diverse items such as tyres, seats, rear-view mirrors and the like. Similarly, one can think of fixed voice telephony as an overall service, which includes a collection of narrowly defined component services such as access, call termination, local call origination, long distance call origination, directory services, etc. These component services may sometimes not be consumed independently (for example, no consumer wishes to have access to the network without using it, much as there is no independent demand for steering wheels), and they will have varying degrees of complementarity (very strong between access and

use, more moderate between long distance and local call origination). How fine a distinction should be made of the different component services becomes a difficult decision for antitrust analysis.

Moreover, if one takes a broad approach to product/service definition, the problems do not disappear. As with automobiles, it is not clear what the client perceives as “the service or product”. Does it include “extras” such as air conditioning or fog lights? Are these components part of a car, or should they be considered as different products?

Let us consider first an approach which starts with a narrow definition of telecommunication services, assuming that a distinguishable service exists if it can be supplied by a firm independently (even if the firm has to acquire complementary or input services from other firms to be able to provide that particular service). This approach leads to the determination of very narrowly defined services. For example, in mobile we could have markets such as access, origination of calls, termination, SMS, roaming, etc. Conceivably, a consumer can have different providers for all these services and the operators could, in principle, provide a few or only one of these services. There are, of course, significant complementarities in demand and scope economies in supply which will influence the actual choices of consumers and service providers, but it is instructive to discuss briefly the antitrust implications of this narrow approach to service definition.

Consider a narrowly defined service such as retail call termination in the mobile industry. The first thing to note is that the existence of strong complementarities with other mobile services reduces the own-price elasticity for this individual service⁵⁵, i.e. a price increase for this service will not diminish demand significantly if the prices of

complementary services do not change. In other words, the existence of complementarities increases the costs of switching for an individual service. Moreover, the extent to which the consumer will change to an alternative provider for call termination will be limited by the reduced number of providers of call termination (we assume here that there is no mandated call-by-call or carrier preselection for this service) and the fact that changing providers may also lead –due to the same complementarities– to a change in the provision of the other services.

In this context, the conventional method of market definition does not work. Starting from one individual product and enlarging the product set will yield unreasonable results. First, the own-price elasticity for a given service such as call termination will typically be very low due to the existence of complements and the limited number of substitutes. Call termination by a single provider may then be in itself a relevant market! Second, if the collection of services is enlarged to include complementary services such as access or call origination, we will usually find that the residual demand faced by the firm becomes more elastic, reflecting the loss of market power⁵⁶. That is to say, starting from an individual service, market power increases if comparable services provided by other operators are included, but declines if what is added are complementary services.

It is, therefore, apparent that if complementarity effects are important (as well as scope economies from the supply side), the standard market definition approach will make sense only when applied to a system of services. Indeed, we can say that a bundle of goods will constitute a system if the goods are complements⁵⁷.

If firms indeed produce jointly these services and consumers buy them as a system, the analysis of observed mark-ups can be a very misleading indicator of market

power and, as a consequence, lead to very poor decisions in terms of defining markets. Firms producing a collection of goods in the presence of fixed costs will optimally charge higher mark-ups in those services which face more inelastic demand. A high mark-up in an individual service need not reflect an overall high level of market power⁵⁸.

As shown in Table 1, the available evidence with regard to the complementarity of telecoms services is very significant (see figures in bold), in particular with regard to pairs of services such as access (subscription) and use, whether in mobile or in fixed telephony.

Table 1. Complementarities in telecommunications

Mobile telephony

Terminating mobile calls [Access Economics, 1999]*

own-price -0.1

Mobile subscription [dotecon, 2001]

own-price -0.37

cross-price mobile calls **-0.25**

cross-price fixed to mobile **-0.21**

Mobile calls [dotecon, 2001]

own-price -0.62

cross-price mobile subscriptions **-0.48**

cross-price fixed to mobile **-0.27**

Originating mobile calls [Access Economics, 1999]*

own-price -0.8

Originating mobile calls [Hausman, 1999]

own-price -0.5/ -0.6

Fixed telephony

Basic Access Service

[Hausman, Tardiff and Belifante, 1993]

own-price -0.005

cross-price intraLATA **-0.0086**

cross-price interstate **-0.0055**

* Quoted by Hausman, 1999.

