Haucap, Justus

Working Paper
The Economics of Mobile Telephone Regulation

Diskussionspapier / Helmut-Schmidt-Universität Hamburg, Fächergruppe Volkswirtschaftslehre, No. 4

Provided in Cooperation with:
Fächergruppe Volkswirtschaftslehre, Helmut-Schmidt-Universität (HSU)

Suggested Citation: Haucap, Justus (2003) : The Economics of Mobile Telephone Regulation, Diskussionspapier / Helmut-Schmidt-Universität Hamburg, Fächergruppe Volkswirtschaftslehre, No. 4, http://nbn-resolving.de/urn:nbn:de:gbv:705-opus-1326

This Version is available at:
http://hdl.handle.net/10419/23522

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes. You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public. If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
The Economics of Mobile Telephone Regulation

Justus Haucap
The Economics of Mobile Telephone Regulation

Justus Haucap*

March 2003

Keywords: Mobile telephony, Competition, Regulation, Interconnection, Call Termination, Number Portability, Roaming

JEL Classification: K21, L13, L51, L96

* Institute für Economic Policy, University of the Federal Armed Forces Hamburg, Holstenhofweg 85, D-22043 Hamburg, Germany. Fax: +49 40 6541 2042, email: justus.haucap@unibw-hamburg.de
Extended Abstract

This paper analyzes how competition works in mobile telecommunications markets and, bases on this analysis, we discuss whether regulatory intervention in mobile telephone markets is justified from an economic perspective. Starting point of our analysis is the observation that an evaluation of regulatory interventions into mobile telecommunications markets cannot be made without a deeper understanding for competitive processes in mobile telephony.

What is of decisive relevance for understanding competition in mobile telephony, is the fact that building a mobile telephone network requires highly specific investments, which take place under significant uncertainty, as investments in 3G networks such as UMTS illustrate. An inevitable consequence of specific investments are sunk costs. Hence, one can only expect firms to extensively invest and innovate if firms can hold a justified expectation to work profitably after they have invested. To cover their capital costs, which are largely fixed and not avoidable, firms need to follow a pricing policy that involves prices above incremental costs.

Hence, a key determinant for mobile operators' price policy lies in their cost structure, which is characterized by high fixed and common costs that are also sunk and relatively low incremental costs. In such situations, efficiency demands so-called Ramsey pricing structures, which involves different mark-ups for different services. In contrast, a situation with uniform mark-ups will generally be inefficient. Instead, services with an inelastic demand should carry relatively high prices, while services, for which the demand is rather elastic, should be priced close to marginal costs. Exactly such a pricing structure results when unregulated firms are left to maximize their profits. Hence, the factor that prices and mark-ups differ between different services and markets is an efficiency imperative and not a sign for market failure.

Nevertheless the necessity of interconnection and fixed-to-mobile termination may give rise to competition problems. As we argue in this paper, closer analysis shows that these problems do not automatically imply that sector specific regulation is warranted. The same hold for the question of regulated mobile number portability.

Instead, an ex post introduction of sector specific regulation can be regarded as a breach of the implicit regulatory contract by the State. This Government hold-up socializes and redistributes operators' profits, while the operators carried the initial investment risk. Such a Government hold-up reduces firms' incentives for investment and innovation and, thereby, also harms consumers in the long run. In addition, there is a real risk of regulatory failure, as empirical evidence demonstrates.

Based on these considerations, this paper fiercely advises against sector specific regulation of mobile telephone markets. The social welfare loss that would arise from such regulations are estimated to be enormous.
1. Introduction

The development of mobile telephony in Europe is generally considered a success story. During the last 15 years the mobile telephone industry has experienced a dramatic growth, which can be attributed to technological, regulatory and competitive factors (see Gruber and Verboven, 2001a, b). In some countries, the penetration rate now approaches 100 percent, and the number of mobile telephone customers now exceeds the number of fixed telephone lines in some regions (see OECD, 2001). Today, there are more than 80 times as many mobile telephone customers worldwide than there were in 1990; this corresponds to an average growth rate of more than 40 percent per year. Revenues from mobile telephony have also increased during this time by almost 30 percent per year on average.¹

This strong growth can be at least partly attributed to increased competition following further market entry (see OECD, 2000) and the dramatic price decrease that is related.² As exclusively using a mobile telephone is already less expensive than a fixed-line telephone connection for some consumers today (see OECD, 2001, page 181f.), intermodal competition between mobile and fixed-line telephone networks will also increase even further.³

Despite or maybe because of this extraordinarily positive development, mobile telephone markets have become more of a focus for political decision makers and regulatory authorities in recent times (see e.g. European Commission, 2002). Moreover, stricter market regulation of mobile telephony is also demanded by some market participants, especially those which compete without their own infrastructure.

This paper therefore addresses the question as to whether regulatory intervention in mobile telephone markets from an economic perspective is justified. Put differently, the question is whether regulation of mobile telephony is likely to lead to a welfare gain for society. In order to analyze this question the next section will discuss the economic foundations of competition in mobile telephone markets, before section 3 addresses the regulation of mobile telephone markets from a first principles perspective. Section 4 deals with some special competition problems that arise in mobile telephone markets. These special problems stem from the necessity for mobile network interconnection and for mobile call termination from fixed-line telephone networks. However, these issues do not necessarily warrant regulation. The same applies, as we will see, to questions regarding international and national roaming. Section 5 summarizes the paper and offers concluding remarks.
2. Competition in Mobile Telephone Markets

2.1 Price Competition with Fixed and Common Costs

In all markets, competition and the resulting prices strongly depend on the competing firms’ cost structures. For mobile telephone networks it is characteristic that the cost structure has high fixed and common costs and relatively low variable or incremental/marginal costs. The main part of a mobile telephone operator’s costs therefore does not vary with the number of participants, calls or connection minutes, but it is fixed and, furthermore, to a large extent also sunk. Sunk costs are the result of specific investments. Building a mobile telephone network and developing a brand name are largely specific investments, as these outlays cannot be recovered at all or only partly if an operator goes out of business. If these investments cannot be used as planned they have to be written off completely or at least to a large extent. Exactly such a cost structure with high sunk fixed and common costs and relatively low incremental costs is decisive for competition and price setting in mobile telephone markets.

