Knieps, Günter

Working Paper

Limits to the (De-)Regulation of Transport Services

Diskussionsbeiträge / Universität Freiburg i. Br., Institut für Verkehrswissenschaft und Regionalpolitik, No. 99

Provided in Cooperation with:
Institute for Transport Economics and Regional Policy, University of Freiburg

Suggested Citation: Knieps, Günter (2004) : Limits to the (De-)Regulation of Transport Services, Diskussionsbeiträge / Universität Freiburg i. Br., Institut für Verkehrswissenschaft und Regionalpolitik, No. 99

This Version is available at:
http://hdl.handle.net/10419/23013

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.
Limits to the (De-)Regulation of Transport Services

by Günter Knieps

Discussion Paper
Institut für Verkehrswissenschaft und Regionalpolitik
No. 99 – May 2004

Abstract:
In section 1 the disaggregated approach of network economics is explained which distinguishes between the network levels transport services, traffic control systems and fixed infrastructure. Section 2 starts with an explanation of the theory of monopolistic bottlenecks, which can be used as a basis for distinguishing between parts of a network where competition functions efficiently and those that enjoy stable, network-specific market power. Section 3 explores the opportunities for potential and active competition on the markets for transport services, while Section 4 explains the potential for auction competition in respect of air traffic and train control systems. Section 5 is devoted to regulation in connection with access to transport infrastructure, including the relationship between disaggregated regulation of bottlenecks and the essential facilities theory, and the development of an appropriate approach towards regulation of access charges. Lastly, the paper looks at the ongoing reform of access to European airports (Section 6) and rail networks (Section 7).
1. Introduction

Ever since the European Court of Justice ruled against the Council of Transport Ministers in 1985 for failing to ensure freedom to provide services in the sphere of international transport, the paradigm shift towards full competition on the European transport market has become irreversible. The European Union has played a leading role in this process, and the benefits of free access to transport markets throughout Europe are now largely uncontested.

At first sight the carriage of passengers and goods by road, rail, sea and air seems very heterogeneous. Common to all the different forms of transport, however, is the fact that access to transport services requires the use of transport infrastructure. This means the development and operation of not only transport infrastructure but also traffic control systems.

Unlike past studies, which tended to focus on one specific sector, this paper opts for a disaggregated approach that distinguishes between three network levels (Knieps, 1996):

(1) Transport services (carriage of passengers and goods by train, air, ship, lorry or private car);

(2) Traffic control systems (e.g. air traffic control, train control, road traffic control and information systems);

(3) Fixed infrastructure (e.g. railway tracks, stations, roads, airports).

With this approach it is possible to conduct a thorough analysis of competition potential and the continuing need for regulation in the future, and in particular, a separate study can be made of the question of access to transport infrastructure in connection with today’s network economy.

Efficient competition on European transport markets is conditional upon the existence of non-discriminatory access to infrastructure for all active and potential transport service providers. In addition, however, efforts must also be made to
ensure scant infrastructure capacities are shared out efficiently and costs are covered. This paper presents a disaggregated approach to regulation that enables these objectives to be met as comprehensively as possible.

Section 2 starts with an explanation of the theory of monopolistic bottlenecks, which can be used as a basis for distinguishing between parts of a network where competition functions efficiently and those that enjoy stable, network-specific market power. Section 3 explores the opportunities for potential and active competition on the markets for transport services, while Section 4 explains the potential for auction competition in respect of air traffic and train control systems. Section 5 is devoted to regulation in connection with access to transport infrastructure, including the relationship between disaggregated regulation of bottlenecks and the essential facilities theory, and the development of an appropriate approach towards regulation of access charges. Lastly, the paper looks at the ongoing reform of access to European airports (Section 6) and rail networks (Section 7).

2. Theory of monopolistic bottlenecks

An appropriate economic reference model that exposes the need for action to control market power in network sectors must be capable of grasping essential network characteristics (cluster/bundling effects, externalities, etc.) without automatically assuming these equate to market power. This Section will attempt to show that stable network-specific market power and the ensuing need for regulation only exist in the event of monopolistic bottlenecks.

The theory of monopolistic bottlenecks is central to the disaggregated regulation approach in terms of locating network-specific market power in connection with the efforts to determine the minimum basis for regulation (cf. Knieps, 1997a, pp. 327 ff; Knieps, 1997b), whereby the aim is to come up with a coherent basis consistent with the network economy which can be applied to all network sectors and which regardless of historical or institutional quirks provides justification for ex ante regulatory measures. The remaining network areas come under
general competition law. In this context, the need for regulation is concerned in particular with the need to design a system for controlling access to monopolistic bottlenecks and for charging users. The problems associated with monopolistic bottlenecks, and in particular the problem of network access (Baumol, Willig, 1999, p. 44; Knieps, 1997a, p. 327 ff.; Laffont, Tirole, 2000, p. 98) are currently frequent topics of discussion in the context of the network economy.

Network-specific market power can only be identified by consistently implementing Stigler’s concept of market entry barriers. According to Stigler:

“*A barrier to entry may be defined as a cost of producing (at some or every rate of output) which must be borne by a firm which seeks to enter an industry but is not borne by firms already in the industry*” (Stigler, 1968, p. 67).