In practice, this close complementarity and the fact that services such as access are not typically consumed independently means that the service definition will require the determination of a system. Of course, not all the services that can be supplied over a network need be components of that system. Services may be considered as stand-alone services when the complementarity with the rest of the components of the system is low⁵⁹ and/or when there are alternative (independently supplied) substitute services (this in turn will be affected by the importance of scope economies on the supply side and other market characteristics). A good example of a separate service market is international roaming, where the service is supplied as an option by the operators and users can choose between their own operator and alternative roaming-only service providers⁶⁰.

It is therefore clear that the strong complementarities in demand and the characteristics of supply could justify the analysis of market definition at the level of bundles of services. This is, in fact, something which is well established in antitrust practice. For example, in *US vs. Philadelphia National Bank* the Court defines the product market as “the cluster of products (various kinds of credit) and services (such as checking accounts and trust administration) denoted by the term ‘commercial banking’”. In the *NYNEX/Bell Atlantic*⁶¹ case the FCC indicates that “to the extent that consumer demand for bundled service packages forces carriers to offer such bundles, the bundling of local exchange and exchange access services with long distance services may well become a relevant product market”.

The bundling concept is also acknowledged by Oftel for the case of mobiles, even if half-heartedly. In the Mobile Effective Competition Review (Oftel, 2001 b) the operators argued that access is not a separate retail market since users would not wish to purchase the ability to make calls without actually making them. Oftel accepts that these two services are consumed jointly⁶² but dismisses the importance of this fact in terms of market definition since “theoretically for access there are no substitutes on the demand and supply side, so a hypothetical monopolist could feasibly raise the price of access”. Oftel bases its argument (and its market definition) not on actual market practice but rather on a hypothetical situation: “innovations such as indirect access would suggest that calls *could* be supplied by a number of suppliers, not just the access supplier” (own italics).

3.3. The debate on call termination to mobiles.

The difficulties posed by market definition in the telecommunications industry are exemplified by the on-going debate about mobile call termination, particularly in the UK. Oftel reviewed in 2001 the UK regulation of call termination. It issued a consultative document in February and a final statement (September) after assessing the views of the different parties. The UK regulator ended up proposing a CPI-12% price cap on termination charges for the next four years. Since this determination was not accepted by the mobile operators, the issue was addressed to the Competition Commission, which has to decide before early 2003.

The Oftel approach can be summarized as follows. Oftel starts by highlighting that the CPP principle creates few incentives for price competition between service

providers, since customers do not directly pay for the service that they use. Oftel next analyses the importance of several potential offsetting factors such as the significance of closed user groups, the possibility of switching networks, substitution for other type of calls and the countervailing power of the calling network. The agency finds⁶³ that these factors are ineffective in terms of leading to downward pressures on the pricing of termination.

Oftel compares also the rates for call termination with its estimates of long-run marginal costs for this service and finds them to be substantially above, with a large mark-up of up to 70%. The industry claims that the high prices for call termination are compensated by the low prices for call origination and other services, but Oftel dismisses this argument, arguing that these cross-subsidies are neither efficient nor fair and that the broadly defined mobile market is still not effectively competitive and therefore will not ensure that high termination charges will be compensated by low prices elsewhere⁶⁴.

After assessing the non-existence of sufficient price competition in call termination, Oftel goes on to delineate the relevant markets and concludes that –due to insufficient demand and supply substitutability– each mobile network constitutes a relevant antitrust market.

The analytical approach of Oftel, however, does not follow the usual antitrust methodology. The UK agency assesses first the degree of competition in call termination, and then moves on to determine the relevant markets. Yet the procedure should be precisely the reverse. Start with an analysis which determines the relevant markets, and then proceed, for each of those markets, to an examination of the

competitive conditions and the potential for abuse of market power. If the analysis of the mobile market is done in this fashion the results may be strikingly different.

First, an analysis of the nature of demand relationships, without mixing it up with competitive conditions, brings forcefully to the fore the importance of demand complementarities between individual services. This has, in fact, important implications for the study of the overall effect of an increase in call termination prices.

Indeed, if termination charges go up, there is an indirect but important effect on the called party, even if it does not suffer the price increase directly. When the termination charge increases, the costs of the calling network go up and this will prompt an increase in its own termination charges⁶⁵, thus leading to an increase in the price of the outgoing calls of the called party.