If an industry exhibits such a cost structure with high fixed and common costs and relatively low incremental costs, then companies cannot survive if they set all prices at (short term) incremental or average variable cost. Instead prices have to incorporate a surcharge over and above the short run incremental cost level. Through this surcharge firms can generate contributions towards their fixed and common costs which allows them to operate without losses and to survive in the market. If, however, a firm does not expect to operate without losses in the long run, it will not enter the market in the first place.

Given this, it is less of a question whether prices for certain mobile telephone services are above variable or short run incremental cost, but more (1) how high the mark-up on top of these costs is and should be and (2) whether the resulting price structure for the various services is efficient. The second question – what an efficient price structure is when a firm has to cover its fixed and common costs – has already been analyzed in the 1950s. Based on work by Frank Ramsey (1927) on tax theory, Marcel Boiteux (1956) examined just this very question: Which price structure is efficient when a firm supplies many different services the production of which has common costs, and when the firm should be able to fully recover its costs at the same time? The now well known answer is that the price for a product or service should be the higher, the less sensitive demand reacts to price changes. More specifically, the price for a service should be proportionally inverse to the elasticity of demand for this service. Prices that correspond to such
a structure are also called Ramsey-Boiteux prices. Accordingly, it is efficient to recover the fixed and common costs especially from those services where demand reacts only slightly to price changes. In contrast, the mark-up should be small for those services where demand elasticity is high.

In fact, firms themselves set prices according to the structure just described when they maximize their profits, i.e. firms use exactly such a price structure even without regulatory intervention, just out of their self-interest to maximize profits. Therefore, the structure of Ramsey-Boiteux prices exactly corresponds to a profit maximizing firm’s price structure that it is not regulated (see Laffont and Tirole, 2000, p. 65), regardless of whether retail prices for final customers or wholesale input prices such as access fees or both are concerned. If the efficient price structure results from unregulated profit maximizing behavior by an unregulated firm however, then Government regulation of the price structure is not only superfluous, but may even reduce social welfare.

2.2 Competition Intensity in Mobile Telephone Markets

Even though a firm’s price structure is generally efficient even without competition, the price level may be inefficiently high so that an allocative efficiency loss can result. In general, prices will be the closer to the allocatively efficient price level the more intensive actual or potential competition is in a market.

The intensity of market competition in turn depends to a large degree on the extent to which barriers to entry exist. Entry barriers are a key factor in determining the firms' market power and also the number of firms in a market. In fixed cost intensive industries market entry will occur as long as additional firms expect that they can operate at least at cost covering levels. Formally, the number of firms \( m \) that are active in a market is given by

\[
m = \left(\frac{1}{F} \sum_{j=1}^{n} (p_j - c_j)q_j(p_j)\right),
\]

where \( F \) are the sunk fixed and common costs that are on average associated with market entry, while \( p_j \) and \( c_j \) is the (expected) price and variable cost for services \( j=1, ..., n \), and \( q_j(p_j) \) is the (expected) market demand for this service at a price \( p_j \). Hence, the higher the sunk fixed and common costs \( F \) are, the lower is, ceteris paribus, the number of firms in a market, and the higher tends to be the contribution towards the fixed and common costs.
In a market without significant entry barriers firms will enter up to the point where an additional firm would not be able to operate without losses. Existing firms are disciplined in their price setting behavior through actual or potential market entry, and the firms' profits are thus restricted. In such a situation prices not only correspond in their structure, but also in their level to Ramsey-Boiteux prices. While the resulting price level will not correspond to incremental costs, it is important to note that the fiction of competition with prices set at incremental costs cannot be upheld in an industry that is fixed and common cost intensive. Correspondingly, the idea that mobile telephone service prices should be "cost-oriented" demonstrates a lack of understanding for competition within the mobile telephone industry in both theory and practice.

2.3 Market Entry Barriers

Hence, while it is reassuring to know that prices tend to be efficient when market entry barriers are low, the question still remains as to whether barriers to entry into mobile telephone markets are sufficiently low to discipline firms in their pricing policies. There are two potential factors suggesting that barriers to entry into mobile telephone markets may not be low, but rather play a major role: (1) The limited availability of frequency spectrum and (2) possible consumer switching costs, i.e. a potential reluctance on the consumers' side to switch from an existing provider to an cheaper alternative. Let us consider these two factors in turn.

2.3.1 Limited Spectrum Availability as a Barrier to Market Entry?

The limited availability of frequency spectrum is important because operating a mobile telephone network is not possible without access to spectrum. In practice, the necessary access or usage rights are allocated by the Government in some way or the other in all jurisdictions. The number of mobile telephone networks is, therefore, limited by the number of licenses issued by the Government, given the state of technology.

The question poses itself therefore whether the number of operators, so limited by the available frequency spectrum, is sufficient to guarantee workable competition within the mobile telephone industry. From an empirical perspective the OECD (2000) reports that markets with four and more operators have grown faster than markets with one, two or three operators. The causality, however, is unclear. On the one hand, markets with four operators may exhibit particularly intensive competition, which may lead to faster growth. On the other hand, however,
markets may attract a larger number of operators exactly because they are particularly attractive, so that the number of operators is not the cause, but the result of a market’s characteristics.