Providing inputs are available to active and potential market players under the same conditions, Stigler considers there are no barriers to entry. Therefore, economies of scale, for example, do not constitute entry barriers providing newcomers to the market also have access to the same cost function. Stigler’s concept also implies that traditional competition parameters such as product differentiation coupled with the need to build up a good reputation and develop goodwill, or the capital required are not entry barriers because they affect all active and potential enterprises equally. In other words, they are situations where the cost functions depend only on factors that are systematically available to all enterprises.¹

The conditions governing a monopolistic bottleneck are met when:

¹ On the other hand, the different entry barriers found, according to Bain (1956), in the traditional industrial economy (economies of scale, product differentiation, high capital needs, etc.) are not reliable proof of stable market power (cf., for example, B. Schmalensee, 1989). Von Weizsäcker (1980a; 1980b) shows, for example, that reputation and goodwill are effective ways of reducing insecurity, which can enhance social well-being. According to Stigler, the development of goodwill is not a barrier to market entry because it does not result in cost asymmetries between established firms and newcomers to the market.
(1) a facility is essential for reaching customers, i.e. if no second or third such facility exists, in other words if there is no active substitute. This is the case when cluster effects produce a natural monopoly and a single provider is able to make the facility available more cheaply than several providers;  

(2) at the same time the facility cannot reasonably be duplicated as a way of controlling the active provider, in other words when there is no potential substitute. This is the case when the costs of the facility are irreversible and there is therefore also no second-hand market in operation for such facilities.

Consequently, network-specific market power in the hands of the established enterprise is only to be expected part-areas characterised, at one and same time, by cluster effects and irreversible costs. Although they are no longer relevant for decision-making by established enterprises, as far as potential competitors are concerned irreversible costs are a crucial factor, insofar as they must decide whether to invest such costs in the market or not. Established firms therefore have lower decision-relevant costs than their potential rivals. This means there is room for strategic manoeuvring, with the result that inefficient production or surplus profits no longer necessarily enable newcomers to enter the market. The market power of the firm that enjoys such a monopolistic bottleneck is therefore stable, even if all market players are fully informed, all users are prepared to switch to another provider, and small price adjustments have an effect on demand.  

In the absence of irreversible costs, however, and as a result of the controlling effect of potential competition, cluster effects do not produce stable market

---

2 A natural monopoly exists when the cost function in the relevant sector is subadditive in relation to demand. In the case of single products, economies of scale are sufficient for there to be a natural monopoly. In the context of a study of the cost factor in the case of multiple products, cluster effects stand out, owing to the economies of scale and bundling effects associated with the provision of services. As a result of these cluster effects it may be possible for a single network provider to serve a given region more cheaply than several providers, and thus to enjoy a natural monopoly (cf., for example, Baumol, 1977).

3 This is Bertrand Nash’s behavioural assumption based on the theory of contestable markets (cf., for example, Baumol, Panzar, Willig, 1982).
power, regardless of the size of the relevant network operator’s market share, insofar as inefficient providers of non market-oriented services will be replaced by new entrant owing to the pressure of competition. In this case there is no need for regulation to limit the active operator’s control over the market.

The bottleneck theory does not set out to deny the information problems encountered to varying degrees by real markets. *Ex ante* stable market power cannot be deduced from the existence of information problems, however, insofar as markets tend to be very good at (endogenously) developing institutions to overcome their information problems. Switching costs, which occur in many areas of the economy, are no explanation for monopolistic bottleneck situations either. Examples of switching costs include monthly or annual season tickets for concerts that cannot be transferred if the holder moves house, or the costs incurred by firms when employees leave as soon as they have learnt the ropes, etc. They are no justification for regulatory measures and can be left to the market’s own problem-solving ability (cf., for example, von Weizsäcker, 1984; Tirole, 1989, Chapter 8). However, the existence of network externalities is no justification for sector-specific regulation either. The essential feature of such externalities is that for an individual the advantage of being part of a network depends not only on its technical specifications – its standard – but also on how many others are involved in it. Where there are positive network externalities, the benefit for the individual increases with the number of other network members, in other words the number of those using the same standard. In the absence of network-specific market power, negotiations between network operators can prove effective because both sides stand to benefit from the agreements. On the other hand, access to bottlenecks does present a need for regulation, given that network-specific market power allows for strategic manoeuvring that also hampers full enjoyment of positive externalities associated with access to the network (cf., e.g., Blankart, Knieps, 1995).

---

4 In the absence of irreversible costs, there is no evidence in the case of a natural monopoly of market power capable of withstanding alternative behavioural assumptions (cf. Knieps, Vogelsang, 1982). Market power based on the Cournot-Nash assumption becomes immediately unstable with the switch to the Bertrand-Nash behavioural assumption. Action taken by competition authorities would therefore have to refer to behavioural assumptions that are difficult to verify in practice.
Indeed, one of the essential features of the ability of competition to operate on the free markets for transport services is that corporate strategies such as product and price differentiation, the build-up of goodwill and the development of an efficient distribution network, etc, can also be used for strategic purposes. Information problems (search costs, asymmetric information, etc.) can also play a role. This must not lead to the opposite conclusion, however, namely that basically competition does not work on transport markets, and nor does it mean that general competition law should not be applied on these markets. What it does mean, however, is that, as on any other market characterised by organised competition, the burden of proof as to the existence of market power and as to whether such power is abused, rests with the competition authorities. In contrast to general ex ante regulation, such interference in competition should always be carried out only on a case-by-case and ex post basis.

3. Competition potential on the markets for transport services

Active and potential competition operates on the transport markets. The very fact that transport services are on offer in the form of a network with their associated bundling effects implies that there is no monopolistic control where transport undertakings have free access to the market, since high profits recorded by one undertaking have the immediate effect of attracting others. There is no danger of preventing competitors from entering the market insofar as the decision-related costs in respect of transport services are similar for established

---

5 The Bertrand Nash assumption, based on the contestable markets theory, does not set out to deny the information problems encountered on real markets either. Stable market power cannot be deduced simply from the existence of information problems, insofar as markets tend to be very good at (endogenously) developing institutions to overcome their information problems, for example, by building up goodwill. Conversely, stable market power is also to be found in natural monopolies with irreversible costs when all market players have all the information they need.

6 In this context, the competition authorities must weigh up two potential sources of error. Firstly, false positives can occur when the authorities interfere in the competition process even though competition is working and there is no need at all for action in terms of competition policy. Secondly, false negatives occur when the competition authorities fail to act even though competition policy calls for action (cf. Knieps, 1997c, p. 51).
undertakings as for potential rivals. As a result, the irreversible costs associated with providing rail services on a railway network, for example, play no significant role. The use of trains is not confined to certain lines; they are just as mobile geographically as aeroplanes or lorries.