Oftel's analysis dismisses this effect because substitutability is assessed under the assumption of "all other prices constant". However, as argued in section 2.1 of this paper, a complete analysis of supply substitutability should take into account the (price) reaction of the rivals. This is further reinforced in this case because the complementarities between the services imply that it will not make sense for any of the providers to consider what happens to the demand for one particular service when other prices are kept constant. If the provider controls those prices, the proper benchmark is to consider how this firm is going to adjust them accordingly (like the Cournot monopolist of complementary products) and we have to consider the reaction of demand to the whole new set of prices determined by the monopolist –given the strategies of its competitors. That Oftel's analysis is far from this framework is clear in subsequent documents issued by the regulator. For example, in a recent paper entitled "Oftel's further documents on market definition" the agency stresses that "the prices of other

services are held constant” and it argues that the companies have “not identified termination and retail services as either demand or supply substitutes”. This is not surprising, since they are complements⁶⁶.

To conclude, it is worth pointing out that the whole question of market definition in mobile call termination should not be mixed up with the issue of whether a regulation of termination charges is required. Even if the market is broadly defined as the bundle of mobile services offered by operators and jointly acquired and consumed by end users, there may still be scope for regulating one or more prices within that bundle. This could be so under several circumstances.

First, one could argue that the overall level of competition in the mobile industry is too low, so that overall mark-ups are excessive, although in an industry with high fixed and sunk costs some degree of mark-up is needed and the determination of excess profitability is bound to be complex. Since the CPP principle creates an artificial bottleneck for the calling party, one can argue that, as with other essential facilities, access to this bottleneck has to be guaranteed to all market players under fair conditions. And furthermore, one could argue that the small number of players and the reciprocal nature of termination charges make it natural for operators to use these charges as “facilitating mechanisms” for non-cooperative price co-ordination (see Armstrong 1998 and Laffont et al. 1998).

All these are serious reasons for possibly regulating termination charges, without necessarily requiring the establishment of termination in each network as a relevant antitrust market⁶⁷. In fact, the Competition Commission of Australia has regulated termination charges, while recognizing both the relevance of the pricing

interdependencies from the point of view of mobile operators and a broader definition of the market⁶⁸.

3.4. Market definition in rapidly changing markets

As discussed previously, section 2.2, the SSNIP test is less useful in industries where fixed costs are significant, specially if these costs are sunk and related to strategic competition in service or product innovation.

We also stressed above that in dynamic industries the mark-up becomes a misleading tool for measuring the efficiency of the market and that competition should be assessed across several performance dimensions.

As we have seen, in actual practice regulatory agencies and competition authorities use the SSNIP test as a reference concept and base their analysis on the consideration of demand substitutability and to a lesser degree, substitution in supply.

In dynamic markets, the concept of substitutability has to be interpreted broadly, since substitution will take place not only on the basis of price differences but also depending on other relative performance indicators. From the demand side this often means taking a forward-looking view of the substitutability between different **technologies**, considering how different technical solutions may end up satisfying comparable consumer needs. From the supply side, a broad perspective implies that the analysis should take into account a larger range of potential **suppliers** of the services.

The focus on the substitutability between **technologies** is an approach similar to the one suggested by authors such as Teece and Coleman (1998) or Evans and

Schmalensee (2001) when they claim that in dynamic high-tech markets market delineation should basically assess “technology” competition.

To a certain degree, this is taken into account by the EC approach to market definition in telecoms. For example, the recent Guidelines on market analysis recognize the importance of technological convergence⁶⁹ and consider the extent to which the availability of alternative technological solutions may diminish the control of the local loop by the incumbent PSTN operators⁷⁰. In the Sprint/MCIWorldCom case, for example, the EC considered in detail new technologies (caching, mirroring, etc.) that could reduce the importance of the first-tier internet backbone. Regulators like Ofcom have also recently examined how technological advances may reduce the importance of the call termination bottleneck⁷¹. In all these cases, the potential technologies are at least carefully considered, even if the final decision has almost always involved the definition of a fairly narrow market.

As for a broad approach to potential **suppliers**, the experience is less encouraging, since the perspective of the regulators has been usually quite restrictive. In rapidly changing markets, which are being created by firms exploiting first mover advantages with new technologies or network effects, it is not surprising that firms enjoy very large market shares, since markets are in their infancy. In these contexts, one can argue that the agencies should adopt a broad view of potential suppliers, taking into account, first, all firms that own assets which could lead to the development of substitute technologies and second, a longer time horizon to assess the potential supply response⁷².