In another empirical study, Crandall and Hausman (2000) have shown that, in regional US mobile telephone markets, one additional operator has been sufficient to lower prices significantly when compared to a monopoly, while further market entry did not have significant effects on mobile telephone prices according to the study. Other studies such as Parker and Röller (1997) and Busse (2000) however, show that a small number of mobile network operators is likely to lead to collusion with correspondingly higher prices. Hence, the empirical literature does not arrive at clear conclusions about which market structure or how many operators are necessary to compete away any potential “super-normal” profits.

Market entry could, however, be facilitated more easily if some regulatory barriers were removed. For example, private property rights in frequency spectrum could be broadened so as to allow spectrum use for different purposes and spectrum trade, as already suggested by Coase (1959). In this case, situations were avoided like that in which companies such as Quam in Germany hang on to their license, once acquired, if only for the reason that the license expires otherwise without compensation. With the possibility of spectrum trading the opportunity cost of license holding would be strictly larger than zero, so that frequency holders would transfer their rights sooner than they do now (without transferability), where they simply have to hand them back to the Government without compensation. For potential newcomers obtaining access to spectrum would be easier with tradable spectrum rights. As with all other markets, the institutional framework for such a frequency trade could be set up either by the State or by private bodies as is the case in most stock markets or other exchanges. Furthermore, measures such as spectrum caps and the general competition law could also be applied in order to prevent operators acquiring a dominant position in a market or significantly lessening competition.4,5

Independent from the transferability of spectrum rights, in many European countries the availability of frequency spectrum does not appear to be a limiting factor to market entry at the moment. On the contrary, in a number of countries more than enough spectrum is available at the moment, without any interest in further market entry.6 Apparently there are no further entrants who expect to be able to profitably enter the market and to operate without losses in the long run.

This may point towards barriers to market entry other than spectrum availability that effectively keep newcomers from entering the market. A candidate for such a barrier to market
entry is the customers’ loyalty for already existing operators or, as the other side of the coin, customer switching costs.

2.3.2 Switching Costs as Barrier to Market Entry?

Switching costs make market entry more difficult, as the established operators’ customers only change over to a new entrant when the price advantage exceeds the customers’ switching costs. It should be noted here that switching costs not only consist of the direct monetary expenses when changing providers such as a contract penalty when canceling a mobile telephone contract prior to expiry. Instead, switching costs also include indirect costs such as the costs associated with giving up one's mobile telephone number or, alternatively, the fee for having the number ported. In this context, the introduction of mobile number portability (MNP) can reduce consumers' switching costs and thereby not only make market entry easier for new operators, but also strengthen the competition for customers between already established operators.

The question whether introducing MNP is also efficient cannot be answered so easily, however. For one thing, the system costs of introducing and implementing MNP can be quite substantial (see Aoki and Small, 1999). In addition, it becomes more difficult for consumers to recognize which mobile telephone number belongs to which respective network with MNP so that it is also more difficult to retain an oversight of the prices one has to pay. This again can lead to a price increase for mobile phone calls (see Gans and King, 2001). Moreover, with MNP new operators may have a tendency to price less aggressively than without MNP. This is because switching costs are lower with MNP so that the critical price advantage that entices customers to change is correspondingly smaller (see Bühler and Haucap, 2003). The total welfare effect of MNP is, therefore, not obvious and the introduction of MNP not necessarily efficient.7

It is also not clear, at least in theory, whether the existence of switching costs really leads to “too little” market entry and competition at all (see Klemperer 1987, 1988). And finally operators may have incentives themselves to reduce consumers’ switching costs in order to become more attractive for potential customers (see Chen, 1997; Garcia Marinoso, 2001).

Since churn rates are currently not exactly negligible in mobile telephony, the pure existence of some consumer switching costs does not seem to substantially impede the competitive process in mobile telephone markets either.8 That otherwise profitable market entry does not occur because of switching costs, therefore, appears largely improbable and
implausible, particularly since MNP has in the meantime also been introduced in many jurisdictions or its introduction is just about to take place.

In addition, certain switching costs also exist in a whole range of other markets. For example, changing banking connections is associated with switching costs, as is canceling a magazine subscription or changing a liability or other insurance. Nevertheless, there is generally lively competition in these markets, without ex ante regulation of all fees and rates. It is therefore unclear why the pure existence of some switching costs should justify regulatory intervention.

2.3.3 Market Equilibrium Without Market Entry?

That no further market entry occurs may alternatively be an expression of a competitive market equilibrium. If an additional entrant's fixed costs are not covered by the expected net revenues (i.e., revenues minus imputable variable costs), the market may be in a competitive equilibrium. If this is the case, then there is obviously no need for regulatory intervention. On the contrary, inadequate regulation may disturb the market equilibrium and induce uncertainty and inefficiencies in the first place.

2.3.4 Regulation Induced Uncertainty as Market Entry Barrier

Rules and regulations that prevent market entry and induce uncertainty can considerably weaken incentives for market entry and investment and innovation in general. This is above all the case when firms that specifically invest into infrastructure development are required to grant third parties access to their facilities against regulated fees.

To undertake specific investment into infrastructure development, firms need planning certainty and security. Of particular importance is that firms are allowed to recoup their sunk cost via an appropriate price policy if their investment and/or innovation is successful.

In many industries it has been recognized that appropriate incentives are necessary for innovation and specific investments to occur. For example, in the pharmaceutical industry patent protection provides incentives for firms to develop new medicines. Copyrights serve similar functions in other industries. Although these mechanisms induce static inefficiencies, as firms are granted monopoly rights (for a limited time) so that they will not supply at prices equal to incremental costs, these mechanisms are by all means efficient from a dynamic viewpoint, as
they drastically increase the incentives for developing the work in question, i.e. developing new medicines.