For competition to be effective, however, the conditions of access to the transport infrastructure must be the same for all (active and potential) service providers. If established undertakings have preferential access to scant infrastructure capacities, they enjoy unwarranted advantages over others that can result in their gaining control over markets that were otherwise competitive.

Whereas the theory of contestable markets examines only the role of potential competition with identical cost functions for both active providers and potential rivals (cf. Baumol, 1982; Panzar, Willig, 1977), effective competition on the markets for transport services does not only mean potential competition. Often a newcomer enters the market with no intention of duplicating the established undertaking. What is important is active competition achieved by means of technological and product differentiation, and the introduction of new products and processes. As a direct consequence of this, it is misleading to assume that newcomers have as their reference point the belief that ideally there can only be one transport network on the markets for transport services.

In the rail sector too active competition on busy lines should mean more efficient pricing, including more incentives for cost-efficiency and pressure to offer services tailored to meet demand. In the passenger sector, the pressure of competition reveals whether the length of the trains deployed and the intervals between them matches demand. Past supply concepts (e.g. clockface timetabling) are then brought into question when customers fail to honour them with a corresponding demand. Regular runs with (almost) empty “ghost trains” are no longer sustainable in a competitive context but, on the other hand, there are new incentives for providing a flexible supply of extra services at peak times. Newcomers’ entry onto the market broadens the range of services offered extensively as well as widening consumers’ choices in terms of the price and service quality. Opportunities for new entrants include the detection and exploitation of gaps in
the market, such as the development of a Europe-wide express service for passengers and goods based on a high-performance, computer-assisted logistics system. Service improvements are also possible over shorter distances, however, with examples including a denser timetable offering better connections. In addition to the pressure of potential competition, active competitive between different undertakings is therefore also a source of potential that should not be underestimated.

4. Competitive potential of traffic control systems

The provision of a transport service requires not only a vehicle (e.g. a train or aeroplane) but at the same time access to infrastructure (e.g. a railway line or landing slot). With rail and air transport, traffic movements must also be constantly monitored and coordinated. Train and air traffic control systems are needed not only to guarantee traffic safety but also to allocate the available infrastructure capacity. Traffic control systems also look set to play an increasing role in the road sector.

It is important to remember that the provision of transport services requires simultaneous access to infrastructure and a traffic control system regardless of whether these functions are vertically integrated within a single undertaking or whether they are performed by several different undertakings. Although airport operators, air carriers and air traffic control authorities can only guarantee smooth services by working together, they have always been separate from each other, both in terms of the way they are organised and institutionally. Such is not the case in the rail sector, however, where all functions used to be vertically integrated within the national railway companies and where cooperation between national companies tended to be minimal. But in this sector too there have been recent moves towards a disaggregated approach towards regulation. Competition on the railway networks is only possible if railway undertakings have unimpeded access to railway lines and at the same time can use the services provided by the different train control systems (cf., for example, Berndt, Kunz, 2003, p. 186 ff.).
Train control systems are the crucial link between railway infrastructure and operations. Both the throughput of traffic and repairs carried out on the track must be coordinated by such systems. Like in the air transport sector, the cost of such coordination is basically the same regardless of whether only one, or more than one railway undertaking is operating over the network. Rather, it depends on the number of trains and their operating speed.

Traffic control systems do not constitute monopolistic bottlenecks. They are natural monopolies, whose geographical limits have to be clearly defined (control jurisdiction). This still does not mean they enjoy network-specific market power, however, since the computer software and know-how needed to develop such systems are not tied to any particular place. Whereas with respect to transport services the pressure of competition can also be achieved by selective, sequential (time-wise) hit-and-run entry (which does not necessarily result in total substitution of the established undertaking), in the case of traffic control systems auctioning competition is needed, where the subject of the auction is the predefined geographical traffic control area for a set period of time, and where the contract is awarded to the bidder who is able to offer a service at the lowest prices while at the same time covering the costs involved.

The move towards an EU infrastructure policy with the aim of developing trans-European networks automatically means interoperability must be promoted between the individual state networks. There is a great need for coordination, particularly with regard to traffic control and monitoring systems. Clearly, active competition between different providers of traffic control systems cannot work. An individual aeroplane or train can only be monitored by one institution at a time if chaos and accidents are to be avoided. Such systems are therefore natural monopolies with geographical limits that have to be clearly defined at the level of institutions. Responsibility for traffic monitoring must remain in the hands of a single authority for a set period of time. This raises the question of the “natural” limit to a regional monitoring area and the coordination needed between different areas.

---

7 Cf. Treaty of Amsterdam, Title XV, Trans-European Networks, Art. 154(2).
In the past, the railway monopoly lead to a predominantly national approach to capacity management of railway lines and timetabling influenced by national considerations. International coordination and cooperation within the International Union of Railways was therefore minimal, both in terms of standardisation efforts and route management. Optimisation efforts were confined to the national rail systems (cf. Knieps, 1995). Since then the tendency has been increasingly towards integration, including the development and introduction of a new standard European Train Control System.

Like air traffic control systems, however, train control systems have considerable potential for operating across frontiers. Competition on the European rail transport markets and the ensuing rise in demand for European rail traffic mean there is a need for systematic internalisation of cross-border restrictions. For example, the technical limits of telecommunications and radio equipment should no longer coincide with countries’ political borders. The full benefits of systems able to operate across borders need to be exploited so that competition can develop to its full potential on European rail transport markets.

The development of an integrated European train control system would be considerably facilitated by the establishment of independent agencies like those that exist for air traffic control. Insofar as no such integrated European system is being developed, steps should at least be taken to ensure full advantage is taken of the opportunities that exist for intensive coordination and harmonisation of train control systems, for example by stepping up standardisation efforts and coordination of timetabling.