⁶⁹ See the Guidelines, # 47.

⁷⁰ See the Guidelines, footnote 62.
Technological Change”, Ofcom 24 May 2002.

This forward-looking view of markets is usually missing in the definition of markets in the telecoms industry. Consider, for example, the Sprint/MCIWorldCom case. In the rapidly changing backbone market, the merging firms enjoyed a large market share, but one that was predicted to deteriorate rapidly in the forthcoming years⁷³. This was not taken into account by the antitrust authorities in the establishment of the relevant market.

The lack of a forward-looking approach may also lead to a different sort of problem. As new markets are being created, some firms lead the way with the introduction of innovative services. Early recognition by the regulator or the antitrust authority of this type of markets is bound to detect competition problems since, by definition, the firms that master-mind the creation of a new market, start by enjoying very large market shares. In the limit, an innovative firm will have a 100% market share when it introduces a successful new service category.

In the EU, the Commission has recognized very quickly the existence of new relevant markets, such as the market for global broadband data communications services⁷⁴, the pan-European market for wholesale access to mobile infrastructure⁷⁵ and the emerging EU-wide seamless retail mobile market⁷⁶. The problem also arises in the US. If we consider the approach set forth in the LEC Order, it implies the absence of regulation of individual services unless the agency suspects lack of competition. Again, with a 100% market share, the suspicion of competition problems is likely to arise very soon.

In sum, the regulatory authorities have tended to be fairly conservative when establishing relevant markets in the emerging segments of the telecoms industry. In case

of uncertainty, they have preferred to err on the side of the early (and as a consequence, narrow) definition of new markets for antitrust and regulatory purposes. This is a pity, since in this type of industries there are sound reasons to believe that such an approach may easily lead to the imposition of undue restrictions on the activity of innovative companies and may not promote economic efficiency in the longer term.

4. Conclusions

The definition of markets in the telecommunications industry for antitrust purposes is based, in the main jurisdictions, upon the principle of the “hypothetical monopolist”. Both in the US and in the EU (as well as in other areas such as the UK and Australia), the framework is also used when delineating markets with regulatory objectives.

The “hypothetical monopolist” framework, or the small but significant non-transitory increase in price (SSNIP) test, can be closely linked to a well grounded model of oligopolistic competition. It provides thus a rigorous statistical methodology for the empirical determination of markets. This methodology, however, requires a large amount of data and in practice, it is most useful as a conceptual framework which guides the analysis of the available quantitative and qualitative information.

When applied to the telecommunications industry, the framework sheds useful light on some of the issues that clutter market definition in this and related industries. Yet this paper argues that the standard approach needs to be carefully adapted along three basic dimensions.

The first concern is that supply substitutability should be fully incorporated in the market definition exercise. Giving full consideration to supply substitutability has two important implications. It means enlarging the relevant market to include all firms which have the assets and capabilities to provide, within a reasonable amount of time, a good substitute service to the one being considered. When defining the market in telecommunications, supply substitutability provides a powerful aggregation device, so that authorities do not have to analyze the competitive conditions for a very large number of individual services.

Full consideration of supply substitutability also means that when we undertake the thought experiment of considering substitute services that can impose competitive

constraints on the pricing of a particular service, not only do we have to consider first the best substitutes in use. We also have to take into account the types of firms that are supplying those services and whether they are currently more or less aggressive competitors. The more aggressive they are, the more likely it is that they belong to the same relevant market.

The second area where the conventional framework has to be adapted when applied to telecoms is the analysis of demand. It must be recognized that systems of services constitute a potentially relevant unit of market analysis. The importance of systems in this industry arises both from the nature of user demand and the characteristics of supply. The analysis should fully recognize the existence of very significant complementarities in demand, which imply that consumers demand many of the services in bundles (and in some cases there is, in fact, no independent demand for some particular services, such as access). Similarly, from the supply side, the importance of scope economies implies that the services can be more efficiently provided as joint products and this is the way they are offered to consumers by the operators.

The importance of complementarities and the technological advantages of joint production are bound to change over time and be influenced by regulatory changes and technological progress, but for each geographical market at any point in time a careful distinction must be made between those services which are linked by strong complementarities and provided and consumed as systems, and those which are consumed and produced as stand-alone services. In the mobile segment, for example, depending on the regulatory and technological situation, the core group of services (access, call termination and origination, SMS, etc.) can form a single system provided by all competing operators, while other services such as roaming can be separate

services that can be bought independently from a non-coincident set of suppliers. The fact that some competitive problems may be detected which affect a few individual services within a system (for example access or mobile call termination) should not obscure the analysis at the market definition stage, where the focus should be on the nature of consumer behavior and the existence of alternative sources of supply given what consumers demand.