Telecommunications network operators need a similar protection in order to undertake specific investments in new infrastructure assets. Without such safeguards (e.g. because investors fear that they have to grant access to their facilities in case of success), investment incentives are significantly reduced. If third parties obtain the right to use another operator's facilities even without his consent and without any consideration to the initial investment risk, this is equivalent to a free option that these third parties are granted by the regulatory authority at the investing firm's expense (see Hausman, 1999, 2001.).

A regulatory policy that privatizes risks and socializes profits therefore does not only reduce the incentives for investment by established providers. For newcomers, the incentives to enter a market with an own infrastructure are also drastically reduced. Instead it becomes attractive to enter markets as a service provider without own infrastructure, which tends to result in to little investment and innovation.9

2.4 Interim Summary

To understand competitive processes and outcomes in mobile telephone markets it is of decisive importance to note that the development of mobile telephone networks requires large specific investments, which have carried quite significant risks and still do, as the current investment into 3G-networks (such as UMTS) illustrates. Such extensive investments and innovations will only be undertaken if firms can expect a return on investment after they have invested. However, for a return on investment firms have to cover fixed and common costs, so that prices need to exceed short-run incremental costs. In such situations, so-called Ramsey-Boiteux prices are efficient.

While unregulated firms will use such a price structure more or less automatically, the price level has to be driven down by effective competition. A competitive equilibrium is now reached when barriers to market entry are low and further market entry is not profitable. Such a situation is not implausible for many mobile telephone markets, as the availability of frequency spectrum often does not seem to be a limiting factor, and switching costs do not seem to overly impede competition in mobile markets either.

However, regulation induced uncertainty can lead to under-investment and too little infrastructure based market entry. Hence, the question presents itself as to how strong the need
for regulatory intervention is under these circumstances and how mobile telephone regulation is to be evaluated from an economic perspective in general. The next section will deal with just these questions.

3. Regulating Mobile Telephone Markets

3.1 Principle Considerations

In order to approach the question as to how regulatory intervention in mobile telephone markets is to be evaluated in principle, one has to consider the objectives of regulation. From an economic perspective, market regulation should pursue only an efficiency objective, whereas other social objectives (such as income distribution) should be pursued with other instruments (e.g. tax and social policy). In practice, however, this is generally not the case, for obvious political economy reasons (see e.g. Baldwin and Cave, 1999), as, for example, universal service obligations in telecommunications demonstrate. Nevertheless, from an economic perspective, efficiency is the criterion against which regulatory interventions are judged.

Generally economists distinguish between static and dynamic efficiency. Static efficiency refers to a specific point in time, and it is given when social welfare is maximized at this point in time. In the simplest case (without transaction costs) static efficiency is given when all prices correspond to the according marginal costs. In contrast, dynamic efficiency refers to a time period and means that social welfare is maximized over the entire time period. Hence, a dynamic efficiency criterion also incorporates investment and innovation. Dynamic efficiency therefore almost always requires a deviation from prices set at marginal cost.

The differences between static and dynamic efficiency are the larger, the more important innovations and (other) specific investments, i.e. sunk costs, are and the higher the uncertainty associated with innovation and specific investments is. Hence, there is a trade-off between static and dynamic efficiency when deciding upon the regulatory framework.

3.2 Basic Rules of Good Regulation

To achieve long-run market efficiency the regulatory framework should on the one hand be designed to protect consumers from firms’ potential market power (or its abuse by the operators), but on the other hand it should also protect firms from expropriation through the Government.
From an institutional economics perspective a regulatory framework can be seen as an implicit contract between operators and customers, which is administered by a neutral third party, the regulating authority (see Goldberg, 1976; Crocker and Masten, 1996).

The more important specific investments are in an industry, the more important it is that the regulatory framework is stable and predictable. An example may illustrate this point: Let us consider a project that requires a specific investment of 10 million Euros. The assets’ expected time of use is 20 years, and both operational expenditures (OPEX) and capital costs (CAPEX) each amount to 1 million Euros per year. As the investment is specific and cannot be sold in the case of failure, the capital costs will arise in any case, independent of whether and how the assets will actually be used. The costs are therefore sunk. In contrast, the operational expenditures are avoidable by ceasing operation.

It is then socially efficient to undertake this investment as long as there are sufficiently many consumers whose willingness to pay is high enough so that the firm can cover its cost in the long-run. For this to happen the firm has to be reasonably sure that the regulatory framework will not suddenly be changed and the firm not be forced to lower its prices after the investment has taken place. From an economic perspective this would be equivalent to expropriating the investors (see Sidak and Spulber, 1996).

Hence, to facilitate investment it is decisive to have a stable, predictable regulatory framework. The promise by the State not to “hold up” the investing firm ex post (i.e. after the investment has been made) and not to shift access rights by changing the regulatory framework must be credible. At the same time, however, the regulatory authority must possess a certain flexibility to consider new developments. Good regulation should therefore be open, transparent, consistent and accountable. This requires, amongst other things, open and transparent processes and criteria, which can be subject, in principle, to judicial review (see also Levy and Spiller, 1994, as well as Williamson and Mumssen, 1999).