Increased harmonisation of systems can be introduced in the context of a competition organised on an institutional level between national train control systems. If calls for tenders are issued throughout Europe (as is generally the case now with other services), train control agencies that are very successful in a given country are also likely to be successful with their auction bids in other countries. As result, new train control developments in one country will gradually spread to other countries. In addition, competition between institutions will probably also result in full utilisation of cost-saving opportunities and a better
supply of transport services. The monopoly of national train control agencies on information is not as strong as it was. Transport undertakings now have an opportunity to put pressure on their respective agencies.

5. Regulation of access to transport infrastructure

On account of bundling effects combined with irreversible costs, transport infrastructure (railway track, stations, airports, etc) can result in the development of network-related market power. Even in open markets, the State has an important responsibility to regulate this monopolistic bottleneck. In particular, steps must be taken to ensure that there is no abuse of this market power in order to distort active and potential competition on complementary transport markets.

5.1. Disaggregated regulation of bottlenecks versus essential facilities doctrine

The concept of essential facilities plays a central role when competition rules are used to control network-specific market power. Facilities (infrastructure or equipment) are considered “essential” if they are simultaneously:

- indispensable for reaching customers and/or enabling competitors to do business;
- not present anywhere else on the market;
- such that they cannot objectively be rebuilt at a reasonable cost.

This concept is closely related to the essential facilities doctrine that derives from American anti-trust laws and is now increasingly applied in European competition law (cf., e.g., Lipsky, Sidak, 1999; Haus, 2002; Aberle, Eisenkopf, 2002). The doctrine affirms that facilities can only be considered essential if the following two conditions are met: without access to the facilities, it is impossible in practice for competitors to enter the complementary market; and it is im-
possible for a provider on a complementary market to duplicate the facilities at a reasonable cost\(^8\) and there is no substitute for them.\(^9\)

In the context of the disaggregated approach to regulation, the essential facilities doctrine is no longer applied on a case-by-case basis, as is usual in anti-trust laws, but instead to a category of cases, namely monopolistic bottleneck facilities. The form taken by the conditions of non-discriminatory access to the essential facilities must be defined (cf. Knieps, 1997a; Knieps, 1997b). Insofar as monopolistic bottlenecks exist in liberalised network sectors, they call for specific residual regulatory measures to control the remaining market power. In particular, symmetrical access to the monopolistic bottlenecks must be guaranteed for all active and potential network service providers, so that competition stands a chance on all complementary markets.

The starting point of such a regulatory policy should be that regulatory measures are strictly confined to those network areas where market power potential actually exists. Regulation of access fees in respect of monopolistic bottlenecks must not therefore be accompanied at the same time by regulation of prices on the complementary markets for transport services (cf. Knieps, 2000, p. 100ff.). There are two further aspects that must be taken into consideration. First, it is wrong “automatically” to assume from competition on the service markets that there is no potential for market power on the infrastructure level, insofar as infrastructure fulfils the criteria governing monopolistic bottlenecks (cf. Brune-kreeft, 2003, p. 89 ff.). Second, there is the question of minimum regulation, which while being sufficient to guarantee non-discriminatory access to essential

---

8 It is not possible, for example, to operate a ferry service without access to ports.
9 For a summary, see Areeda, Hovenkamp (1988). A further criterion for the essential facilities doctrine is sometimes that shared use of the facilities is essential for competition on the complementary market because it lowers the prices there or increases the volume of services offered. However, this criterion describes only the effects of access.
facilities stops short of infringing excessively on the regulated enterprise’s property rights. 

5.2. Price-level regulation of infrastructure user charges

The effect of a total denial of access to infrastructure would also result if capacity was made available at prohibitively high prices. This alone shows that effective application of the essential facilities doctrine must be combined with appropriate regulation of access conditions. The identification of monopolistic bottlenecks is always based on an intramodal perspective, a decisive factor being the need for complementary service providers to have non-discriminatory access to such facilities. However, the existence of monopolistic bottleneck facilities does not necessarily guarantee that there will be long-term surplus profits. Firstly, there is the possibility of the “necessary case”, where even unregulated train path providers are unable to meet their costs (Berndt, Kunz, 2003, p. 207ff.). Secondly, competition between modes can severely limit an infrastructure provider’s profit potential (Fritsch, Wein, Ewers, 2003, p. 208).

Regulation of railway infrastructure access charges should always be confined to the parts of the network where market power potential actually exists. Price/profit regulation in the complementary competitive parts of the network would go against the principle of minimalist regulation and lastingly obstruct the goals of a fully open market. Regulation of infrastructure user charges must not therefore lead at the same time to regulation of prices in the complementary parts of the network where there is no market power potential.

Regulation of railway infrastructure access charges should be limited exclusively to price-capping. The basic principle underlying price-capping regulation is that price levels should be regulated in areas where there is network-specific market power. The benefits of price-capping in terms of efficiency improve-

---

10 In principle, a distinction must be made between the question of whether network-specific market power exists as a result of a monopolistic bottleneck and the question of what constitutes appropriate regulatory measures.
ments and future investment activities can only unfold if price-capping is applied in its “unadulterated” form and not combined with input-based profit regulation. Individual pricing agreements amount to over-regulation that is harmful to competition.

The reference point for monopolistic bottleneck facilities in the sense of quasi competition where the criticism of abuse of market power is not justified should be overall cost recovery. Regulatory authorities should not force undertakings to apply specific price rules, such as Ramsey prices or two-part tariffs, as this would hamper their quest for innovative pricing systems. It is always possible that better rules will be found in future.

5.3. Flexible innovative pricing structures for network access

5.3.1 Advantages of the subsidiarity principle

Intramodal competition on European transport markets requires that all transport service providers, domestic or foreign, should have non-discriminatory access to infrastructure. The criterion of non-discrimination must refer here both to the quality of the available infrastructure (avoidance of grandfather rights, etc) and the access tariffs.