The use of the traditional market definition framework in telecoms has to be adapted also because this is an industry characterized by the existence of significant fixed and sunk costs. This is relevant at least for two reasons. First, it implies that the analysis of market power –and as a consequence that of a hypothetical monopolist– cannot be based on the use of marginal cost pricing as a competitive benchmark. In industries with fixed costs, a positive mark-up is fully compatible with a competitive market and if this is not taken into account, markets may be defined too narrowly. Moreover, to the extent that the sunk costs incurred by telecommunications service providers are of a continuous nature, the result of competition in service innovation and technology, it becomes harder to apply a methodology based upon a static measure of competition such as the Lerner index. If telecoms firms compete through new services and better performance (and this may be the case in some segments of the industry), the conventional analysis of demand and supply substitutability based upon cross-price elasticities may become irrelevant and the analysis of the boundaries of the market should take into account several performance variables from the point of view of demand, as well as a broader set of potential firms engaged in the introduction of new services and technologies that could end up satisfying the same consumer needs.

Finally, the paper provides a discussion of the market definition problems associated with the existence of bottlenecks, with a particular focus on the mobile call

termination bottleneck caused by the “calling party pay” principle. This discussion is specially relevant in view of the new EU regulatory framework. The new EU approach is based upon the idea that regulation will be based on competition policy principles. That is to say, regulatory action will be justified only when there is insufficient competition. The absence of significant competition constraints will be assessed using well-established competition policy tools, the first of them being the definition of the relevant market.

This regulatory procedure poses, however, significant methodological problems, illustrated by the mobile call termination controversy. The widespread perception of insufficient competitive pressures in the provision of this service has led in some jurisdictions to the definition of a very narrow relevant market: call termination in each network. This market definition does not take into account some of the key characteristics of the telecoms industry (the complementarity between services) and runs counter to some of the main principles that should guide market definition (what services, or bundles of services, are in fact perceived by users as constituting viable alternatives in consumption and are being or could be offered by service providers). Nevertheless, the overriding objective of regulating call termination and the need to follow the conventional antitrust methods have led to the definition of call termination in each mobile network as a relevant –and by definition monopolized– market. There may be sound reasons to regulate a bottleneck such as mobile call termination. But these should not lead to a biased determination of the relevant market.

Notes

¹ See the Notice on the definition of the relevant market (European Commission, 1997) and the Merger Guidelines in the US (FTC, 1997).

² See Landes and Posner (1980) for a detailed exposition.

³ It is more convenient now to use this demand system rather than the direct system (see Baker and Bresnahan, 1985, for details).

⁴ Following Baker and Bresnahan (1985) we assume constant elasticities. These authors also show that the residual demand curve approach can encompass both quantity and price competition, as well as different degrees of non-cooperative and cooperative behaviour (op. cit. page 431).

⁵ The relationship between the corresponding direct demand elasticities is going to be reversed. That is, $\epsilon_i^R > \epsilon_i$. The elasticity of the direct residual demand will usually be larger than the direct own-price elasticity (and, again, the markup smaller).

⁶ Strictly speaking, with more than one firm the analyst will have to compute a set of partial residual demands and add their elasticities (see Baker and Bresnahan, 1988). A residual demand curve faced by one firm has a slope which depends (in part) on how the customers of the firm defect to all other firms as prices are increased. The slope of a partial residual demand curve takes into account that some of the competing firms also raise prices.

⁷ See Landes and Posner (1981, page 961) and White (1999). As Posner (2001, page 156) points out “the fact that a further price increase though modest would cause substantial substitution either in consumption or production is not a reason to define the market more broadly”.

⁸ See Krueger, Alan B. (2001) page 220 “An Interview with William J. Baumol”

⁹ Teece and Coleman (1998), Pleatsikas and Teece (2001) and Evans and Schmalensee (2001).

¹⁰ See Baker and Bresnahan (1985, 1988), Scheffman and Spiller (1987) and Tollison et alia (1991).