The ex post introduction of sector specific regulation can therefore be regarded as a Government breach of the implicit regulatory contract. Through this “hold up” by the State firms’ profits are socialized and redistributed even though they carried the original investment risk. Such a regulatory “hold up” (or the expectation of eventually being held up by the State) considerably reduce firms’ investment incentives and therefore also harm consumers in the long run.
3.3 Problems of Long-Run Incremental Cost (LRIC) Regulation

In regulatory practice, essential facility operators are usually allowed to recover their costs as they do not have to grant competitors access at short-run, but at long-run incremental costs. In calculating the long-run incremental costs (LRIC), fixed and common costs are either distributed over a networks elements (TELRIC) or over the services delivered over the network (TSLRIC). Generally the so-called cost of ‘efficient service provision’ is used as the regulatory standard.

Using such a LRIC standard is highly problematic as the investments’ specificity is not appropriately reflected, i.e. the fact that specific investments result in sunk costs. Regulating at long-run incremental costs means that, in the presence of sunk costs, the investing firm has to carry the risk of failure alone, while competitors that operate without their own infrastructure receive at no cost an option by the regulating authority to share the investor’s success (see Hausman, 1999, 2001).

Let us illustrate this again using the above example: Suppose that ex ante there is a 50 percent chance that a sufficient market demand arises so that an investor can earn revenues of 3.2 million Euros per year on the investment, while with a counter probability of 50 percent not even the OPEX of 1 million Euro per year can be earned. The latter means that the investment would strand and had to be completely written off. Nevertheless, the investor would still face the capital costs of 1 million Euros per year in this case. From an ex ante viewpoint it is still efficient to carry out such an investment, as the expected return from the investment is positive at 100,000 Euros. However, if the investor is forced to share his infrastructure at long-run incremental costs (i.e., 2 million Euros) in the case of success, then the investment does not take place. The looming TELRIC regulation therefore prevents investment and consequently also harms consumers.

The idea behind LRIC-regulation is that one can create a framework that resembles the features of competition in perfectly contestable markets, where firms operate without any sunk costs. As Hausman (2001, p. 55) has explained, however, the contestable market model cannot offer any insight into competition in telecommunications, exactly because the contestable markets model presumes that there are no sunk costs. However, specific investments and sunk costs are a key characteristic for telecommunications markets. According to Hausman (1999, p. 62 ff), when taking into account the risks of investing with sunk costs, access fees for fixed-line
telephone networks should lie more than 100% above their long-run incremental costs in order to provide incentives for efficient investment.

3.4 Regulation Must not be Based on a Nirvana Approach

More than 30 years ago Demsetz (1969) has, based on Coase (1960), sharply criticized the approach to base economic policy advice on comparisons between real world situations and unrealistic textbook ideals and labeled this nirvana approach. The observation that reality does not correspond to the ideal textbook model does not in turn imply that better policies exist and can be implemented. Instead, a comparative institutional economic approach is required, i.e. the comparison of alternatives that can actually be realized, in order to deduce economic policy recommendations (see Dixit, 1996). Comparing reality with a hypothetical ideal, a nirvana, is not suitable for advising economic policy.

A current example for this nirvana approach is the discussion on whether regulatory mechanisms such as carrier pre-selection or call-by-call, which have been successfully applied in fixed-line telephony in many jurisdictions, should also be mandated for mobile telephone networks. If this regulation is to be based on long-run incremental costs, two drastic problems will result:

One is that mobile network selection, regulated at long-run incremental costs, will prevent an efficient pricing policy á la Ramsey-Boiteaux; the other thing is that such regulatory intervention is a “hold up” by the Government, which destabilizes expectations and destroys future investment incentives, also in 3G-networks such as UMTS. Though the regulatory intervention may be efficient from a static perspective, it has devastating effects from a dynamic perspective, especially with regard to UMTS and other future infrastructure development, as investment and innovation incentives are reduced.11

3.5 Interim Conclusion

As we have seen, regulation carries risks that should not be underestimated and can, even if regulation is well-meant, lead to adverse consequences, particularly when specific investments are required for building an infrastructure. A good regulatory framework should therefore protect consumers from (the abuse of) firms’ market power, but also protect investing firms from a
breach of the regulatory contract by the State. Transparency, openness, consistency as well as comprehensibility of and the ability to review regulatory decisions are therefore criteria for a suitable regulatory framework. The ex post introduction of regulatory measures can be interpreted as a breach of the regulatory contract by the State, which can have devastating effects on investment and innovation incentive, particularly when regulation is based on long-run incremental costs.

Nevertheless, under certain circumstances competition problems can arise through the necessity of interconnection and call termination from fixed-line telephone networks. Further particularities of mobile telephone competition arise from national and international roaming requirements. These special problems will now be closer examined in the following section.

4. Special Problems of Mobile Telephone Competition

4.1 Interconnection and Termination in Mobile Telephone Networks

A special feature of telecommunications competition is competing networks usually benefit from interconnection so that customers on one network can also call customers on other networks and can likewise be called by them. Therefore, mobile network operators need to terminate calls from other mobile networks on their own network and they need to terminate calls from their own network on other networks. The same applies for fixed-to-mobile and mobile-to-fixed calls. For these termination services, mobile network operators usually pay and charge so-called termination rates.

4.1.1 Collusion Through High Termination Fees?

Mobile termination fees directly influence the incremental costs of the respective calls, so that they generally also affect the respective consumer retail prices.¹² Since, however, every mobile network operator can now raise its rivals’ call prices through high termination fees, a mutual price increase for termination services can result in a collusive equilibrium with strategically high termination fees and accordingly high mobile call prices, as Laffont and Tirole (1998a) and Armstrong (1988) have shown.