Financing of transport infrastructure (roads, canals, airports and railways) used to be seen as a typical responsibility of the State. Providing capacity utilisation is so low that there is no rivalry about use, market pricing based on load factors is not appropriate.

Unlike transport infrastructure where capacity is in short supply, parts for which demand is low have all the features of a public commodity because of the lack of rivalry. Accordingly, in the case of little used transport infrastructure it is still necessary to decide on the level of investment that is socially (politically) desirable and to guarantee state funding. By themselves, however, high fixed costs for making transport infrastructure available are no grounds for unlimited state
subsidies. Above all, the scale of subsidies must not be left to chance (Scientific Advisory Board at the Federal Ministry for Transport, Construction and Housing, 1999, p. 442). Rather, what is needed is transparent implementation of policies based on the ‘Orderer Principle’, according to which lines operating at a loss, for example, may be funded via invitations to tender. More fundamental is the question of public investment policy, particularly with relation to federal infrastructure planning (Aberle, 2003, p. 453ff.).

One question that has to be asked is to what extent the goal of international intramodal competition calls for pan-European harmonisation of the pricing principles applied to infrastructure user charges. Providing the same conditions apply to all transport service providers in each country, there is no discrimination against foreigner providers (cf. Scientific Advisory Board at the Federal Ministry for Transport, Construction and Housing, 1999, p. 443). On the other hand, different infrastructure charges in different countries (ceteris paribus) lead to different transport tariffs, which in turn can encourage users to look for other solutions. It is to be expected, however, that the volume of traffic will also grow in international traffic as a result of the introduction of innovative, flexible pricing systems.

As far as possible, systems for charging for infrastructure access should ensure non-discrimination, the efficient allocation of scant infrastructure capacities (efficiency requirement), and a harmonised degree of cost-recovery (financial requirement), all at the same time. Conventional full-cost calculations based on administrative apportionment keys for assigning shared infrastructure costs to the different user groups make no sense economically and are obviously not the solution (cf. for example Baumol, Koehn, Willig, 1987). However, pricing based on marginal social costs cannot fulfil all these criteria at once either. In particular, fixed cost recovery is a residual value.

A critical feature when it comes to making infrastructure capacity available are the high fixed costs and the economies of scale that ensue with respect to service provision. Where there are economies of scale it is known that marginal cost prices no longer produce total cost efficiency. By themselves, the high fixed
costs associated with providing access to network infrastructure are no grounds for state subsidies. Owing to the sharp rise in demand for transport in recent decades infrastructure capacity is in increasingly short supply, creating a need to find free market solutions.

5.3.2 Welfare-enhancing price differentiation

Pricing principles should aim to incorporate the financing requirement as an *ex ante* condition, so that the State, which has to be brought in to mop up any shortfall, is not faced with any incalculable debt requirements. With the help of an appropriate two-part pricing system, however, consisting of a fixed use-related component (Infracard) and a variable component lower than a linear price, it is also possible to meet the underlying need for market-oriented infrastructure user charges. For this, inclusion of the demand side is vital. Tracking down pricing principles is therefore a business matter that ultimately can only be performed by the infrastructure companies themselves.

Under certain conditions price differentiation can enable an undertaking to survive, but only if a standard (uniform) price fails to guarantee the necessary cost-efficiency.\textsuperscript{11} For the national economy, a key advantage of two-part tariffs over one-part tariffs is that the goal of (partial) cost-efficiency can be achieved without severely reducing transport demand by applying major mark-ups to the variable price so that scant infrastructure capacity can be allocated efficiently. Of particular importance for the welfare-enhancing effects is the increase in the volume of traffic (more traffic) typically associated with two-part tariffs.\textsuperscript{12} Major infrastructure users (Infracard holders) can be expected to do whatever they can to make as much use of the infrastructure as possible. Moreover, small users for whom the purchase of an Infracard is not worthwhile are not excluded insofar as they also have access to infrastructure capacity.

\textsuperscript{11} If the average cost curve lies above the demand curve, price differentiation is unavoidable for full cost efficiency.

\textsuperscript{12} It should be borne in mind that environmental costs have to be met by means of separate environment policy measures (mineral oil tax, etc.).
Optional two-part tariffs have the advantage of persuading potential users to disclose information about their own particular willingness to pay (e.g. whether it is worth their while to pay a certain fixed entrance fee) and in so doing assign themselves to a particular user category.

It must be stressed here that there is no one system that is better than all the others and that could be centrally adopted as a pricing objective. Rather, it is a case of testing the limits of further price differentiation via a process of trial and error. The limit to further differentiation is when the transaction costs become too high for the system, in other words when the cost of avoiding arbitrage outweighs the benefits of graduating prices further. This limit cannot be uniformly defined, however, as it will depend on the prevailing “local” circumstances. Consequently, there is a need for a regulatory framework that does not impede infrastructure operators in their quest for new pricing structures.

6. Current reform of access to European airports

European air transport was liberalised over a decade ago. 1 January 1993 saw the entry into force of the third package of measures for liberalisation of air transport in the European Union, which largely replaced the bilateral air transport agreements signed in the past between member States and makes it possible for EU nationals to establish air transport undertakings anywhere in the European Union. The package also provided for free access to all intra-Community routes and flexible fares for the services operating on these routes.

In January 1993 the Council of the European Communities also adopted a Regulation on common rules for the allocation of slots at Community airports which established a legally binding framework applicable in all Member States. The main features of the Regulation include the maintenance of “grandfather” rights, according to which the air carrier that has operated a slot in the previous

---

scheduling period has priority over other air carriers in respect of that slot in the
next scheduling period. Unlike the exchange of slots, no provision is made for
slot trading or slot auctions (Niejahr, 1999).

For competition on European air transport markets to operate efficiently, how-
ever, non-discriminatory access to airports must be available to all active and
potential suppliers of airline services. At the same time efforts must be made to
achieve efficient allocation of scant infrastructure capacities and to cover the
costs involved.