¹¹ *Regulatory Treatment of LEC Provision of Interexchange Services Originating in the LEC's Local Exchange Area*, Second Report & Order, CC Dkt. No. 96-61, FCC 97-142, 1997 WL 193831 (rel. April 18, 1997).

¹² See LEC Order, #28.

¹³ LEC Order # 28.

¹⁴ LEC Order # 36.

¹⁵ LEC Order # 42.

¹⁶ LEC Order # 34.

¹⁷ LEC Order # 55.

¹⁸ LEC Order # 53.

¹⁹ Case NYNEX/Bell Atlantic Docket 97-286, # 50.

²⁰ See European Commission (2002).

²¹ See, for example, the Case N° IV/JV.15 BT/AT&T, #73 and 79.

²² See the Oftel positions as presented in Oftel (2001 b) and Oftel (2002 a). Several other documents in the Oftel web site reinforce this view.

²³ See the paper by Dotecon (2001) and the comments by Hunter and Ioannidis (2002), both available in the Oftel web site.

²⁴ See Oftel 2001 b, # A.1.9 to A.1.11.

²⁵ See Hausman et al. (1996), pages 368-369.

²⁶ Declaration of Professor J. Hausman on behalf of SBC in the case Sprint/MCIWorldCom. Docket # 99-333.

²⁷ See Hausman's declaration in the Sprint/MCIWorldCom case.

²⁸ LEC Order # 42.

²⁹ Case NYNEX/Bell Atlantic, FCC 97-286, # 53.

³⁰ Case BA/GTE, FCC 00-221, # 102.

³¹ See footnote 58 in the Guidelines, and the cases cited therein.

³² See the Guidelines, #46.

³³ See Hausman et al. (1996).

³⁴ See the Guidelines, #46.

³⁵ LEC Order # 64 to 66.

³⁶ Op. cit. #66.

³⁷ Op. cit. #76.

³⁸ See the Guidelines, #55 to 60.

³⁹ Case N° IV/JV.15 BT/AT&T.

⁴⁰ Op. cit. #78.

⁴¹ Case N° IV/JV.15 BT/AT&T, #83.

⁴² See Oftel (2001b, page 46).

⁴³ See the Guidelines, footnote 65.

⁴⁴ See the Guidelines, footnote 64.

⁴⁵ See the Guidelines # 65 and footnote 56.

⁴⁶ See the Guidelines, footnote 61.

⁴⁷ DoJ vs WorldCom and Sprint, # 28 to 30, June 26, 2000. Docket 99-333.

⁴⁸ See case COMP M. 1741 MCIWorldCom/Sprint #53. The EU distinguishes between tier 1 and the other backbone providers not only on the basis of network effects but also taking into account the nature of the relationships between firms (peering arrangements vs. customer relationships).

⁴⁹ See Laffont and Tirole (2000), page 97-98.

⁵⁰ See the Guidelines # 68 and footnote 62.

⁵¹ Of course, one possibility would be to introduce for call termination a regulation similar to carrier selection (call by call) or carrier preselection, whereby the caller is able to choose on which network to terminate the call, which is then rerouted to the network of the called party.

⁵² See the Guidelines # 69.

⁵³ See the Guidelines, footnote 69.

⁵⁴ Oftel (2001c and d).

⁵⁵ This is the case even in the absence of the CPP principle. Of course, with the CPP principle the price increase is born by the calling party and this only indirectly may lead to a price increase of the services paid for by the called party say.

⁵⁶ This is, of course, a variant of the result highlighted long ago by Cournot. A monopolist producing a range of complementary goods will tend to set prices lower than those that would be established by a collection of independent monopolists selling the same set of complementary goods.

⁵⁷ See, for example, Katz and Shapiro (1994).

⁵⁸ It could be, of course, that firms use a service such as termination which constitutes a cost that each firm imposes on each other and is not directly charged to customers, as a vehicle for collusion (see the discussion in section 3.4).

⁵⁹ The literature on aftermarkets (see Shapiro (1995)) has tackled a related issue: should durables and their associated complementary goods (photocopiers and parts in the famous Kodak case) be considered a single product? Similarly, Katz and Shapiro (1999) discuss complementarity in software and hardware markets. They conclude that strong complementarity warrants the definition of the system as the relevant market.

⁶⁰ See European Commission (2000) "On the Initial Findings of the Sector Inquiry into Mobile Roaming Charges", Brussels, Working Document, 13 December.