The mechanism just described does not lack a certain intuition, but closer analysis reveals that it can only be expected under very specific circumstances. In particular, for collusive
termination fees to be an equilibrium the respective networks must be (a) symmetric, i.e. of the same size, and (b), they must not compete in non-linear prices such as two-part tariffs or price discriminate between on-net and off-net calls. If, in contrast, the networks are not completely symmetric, but asymmetric, i.e. of different sizes, then it is often very difficult to determine an equilibrium, if one exists at all (see Bühler, 2002). This also means that one should generally not expect that operators can stabilize collusive equilibria by increasing each other’s termination fees. Especially in reality there are usually large and small, i.e. asymmetric, networks, so that the practical relevance of the mechanism described by Laffont and Tirole (1998a) and Armstrong (1998) is highly questionable.

4.1.2 Collusion Through Low Termination Fees?

Aggravatingly, probably all mobile network operators worldwide offer multi-part tariffs, which generally consist of at least a fixed fee and a price per minute. Moreover, many operators also price differentiate between on-net and off-net calls. In such situations operators’ incentives change drastically when setting the respective termination fees. (see Carter and Wright, 1999; Wright, 1999). Under some circumstances the incentives can even be completely reversed (see Gans and King, 2001). The logic is the following: The higher the termination fees, the higher is the profit from call termination (as long as the monopoly price has not been reached). With this it also becomes more profitable to acquire new customers who are then called by others. Competition for these customers, i.e. for market shares, can therefore be intensified through high termination fees. In contrast, the less attractive it becomes to acquire more customers, the less aggressive firms will compete for customers (see Carter and Wright, 1999).13

As Gans and King (2001) now argue, under such circumstances so-called “bill and keep” arrangements may enable collusion. Under “bill and keep” arrangements each operators bills his customer for calls into other networks, but he does not pay any termination fees and also does not charge other network providers for terminating calls on his network. Hence, the effective termination fee of zero under “bill and keep”. As this termination fee is therefore below the incremental cost of call termination, it becomes relatively unattractive to terminate calls. Consequently, the value of acquiring new customers is reduced, so that competition for new customers is accordingly softened.
4.1.3 Optimal Termination Fees

Wright (1999) extends the analysis of Gans and King (2001) by two essential aspects: Firstly, it is examined how mobile network operators’ incentives change when the market is not covered and demand still growing. And secondly, Wright (1999) does not only consider mobile-to-mobile calls as Gans and King (2001) do, but also fixed-to-mobile calls. As Wright (1999) shows, mobile termination fees that exceed incremental costs can well result if the market is not (yet) covered. Such a pricing policy enables mobile operators to use the termination revenues to “cross-subsidize” the fixed fee and to induce more consumers to enter a mobile telephone subscription. The higher termination fees, therefore, indirectly lead to higher penetration rates in mobile telephone markets.

Altogether Wright (1999) identifies two effects, which work into different directions from a welfare economic perspective: On the one hand termination fees may exceed the incremental costs and so reduce consumer welfare customers on fixed-line networks who initiate calls to mobile networks. On the other hand termination revenues are used to subsidize mobile telephony subscriptions and so to induce more customers to subscribe to a mobile telephone network, which is, particularly in the presence of network effects, a positive thing. According to Wright’s (1999) simulation analysis, the socially optimum termination fee should therefore even be at around 200 to 400 percent above the incremental costs.

Hence, there are relatively clear conclusions concerning the regulation of mobile termination rates: From an efficiency perspective, it may, for several reasons, be desirable to have termination fees that significantly exceed the imputable incremental costs. First of all, the Ramsey-Boiteux pricing rule demands that prices should be particularly high for those services, for which demand is relatively inelastic. This seems to be the case for calls to mobile networks, as Koboldt and Maldoom (2001) have found for the UK. Secondly, high termination fees indirectly intensifies competition for new subscribers between mobile networks and, consequently, this leads to higher penetration rates, so that larger parts of the population have access to mobile telephones.

Nevertheless, termination fees can also be too high from a welfare economic perspective. This may particularly be the case when consumers are ignorant about the prices for fixed-to-mobile or mobile-to-mobile calls (see Gans and King, 2000; Bühler and Haucap, 2003). The idea here is as follows: In a normal market setting, consumers base their purchase decisions on
individual firms’ prices. The firms again set their prices accordingly so that their incremental revenue equals their incremental cost. For calls to mobile telephones, it is now assumed that consumers do not know the individual firm’s prices or, alternatively, that, when placing a call, the callers often do not know which mobile network the called party is on (and, therefore, which price they have to pay). Therefore, callers may not base their calling decisions on the individual firms’ prices, but on an average price. In this case, the individual operator’s price policy has less effect on his demand than with perfect market transparency. This in turn leads to prices being inefficiently high.

If, for example, a mobile operator increases its termination fee, then the perceived \textit{average} price only increases by less than the actual price increase. At the same time the operator’s price increase has a negative external effect on other operators, as the demand for calls to their networks is also reduced. However, as every operator only considers the effect that a price increase has on the demand for calls into one’s own network (but not the effect on demand for calls into other networks), prices tend to be inefficiently high and may even lie above the monopoly price (see Gans and King, 2000; Bühler and Haucap, 2003).

Potential remedies can include a targeted coordination of termination fees or improved information about which participants belong to which network.\textsuperscript{14} In general, it can be stated again that the very fact that prices exceed incremental costs is not necessarily a sign for inefficient price setting or market failure. Quite in contrast, such price policies are \textit{necessary} for effective competition in mobile telephony.\textsuperscript{15}

4.2 \textit{International Roaming}

According to Salsas and Koboldt (2002), a very similar problem can arise with international roaming prices. When abroad, consumers often neither know the different networks’ prices for being called (termination) nor do they exactly know the prices for initiating a call (origination). As Salsas and Koboldt (2002) argue, specific coordination of roaming prices or improved information can here also lead to lower prices. In addition, horizontal mergers between operators in different countries can ease possible problems, as Vodafone’s pricing plan with one single roaming price in Europe illustrates.
In analogy to termination fees, let us again be point out that economic efficiency requires Ramsey-Boiteux pricing, which implies that call prices should be high and exceed incremental costs if demand is relatively inelastic.