The system used to allocate slots at busy airports in a way that distorts competi-
tion is the central question in the ongoing air transport debate (Brunekreeft,
Neuscheler, 2003, p. 254 ff.), which is also focused on revision of EEC Regula-
tion 95/93. In this context, those issues that are proving especially controversial
are the abolition of “grandfather rights” and the feasibility of market-oriented
approaches to slot allocation at saturated airports, in particular slot trading and
slot auctions and the role of optimum user charges (Boyfield, 2003, p. 34 ff.).

6.1. Slots as marketable commodities

It is an indisputable fact that many airports – not only in the USA but also in
Europe – hit their capacity limits at peak times. In view of these capacity bot-
tlenecks, which are getting worse, the public authorities are increasingly being
asked to overcome the problem by developing more capacity. Investment on
such a scale, however, would produce a surplus of slots which, economically
speaking, would be a waste of valuable resources. This is not the same as say-
ing that airport investments (generally) should stop, but rather that they should
continue only for as long as the added benefits of capacity expansion are in
keeping with the extra costs involved. This means that even in the context of
economically optimum investment, at busy airports capacity will still be in short
supply at peak times. The allocation of scarce slots is therefore not only a tran-
sitional problem: the process of transforming a public commodity into a private
one is unstoppable and irreversible.
The following two questions are examined in greater detail below:

• How can the allocation of scant airport capacities be organised in such a way that they can be put to optimum use from the point of view of air travellers (consumers)? and

• How can the allocation of such scant capacities be organised in such a way that distortions of competition on the liberalised markets for air transport can be prevented or kept to a minimum?

6.2. Economic characteristics of slots

As soon as airport capacity is no longer available in excess – i.e. as public commodities – it becomes necessary to specify and define which slots have become in short supply and when. In the usual world of trading in commodities (e.g. grain), there is the microeconomic problem of defining the different categories of a given commodity (e.g. types of grain), but the precise time of the transaction is generally not crucial. The situation regarding airport capacities is completely different. There are a great many resources that must be coordinated with each other, time-wise, to the highest possible degree of accuracy. The very definition of a take-off or landing slot opens up a vast range of alternatives that can be crucial for potential transactions. If a take-off slot, for example, means only the right of a given airline to take off within a relatively long period of time, that right is worth much less than a guarantee that the airline can take off at a specific point in time without being subject to any delays. Some airlines on the other hand may prefer flexible operating times. Trading in slots therefore presupposes that take-off and landing rights have first of all been defined in a way that reflects both the needs of the airlines (and their passengers) and the operational and logistical possibilities of the airport operators.

In 1969 limits were imposed on the number of peak-time take-off and landing rights issued at five American airports (Chicago O'Hare, Washington National, New York Kennedy, La Guardia und Newark) (high density rule). The right to take off or land during this period was referred to as a “slot”, with slots lasting
either half an hour or an hour. Slots were not classed as property and offered no guarantee of punctuality (Report of the Congress, 1995, p. 1 ff.).

The definition of slots contained in Article 2 of EEC Regulation 95/93 also leaves considerable room for manoeuvre (“the scheduled time of arrival or departure available or allocated to an aircraft movement on a specific date at an airport coordinated under the terms of this Regulation”). Here too it is ex ante coordination with no guarantee of punctuality, no rules on priority, and no means of enforcing the right to take off or land as a right of ownership.

This imprecise formulation of the right to use airport capacities is exactly what the airport operators want. There is no incentive for them to issue guarantees of punctuality (for specific flights) and accept the liability rules that would stem from such guarantees without at the same time benefiting from the scarcity rents. On the other hand, it is patently obvious that administrative management of capacities in short supply with no financial incentives for all the parties involved can produce a high degree of inefficiency.

It has become clear that slots cannot be defined independently of the market-economy instruments used in connection with the allocation system. Efficient slot allocation means making maximum use of airport capacities while still complying with the relevant safety standards. Allocation systems that take account of individual air carriers’ priority and punctuality preferences within a given slot period would have the advantage of making do on average with smaller reserve/buffer zones. For example, it would be possible to apply a system based on different slot categories, where owners of expensive slots would have higher priority at take-off than cheap slot owners, who would sometimes have longer to wait.
6.3. Abolition of “grandfather rights” in favour of ex ante auctioning of take-off and landing slots

As in the past, rather than being reallocated according to changing needs, take-off and landing slots in Europe remain in the hands of the airline to which they were initially allocated, even if that airline does not use them or another airline would put them to better use. *(Ex ante)* flight schedule coordination is carried out by the airport coordinators appointed by each individual country. The exchange and transfer of slots is allowed in the context of mutual agreements between air carriers. Voluntary airline associations worldwide also negotiate flight schedule adjustments to take account of airport capacity limitations and avoid unnecessary delays. Even if take-off times booked by airlines are in increasingly short supply, so far they have always been allocated free of charge. Economically, there is no justification for this unless there is sufficient capacity for all airlines to be able to take off and land at any one time. Otherwise, the airlines already well-established at a given airport have an asymmetrical competitive advantage over other airlines.

With *ex ante* auctioning of take-off and landing rights the advantages of long-term flight scheduling could be maintained, but at the same time the market would have to be opened up to newcomers (Wolf, 1995). The danger of stockpiling slots for the strategic purpose of gaining a competitive advantage over rival companies decreases the larger the air transport market in respect of which *ex ante* flight schedules are drawn up. Alternative routes and extensive product diversity create sufficient substitution options between different air carriers.

Given that on expiry of the auction period slots tend to return to the airport operator, the opportunities for air carriers to receive scarcity rents from the sale of slots is limited to trading in slots during an auction period. Insofar as airport operators are now party to the scarcity rents received from slots, scant airport capacities are allocated to the bidders who show the greatest willingness to pay.

---

Income from the auctions can also be ploughed back into airport development projects given that very high scarcity rents send out a signal to the economy that airport capacity is insufficient and that further airport development is required.