⁶¹ Case NYNEX/Bell Atlantic # 52.

⁶² Oftel (2001 b) A.1. #29.

⁶³ Oftel refers to quantitative evidence when in fact all the analysis is qualitative in nature. The agency uses data (survey evidence) but it does not attempt to measure the key variable: the extent of substitutability between different types of calls. (See Oftel 2001 b, Pages 7 to 9). Using survey evidence leads to an informal assessment of the importance of consumers switching demand under certain circumstances (for example Oftel 2001 d, Annex 2, page 58), but as we have seen (see above section 3.1.) this is a particularly troublesome approximation to elasticities of residual demand in industries with fixed costs.

⁶⁴ Oftel recognizes (see Oftel 2002 b, # 16 and 17) that to pay for the fixed costs, a positive markup will be required, but it is less clear what the markup level compatible with effective competition is. In any case, its criticisms of Ramsey prices miss an important point. Part of the inefficiencies created by the cross-subsidies of the operators may be the result of other regulations: i.e. the CPP principle, or the fact that the incumbent fixed operator is subject to regulation.

⁶⁵ This is the case, of course, for mobile networks, and not for fixed networks due to regulation.

⁶⁶ See Oftel (2002 b). The same document (#32) argues also that access and calls are not substitutes but does not recognize the complementarity between the two. Indeed, these services would be complements even if indirect mobile access was available. The analysis of Oftel mixes up the competitive conditions with the nature of the demand relationships between services.

⁶⁷ The overriding objective of regulating mobile call termination may undermine the analysis. For example, in Oftel (2001d, #3.6) the agency argues that demand for call termination is so inelastic, even below one, that revenues would fall if prices were lower. However, no single product monopolist would

ever set such a price. This highlights the fact that the competitors should be analyzed as oligopolists simultaneously setting the prices of a group of services which are complements in demand.

⁶⁸ See Australian Competition Commission (2001, pages 24 -31). Other regulators, such as OPTA in the Netherlands, follow an approach similar to Oftel.

⁶⁹ See the Guidelines, # 47.

⁷⁰ See the Guidelines, footnote 62.

⁷¹ “Oftel’s response to the Competition Commission’s letter of 7 May 2002 (Oftel 20) entitled ‘Technological Change’”, Oftel 24 May 2002.

¹³ See Plaetsikas and Teece (2001).

¹⁴ Supplemental Internet Submission to the MCIWorldCom/Sprint case, CC docket n° 99-333.

¹⁵ See the Guidelines, footnote 56.

¹⁶ See the Guidelines, footnote 64.

¹⁷ See the case COMP M.1975. Vodafone Airtouch Mannesman, # 12 to 21.

References

- Armstrong, M. (1998) "Network Interconnection," *Economic Journal*, 108, pp. 545-564.
- Australian Competition Commission (2001) "Pricing Methodology for the GSM Termination Service", Final Report, July.
- Baker, J. and T. F. Bresnahan (1985) "The Gains from Merger or Collusion in Product Differentiated Industries," *Journal of Industrial Economics*, vol. 33, pp. 427-444.
- Baker, J. and T. F. Bresnahan (1988) "Estimating the Residual Demand Curve Facing a Single Firm," *International Journal of Industrial Organization*, vol. 6, pp. 283-300.
- Dotecon (2001) "Ramsey Prices for Mobile Services", a report prepared for BTCCellnet, 4 May.
- Economides, N. and S. Salop (1992) "Competition and integration among complements, and network market structure," *Journal of Industrial Economics*, vol. 40, issue 1, March.
- European Commission (1997) *Commission Notice on the definition of the relevant market for the purpose of Community competition policy*, Official Journal: OJ C 372, 9/12/1997.
- European Commission (2002) *Commission guidelines on market analysis and the assessment of significant market power under the Community regulatory framework for electronic communications networks and services*, Official Journal: OJ C 165/03, 11/07/2002.