4.3 National Roaming and Virtual Mobile Network Operators (VMNOs)

An entirely different problem than international roaming arises with national roaming and with virtual mobile network operators (VMNOs). In general, international roaming agreements are concluded between mobile operators from different countries, which do not operate in the same market. As these firms typically do not compete with each other, both have generally strong incentives to enter into international roaming agreements.

In contrast, national roaming agreements concern mobile operators that compete in the same market. Here some operators typically have larger network coverage than other, especially relatively new, operators. Therefore the question arises whether established operators have incentives (a) to conclude roaming agreements with new operators and (b) to let VMNOs into the market.

To answer this question it should first be noted that a monopoly provider has usually little incentive to allow for roaming, and a monopoly would normally not voluntarily allow for VMNOs either (see Valletti, 2003).16

Even with several operators there may still be competition problems: Valletti (2003) has analyzed the case with two duopoly network operators who strategically engage in vertical product differentiation through different network coverage, in order to soften price competition. According to Valletti (2003), only one duopolist will choose full network coverage in such a case, while the other will save costs and only build a smaller network. In fact, as Valletti (2003) shows, the two operators will only conclude a roaming agreement in this case if this can facilitate collusion. Otherwise, neither of the two operators will have an interest in national roaming. Hence, Valletti (2003) concludes that voluntary national roaming agreements should be considered a sign of collusion.

The situation is different, though, if either (a) the operator who has full network coverage also benefits from roaming agreements or (b) if there is already more than one mobile operator with full network coverage in a market. The operator with full network coverage (network A) may, for example, benefit from national roaming agreements because this also allows his own
customers to place more calls (as users on network B can be easier reached). Foros, Hansen and Sand (2002) analyze exactly this situation and conclude that operators will always have incentives to voluntarily enter into roaming agreements in this case.\textsuperscript{17}

The same applies when there is already more than one operator with full network coverage who still has spare capacity. In this case, operators have incentives to sell or lease their spare capacities to third parties. Even though this intensifies competition for customers and may partly cannibalize the operator’s own profit, it is still more advantageous to profit from leasing capacities to third parties rather than having others leasing their capacities, as competition will be intensified in any case (see also Haucap and Wey, 2003). This logic does not only apply to VMNOs, but also to national roaming agreements.

While it must be ensured that established operators do not agree amongst each other to collusively deny market entry, for the above logic to apply, collectively excluding VMNOs and refusing to enter into national roaming agreements may also be dynamically efficient, as it makes infrastructure investment more attractive (see Haucap and Wey, 2003).

### 4.4 Summary

Altogether we can conclude that a need to regulate specific mobile telephone services or aspects cannot simply be deduced from the fact alone that certain services may appear to be especially profitable, or from the observation that newcomers without their own infrastructure do not always find it easy to compete with established operators. On the contrary: Regulatory intervention can often cause more damage than solve problems.

The potential failure of Government regulation in mobile telephony has been impressively analyzed by Hausman (2002, p. 591) in two separate studies: In his first study Hausman empirically examines retail price differences between various mobile telephone networks in the US, where operators in 26 states had been price regulated for the first 12 years of mobile telephony, whereas 25 states left prices unregulated. According to Hausman (2002) average prices in the regulated states were about 39% above the prices in non-regulated states in 1994. For the period from 1989 to 1993 Hausman finds that a price difference of 14 to 18 percent can be explained alone through the existence of price regulation, with a higher price being observed on average in regulated states. Accordingly, after mobile telephony was completely
deregulated in 1995, mobile telephone prices fell by about 14 percent more in the states that were previously regulated when compared to the states that had previously been unregulated.

In his second study Hausman (1997) investigates the welfare loss that can be caused by regulatory failure. According to Hausman (1997) the introduction of mobile telephony in the US were effectively delayed by about 10 years through protracted decisions by the FCC and slow licensing proceedings (also see Hausman, 2002, p. 588). According to Hausman, these delays have resulted in a welfare loss of between US$19 and US$50 billion (basis: 1994).

5. Conclusions

In conclusion, let us once more stress again that the development of mobile telephone networks requires highly specific investments, which have been and are still carried out under significant uncertainty, as the 3G-networks such as UMTS illustrate. Specific investments inevitably result in sunk costs. Therefore, extensive investments and innovations can only will only be undertaken if firms have a justified expectation that they are allowed to make profits after they have invested. However, in order to cover the fixed costs and to operate profitably, a pricing policy is necessary that involves prices above short-run incremental costs.

Hence, the mobile network operators’ cost structure, which is characterized by high, largely sunk, fixed costs and relatively low variable costs, is decisive for their price setting. In such a situation a price structure is efficient, where different services carry different prices and according mark-ups. A situation, where all services carry the same mark-up, is generally inefficient. On the contrary, services, for which demand is relatively inelastic, should be priced relatively high, while prices should be close to incremental costs when demand for a service is relatively elastic. It is exactly this price structure that results automatically out of the firms’ profit maximizing interest. It is therefore a principle of efficiency and not a sign of market failure that mark-ups for different services and markets differ from each other.

An ex post introduction of sector-specific regulation can be regarded as a breach of the implicit regulatory contract by the State. Through this “hold up” by the State, mobile network operators’ profits are socialized and redistributed, while the operators had to carry the original investment risk themselves. Such a regulatory “hold up” considerably reduces investment incentives and also harms consumer in the long-run. In addition, regulatory failure is a real risk in practice, as the empirical evidence shows.
At closer inspection, potential competition problems, which result from the necessity for interconnection and fixed-to-mobile termination, do not justify sector specific regulation either. The same applies to mobile number portability.