6.4. Potential for trading in slots

One of the essential features of Regulation 95/93 is that the Council does not challenge the existence of “grandfather rights”, with the result that a slot that has been operated by an air carrier shall entitle that air carrier to claim the same slot in the next equivalent scheduling period (Art. 8, paragraph (1)a). The Regulation also provides that carriers have an obligation to use 80% of the slots allocated to them. Preference with regard to slots allocated out of the slot pool is given to new entrants; 50% of such slots must be allocated to new entrants (Article 10, paragraph 7). According to the Regulation, slots may be freely exchanged between air carriers by mutual agreement. To date, however, there is no provision for the sale and purchase or leasing of slots.

Slots may only be exchanged by air carriers that already have a slot. The ban on compensation payments means that if no equivalent slots are available there is no incentive for an exchange, even if exchanging a slot would ultimately result in its being put to more appropriate use. It also encourages black-market trading in slots and circumvention of the trading ban. Efficient allocation would therefore be better achieved if trading in slots were officially allowed. In terms of free competition, too, trading in slots is preferable to exchanges. New entrants to the market, however, always have the opportunity to buy slots. If slots are in very short supply, with the result that the economic value of slots used for a specific purpose (e.g. scheduled business flight) is very high, slot owners must accept a considerable loss of income if they decide not to sell. In other words, in an efficient slot trading system the opportunity costs of using slots or selling them will be nearly the same.

There is also the question of to what extent slot trading favours hoarding and therefore hampers competition on the markets for air transport. Even if the de-
clared aim of the rule that unused slots are returned to the pool is to reduce stockpiling (“use-it-or-lose rule”), the possibility of such hoarding cannot be totally excluded a priori. Unlike straightforward swaps, however, the possibility of selling slots increases the opportunity costs of hoarding slots or using them for a less lucrative flight because of the scarcity rents that can be obtained from selling.

6.5. Optimum user charges based on scarcity rents

For as long as airport slots are allocated by applying “grandfather rights” rather than in auction procedures, it is unclear how far a reform of airport charges can achieve more efficient allocation of scarce capacity while at the same time improving the status quo in favour of symmetrical access conditions.

Up until now airport charges have basically depended on the weight of the aircraft, their function being to help finance the airports, not to control the way available capacities are allocated. Aircraft weight and flight distance are no indication of a flight’s (marginal) contribution to the shortage of capacity available to air traffic control authorities and airports nor of the costs that ensue for all other transport players. The decisive factor in this respect is the demand at a particular time for airport capacities and for route capacities. In the short term, airport capacities are essentially unchangeable. In the event of unforeseen bottlenecks, the typical solution is for airports to ration capacity on a first-come first-served basis.

Air carriers tend to ignore the constraints imposed on other aircraft and their passengers by an additional flight at a particular time (e.g. longer clearance times, longer delays, and longer flight times). To take these constraints into account, one solution would be to levy a (time-based) congestion fee equivalent to the congestion costs incurred by all other flights as a result of the extra flight. If demand for infrastructure capacity still exceeds supply, the solution would be to charge a market price that includes not only the congestion costs but also a scar-
city rent. These are therefore capacity bottlenecks where there is direct rivalry in respect of take-off and landing slots.

The congestion charges or scarcity rents would need to be graduated according to the degree of capacity utilisation during a day and depending on the season, insofar as capacity utilisation for the same flight may vary. This would enable peak-time take-off and landing rights to be allocated more efficiently. Congestion charges operate like peak load prices, but are not to be confused with them since a (non time-based) congestion charge would still have to be levied even if there was no change in capacity utilisation over the period and no fluctuation in the level of the congestion costs.

Another advantage of congestion charges with respect to the short-term allocation of slots is that when congestion charges are high during peak periods there is no incentive to hoard slots. Given that, unless it can be proved that 80% of the allocated slot sequences have been used, slots in Europe are returned to the slot pool, the introduction of capacity-based congestion charges also reduces the negative effects of “grandfather rights”.

6.6. Reform solutions with respect to airport take-off and landing fees

To date only a few airports levy capacity shortage-based take-off and landing fees. At London’s Heathrow and Gatwick airports landing fees based on peak load pricing have been charged since the early 1970s. In the morning and evening a standard peak landing fee applies regardless of aircraft weight. The principle of peak load pricing is also applied, however, in relation to aircraft passenger and parking fees. Insofar as peak load times for aircraft landing, passenger clearance, and parking are not the same, different peak periods are defined for each of these services, with different peak load prices.

---

Other airports have also introduced basic charges or minimum landing fees to deter smaller aircraft from using the airport at peak times. They include Toronto, Sydney and New York, as well as Frankfurt, Munich and Dusseldorf.

Airport charges based on capacity utilisation are a (partial) improvement on weight-based take-off and landing fees, although they cannot be compared with optimum congestion charges. There is still a lot of resistance that has to be overcome on the part of the different parties involved.

6.7. Disaggregated regulation of airport market power

The question that now has to be asked is whether there is a danger that with the introduction of slot auctions and scarcity rents airport operators would exploit their monopoly position (at least regionally) and in certain circumstances even reduce the number of slots up for auction, with the result that slot prices would reflect not only scarcity rents but also the airport operator’s market power. This fear cannot simply be brushed aside on the grounds that many airports are still state-owned enterprises which in any case serve to maximise public interests. Unlike the operation of flight services, airport infrastructure is bound up with irreversible costs. Once in operation, they cannot simply be transferred to another location in the way that an aircraft can. This means that the result of inefficiency or excessive airport charges will not be the construction of another airport since two airports would not be able to survive together for long. The introduction of scarcity rents and the move away from the strict principle of cost-efficiency confers a discretionary power on airport operators that needs to be regulated.