- Evans, D. and R. Schmalensee (2001) "Some economic aspects of antitrust analysis in dynamically competitive industries," NBER Working Paper 8268. May 2001.
- Federal Trade Commission (1997) Merger Guidelines.
- Hausman et al. (1993) "The Effects of the Break-up of AT&T on Telephone Penetration in the United States," American Economic Review Papers and Proceedings, May.
- Hausman et al. (1996) "Market definition under price discrimination," Antitrust Law Journal, 64, pp. 367-386.
- Hausman, J. (1999) Declaration in response to the paper by Prof. Gans (University of Melbourne) "An Evaluation of Regulatory Pricing Options for Mobile Termination Service," of 9th December 1999.
- Hunter, J. and C. Ioannidis (2001) "A review of Ramsey Prices for Mobile Services, produced for BTCellnet by Dotecon," 15 June.
- Katz, M. and C. Shapiro (1994) "Systems Competition and Network Effects," Journal of Economic Perspectives, vol. 18.
- Katz, M. and C. Shapiro (1999) "Antitrust in Software Markets," in Eisenach and Lenard, *Competition, Innovation and the Microsoft monopoly: antitrust in the digital marketplace*, The Progress and Freedom Foundation.
- Krueger, Alan B. (2001) "An Interview with William J. Baumol," Journal of Economic Perspectives, Summer.
- Laffont, J.-J., P. Rey, and J. Tirole (1998) "Network Competition I: Overview and Non-discriminatory Pricing," Rand Journal of Economics, 29, pp. 1-37.
- Laffont, J.-J. and J. Tirole (2000) *Competition in Telecommunications*, Munich Lectures in Economics, MIT Press, Cambridge Mass.

- Landes, W.M. and R. A. Posner (1981) "Market power in antitrust cases," *Harvard Law Review*, 94, issue 5.
- LECG (2000) *The economics of mobile call termination*, report prepared on behalf of BTCellnet, One2One, Orange and Vodafone.
- Neven, D., R. Nuttall, and P. Seabright (1993) *Merger in Daylight*, CEPR, London.
- Oftel (2001a) *Effective competition review: Mobile*, February.
- Oftel (2001b) *Effective competition review: Mobile*, September.
- Oftel (2001c) *Review of the Price Control on Calls to Mobiles*, February.
- Oftel (2001d) *Review of the Price Control on Calls to Mobiles*, September.
- Oftel (2002 a) *Protecting Consumers by Promoting Competition*, January.
- Oftel (2002 b) *Oftel's Further Comments on Market Definition*, 27 March.
- Pleatsikas, C. and D. Teece (2001) "The analysis of market definition and market power in the context of rapid innovation," *International Journal of Industrial Organization*, 19, pp. 665-693.
- Posner, R.A. (2000) "Antitrust in the new economy," in *Tech Law Journal* available at www.techlawjournal.com/atr/20000914posner.asp , September 14. Address at the Conference of the American Law Institute.
- Posner, R.A. (2001) *Antitrust Law*, The University of Chicago Press, Second Edition, Chicago.
- Scheffman, D.T. and P.T. Spiller (1987) "Geographic Market Definition under the U.S. Department of Justice Merger Guidelines," *Journal of Law and Economics*, vol. XXX, pp. 123-146, April.
- Shapiro, Carl (1995) "Aftermarkets and consumer welfare: making sense of Kodak," *Antitrust Law Journal*, Issue 2, winter.

- Slade, M.E. (1986) "Exogeneity tests of market boundaries applied to petroleum products," *The Journal of Industrial Economics*, vol. XXXIV, issue 3, pp. 291-303, March.
- Stigler, G. J. and R.A. Sherwin (1985) "The Extent of the Market," *Journal of Law and Economics*, vol. XXVIII, pp. 555-585, October.
- Sutton, J. (1991) *Sunk costs and market structure: price competition, advertising and the evolution of concentration*, Cambridge, Mass.: MIT Press.
- Teece, D.J. and M. Coleman (1998) "The meaning of monopoly: antitrust analysis in high-technology industries," in *The Antitrust Bulletin*, Fall-Winter.
- Tollison, R., D. Kaplan and R. Higgins (1991) *Competition in the Carbonated Soft Drink Industry*, Lexington, MA: Lexington Books.
- Teece, D. and M. Coleman (1998) "The meaning of monopoly: antitrust analysis in high technology industries," *Antitrust Bulletin* 43, pp. 801-857.
- U.S. Department of Justice (1997) "1992 Horizontal Merger Guidelines (with April 8, 1997, Revisions to section 4 on efficiencies)," Source: FTC web site.
- White, L.J.(1999) "Wanted: A Market Definition Paradigm for Monopolization Cases," in *The Computer Industry Newsletter*, Spring. Available at www.abanet.org/antitrust/committees/computer/spr99white.html