Based on these considerations, one is left to fiercely advise against sector specific regulation of mobile telephone markets. The social welfare loss that would arise from such regulations are estimated to be enormous.

Endnotes
1) According to the ITU, in 1991 worldwide revenues in mobile telephone markets amounted to US$19 billion and the number of customers to 16 million. For 2003, the ITU estimates a turnover of US$414 billion with more than 1.3 billion customers (see ITU, 2003).
2) For an exemplary analysis of the Italian mobile telephine market see Cricelli, Gastaldi and Levialdi (2002).
3) So-called network effects may have also played a role in the rapid growth of mobile telephone networks (see e.g. Economides, 1996).
4) This is the case, for example, in New Zealand, where spectrum trade is subject to the Commerce Act 1986, New Zealand's competition law. Particularly section 47 of the Commerce Act is applicable, whereby it is prohibited for firms to acquire assets, when this has the effect of substantially lessening competition in a market.
5) Detailed economic analyses of spectrum trade are provided by Valletti (2001) and Leese, Levine and Rickman (2002).
6) For example, in France, Belgium and Portugal fewer UMTS licenses have been allocated than were available. In other countries such as Germany, market expectation are that not all firms that hold a UMTS license will actually build a UMTS network and enter the market in the end.
7) However, if MNP is mandated, a number of factors suggest that the Government should carry the cost of introducing MNP. Thats is, the Government should basically purchase MNP from mobile network operators because MNP has strong public good characteristics (see Haucap, 2003). However, Gans, King and Woodbridge (2001) arrive at a different conclusion: They argue that the mobile network operators should firstly be obliged to introduce MNP and secondly also bear the associated costs if at least one customer demands MNP. The analysis by Gans, King and Woodbridge (2001) assumes, however, that introducing MNP does not involve any fixed costs.
9) This problem will be further discussed in section 3.
10) In a world with transaction costs, however, it is considerably more difficult to define a clear efficiency standard (see Furubotn, 1999).
11) For details on the regulated selection of connection network operators see Kruse (2002).
12) This does not necessarily always and under all circumstances have to be the case. Marketing considerations and transaction costs may speak for uniform prices for calls in different mobile networks, even if their termination fees differ from each other. In general though, termination fees should strongly affect the corresponding retail prices.

13) Laffont and Tirole concede this themselves in a second paper (see Laffont and Tirole, 1998b).

14) This problem can be solved by changing from a “Calling Party Pays (CPP)” to a “Receiving Party Pays (RPP)” system, where not the caller, but the called party pays the calling charges (or at least the termination fees) (see Doyle and Smith, 1998). However, as empirical studies show, RPP seems to have a negative effect on mobile telephone penetration (see OECD, 2000).

15) A detailed theoretical analysis of mobile termination and its possible regulation can be found in Gans and King (1999a, b).

16) With so-called network effects or consumer “lock-in”, this does not have to be the case. On the contrary, even a monopolist can have incentives to invite entry if network effects are important (see Economides, 1996) or if consumers fear to be locked in by a monopolist (see Farrell and Gallini, 1988).

17) However, as Foros, Hansen and Sand (2002) show in their theoretical analysis of roaming and VMNOs, voluntary roaming agreements may lead to inefficiently low infrastructure investment. If there are VMNOs, in contrast, regulatory intervention is not required according to the authors. Furthermore, Roson (2001) shows that infrastructure investments tend to be at inefficiently low levels if roaming is regulated.

References


**Bisher erschienen:**

**Diskussionspapiere der Fächergruppe Volkswirtschaftslehre**

- Kruse, Jörn, Regulierung der Terminierungsentgelte der deutschen Mobilfunknetze?, Nr. 1 (Juni 2003).
- Dewenter, Ralf, Rational Addiction to News?, Nr. 2 (Juni 2003).
- Haucap, Justus, The Economics of Mobile Telephone Regulation, Nr. 4 (Juni 2003).
- Dewenter, Ralf, Media Markets with Habit Formation, Nr. 5 (Juni 2003).
- Bräuninger, Michael, A Note on Health Insurance and Growth, Nr. 6 (Juni 2003).
- Dewenter, Ralf, Quality Provision in Interrelated Markets, Nr. 7 (Juni 2003).
- Josten, Stefan Dietrich, Dynamic Fiscal Policies, Unemployment, and Economic Growth, Nr. 9 (Juni 2003).
- Dewenter, Ralf, The Economics of Media Markets, Nr. 10 (Juni 2003).
Frühere Diskussionsbeiträge zur Wirtschaftspolitik


**Frühere Diskussionsbeiträge aus dem Institut für Theoretische Volkswirtschaftslehre**

• Bräuninger, Michael, Social Capital and Regional Mobility, Nr. 4/2002.
• Heppke, Kirsten, On the Existence of the Credit Channel in Poland, Nr. 8/1999.
• Bräuninger, Michael, Unemployment and International Lending and Borrowing in an Overlapping Generations Model, Nr. 8/1999.
• Henning, Andreas und Wolfgang Greiner, Organknappheit im Transplantationswesen - Lösungsansätze aus ökonomischer Sicht, Nr. 7/1999.
• Carlberg, Michael, European Monetary Union: The New Macroeconomics, Nr. 4/1999, erschienen in: Gerhard Rübel (Hg.), *Real and Monetary Issues of International Economic Integration*, Berlin 2000, S. 155-175.


**Frühere Diskussionsbeiträge zur Finanzwissenschaft**