This is where the modern regulatory theory of price-capping can provide appropriate solutions. Regulation of airport market power is sometimes difficult to reconcile with user charges based on the scarcity of capacity. Since it is not

---

15 In 1998, a reform of charges levied at Boston airport with a view to significantly raising landing fees for smaller aircraft, in the interests of more efficient capacity utilisation, met with such fierce resistance that it had to be abandoned.
possible to expand airport capacity in the short term, the levying of optimum user charges at airports where capacity is in persistently short supply gives rise to scarcity rents which are not necessarily compatible with a given regulatory restriction.

7. Reform of access to the European rail network

Even if the markets for rail transport were not initially in the forefront of the deregulation debate, competition on these markets is now also seen as a central coordination instrument. Of particular importance in this context was the package of Directives adopted on the 26 February 2001 by the European Parliament and the Council on the allocation of railway infrastructure capacity and the charging of infrastructure fees, building on the earlier Directives adopted in 1991 and 1995.

The provision of rail services requires not only trains but also access to a network of railway lines. Train control systems are also needed, not only to ensure traffic safety (traffic control), but also to ensure real-time train path management. Efficient competition on the markets for rail transport is conditional upon non-discriminatory access to the rail network for all active and potential providers of railway services.

7.1. Non-discriminatory access to railway infrastructure

Railway infrastructure (unlike rail services) is characterised by a monopolistic bottleneck situation insofar as infrastructure operators have a natural monopoly and the construction of railway lines involves irreversible costs.

---

16 Directives 2001/12/EC, 2001/13/EC, 2001/14/EC.
17 Directive 91/440/EEC.
18 Directive 95/19/EC.
Insofar as monopolistic bottlenecks exist in network sectors, there is a need for specific residual regulation to control the remaining market power. In particular, all active and potential providers of network services must be guaranteed symmetrical access to the monopolistic bottlenecks, so that competition can operate properly on all complementary markets.

7.2. Efficient use of train paths based on scarcity prices

In the past, railway infrastructure capacity was traditionally allocated with the help of administrative measures defined by the national railway monopolies (e.g. timetable conferences, priority rules for determining train sequences in the event of delays, and discretionary one-off measures with respect to train control). There were no user charges to reflect the scarcity of capacity, even if at certain times of the day or year there were severe bottlenecks on certain sections of line. The decision when to use a particular line section therefore had no impact on fares, with the result that there was no incentive for peak-time customers to switch to less busy times. Customers who lay great store by punctuality and were willing to pay for it had no possibility of travelling in trains guaranteed to arrive on time.

One way of overcoming this problem is to levy a (time-based) scarcity price for using the railway network. The right to operate trains on busy sections of line could be auctioned among the different operators. Undertakings wishing to operate a train on a busy section would then have to pay a market price that reflects the opportunity costs of using these capacities, whereby it is possible that a freight train with traditionally low priority might be prepared to pay more than a traditionally higher priority Intercity train to make sure certain just-in-time production processes do not come to a standstill. If railway operators then pass these efficient access fees onto the user, customers must be required to pay higher fares at peak times and lower fares at less busy times. The function of this peak load pricing system is to manage available capacity.
7.3. Price differentiation versus discrimination

Even on busy routes optimum user charges still may not enable all costs to be recovered. Economies of scale with respect to the construction of railway infrastructure are such that optimum access charges are unable to cover the infrastructure investment costs. This raises the question of how to finance the shortfall as well as the related question of the *ex ante* politically determined degree of cost-recovery. To ensure that the incentives for achieving the necessary cost-recovery are credible for the infrastructure operator, the degree of total cost-recovery must not be left to chance (*ex post*), and nor must it therefore be subject to constant scrutiny in terms of its level and validity.

The goal of efficient allocation of train path capacities in a context where cost-recovery is limited calls for price differentiation strategies. In terms of access to train paths of different qualities, price differentiation has to take account of differences in the load factor, as reflected in differing degrees of willingness to pay (price elasticity of demand for rail infrastructure capacities). In particular this means that differences in train path prices are due not only to differences in the cost of the different qualities of train path available but also to the different additional price components that have to be included to cover the fixed infrastructure costs (cf. Berndt, Kunz, 2003, p. 195 ff.).

In addition to a variable use-based component there can also be a fixed pricing component (Infracard). The legal concept of discrimination must not be used to prohibit economically desirable price differentiation as such. This can be a danger, not least because another name in English for the economic concept of price differentiation, which is neutral in terms of competition, is “price-discrimination”. Price differentiation means that differences in price are not solely due to differences in the costs directly attributable to a given customer but also take account of demand factors.

---

19 Two-part tariffs were levied by Deutsche Bahn AG in the context of the 1998 route pricing system (cf. Knieps, 1998).
Bibliography


Als Diskussionsbeiträge des
Instituts für Verkehrswissenschaft und Regionalpolitik
Albert-Ludwigs-Universität Freiburg i. Br.
sind zuletzt erschienen:


82. **A. Gabelmann**: Monopolistische Bottlenecks versus wettbewerbsfähige Bereiche im Telekommunikationssektor, Dezember 2001


88. **G. Knieps**: Does the system of letter conveyance constitute a bottleneck resource? erscheint in: Proceedings of the 7th Königswinter Seminar „Contestability and Barriers to Entry in Postal Markets“, November 17-19, 2002


90. **H.-J. Weiß**: Die Doppelrolle der Kommunen im ÖPNV, in: Internationales Verkehrswe- sen, Jg. 55 (2003), Nr. 7+8 (Juli/Aug.), S. 338-342


95. **G. Knieps**: Neuere Entwicklungen in der Verkehrsökonomie: Der disaggregierte Ansatz, erscheint in: Schriften der Nordrhein-Westfälischen Akademie der Wissenschaften, Schöningh-Verlag, Paderborn, 2004


98. **G. Knieps**: Verkehrsinfrastruktur, erscheint in: Handwörterbuch der Raumordnung (HWB) der ARL, 2004

99. **G. Knieps**: Limits to the (De-)Regulation of Transport Services, erscheint in: EMCT Round Table 129, Paris, 2004