

VTI notat 2A-2003

# **Marginal cost pricing of airport use**

The case for using market mechanisms for slot pricing

## **Pricing the use of Sweden's railways**

Are charges in line with marginal costs?

|                   |                     |
|-------------------|---------------------|
| Author            | Jan-Eric Nilsson    |
| Research division | Transport Economics |
| Project number    | 92020               |
| Project name      | MC-FCAM             |
| Sponsor           | EU                  |

| <b>Contents</b>   | <b>Page</b> |
|---|-------------|
| <b>Summary and Abstract</b>                                   | <b>3</b>    |
| <b>Marginal cost pricing of airport use;</b>                  | <b>5</b>    |
| <b>The case for using market mechanisms for slot pricing</b>  | <b>5</b>    |
| 1 Introduction  | 5           |
| 2 What is special about airport infrastructure?               | 5           |
| 3 Relevant objectives   | 7           |
| 4 Current procedures for capacity allocation                  | 8           |
| 5 Implications of current allocation procedures               | 10          |
| 6 Using a market mechanism                                    | 12          |
| 6.1 A benchmark auction mechanism                             | 12          |
| 6.2 Other aspects   | 15          |
| 7 Conclusions   | 17          |
| <b>References</b>   | <b>18</b>   |
| <b>Pricing the use of Sweden's railways;</b>                  | <b>19</b>   |
| <b>Are charges in line with marginal costs?</b>               | <b>21</b>   |
| 1 Introduction  | 21          |
| 2 The original reorganisation and the subsequent deregulation | 21          |
| 3 Marginal costs and present track user charges               | 23          |
| 3.1 Wear & tear   | 24          |
| 3.2 External effects  | 26          |
| 3.3 Conclusions   | 27          |
| 4 Scarcity  | 29          |
| 4.1 Modelling the scarcity problem                            | 29          |
| 4.2 Testing   | 33          |
| 4.3 Conclusions   | 35          |
| 5 Conclusions   | 36          |
| <b>References</b>   | <b>37</b>   |

## Summary and Abstract

As an input for the EU-funded MC-ICAM project, two background papers have been prepared, one for Task 5.3 (air) and the other for Task 5.3 (rail). These two documents are presented here as one single VTI publication.

**Abstract (air):** The marginal cost for use of airport capacity includes congestion at major airports. Based on a review of the shortcomings of an administered committee allocation of take-off and landing rights, this paper argues for using an auction mechanism instead. One particular such mechanism is presented in some detail in order to provide substance to the concept; much more work would be needed to elaborate on the details for a mechanism that could fit a European context.

**Abstract (rail):** Since 1988, Sweden's railways have been vertically separated with (private or public) train operators paying for the use of government-owned railway infrastructure. The present paper scrutinizes today's charging regime. The revenue from these charges generates insufficient revenue to recover the total spending on infrastructure. The charges may, never the less, be in line with an efficiency-enhancing pricing policy. To answer whether or not this is so, the paper seeks to compare state-of-the-art knowledge about marginal costs for using infrastructure with current charges. It is found that there may be reason to increase charging on at least two counts: Current tariffs fail to mirror marginal reinvestment cost and scarcity is not priced.

# **Marginal cost pricing of airport use; The case for using market mechanisms for slot pricing<sup>1</sup>**

## **1 Introduction**

The overall purpose of this task is to identify and assess the main barriers to the implementation of marginal cost pricing in airport use and to recommend strategies to overcome these barriers. The present text focuses one specific cost aspect, viz. the marginal capacity cost for using existing capacity at major airports. We therefore disregard possible problems with respect to cost recovery etc. of airports that have sufficient capacity during most of the day.

In the same way as in the rest of the transport sector, congested airport capacity need to be properly priced in order to ascertain its efficient use. The way in which these costs, and the consequent prices, should be calculated however differs from the way in which capacity shortages can be dealt with in the road sector. We could think about congestion pricing for road use as a price scheme announced before the driver embarks on a trip into a crowded area. It is also reasonable to envisage this scheme as a successive number of iterations to adjust price to the actual consequences it has on traffic flows. In contrast, we will argue that airport congestion should be dealt with by way of auctioning slots to airlines operating on an airport, a slot being the amount of capacity needed for an aircraft to take off or land.

There are several important motives behind this position and these will be further elaborated in section 2. At the core of the argument is, however, the problems related to making a deregulated market for aviation function properly, in particular in view of entry problems for new operators at capacity constrained airports. Section 3 briefly summarises the efficiency aspects that have to be observed when we think about airport capacity allocation, section 4 presents the way in which capacity is allocated today and section 5 the efficiency concerns that these “club meetings” provoke. Section 6 elaborates on one particular design of the allocation process suggested more than twenty years ago as well as on possibilities to further update this approach to using a market mechanism. Section 7 finally hints at the way in which congested airports could be organised or regulated in order to ascertain that appropriate efficiency concerns are taken into account in the day-to-day operation of the facilities.

## **2 What is special about airport infrastructure?**

Airports are beset by several types of problems which can give rise to economic inefficiency (for more detail, see Bruzelius, 1997). First, as long as demand for airport services does not add up to capacity, the airport is a natural monopoly, with the standard type of efficiency issues that this gives rise to. Secondly, the (potential) efficiency problems are reinforced by that larger airports often are dominated by one airline. Third, the demand for airport services is not generated from the demand for transport between only two airports, but between a great

---

<sup>1</sup> I am grateful for comments on a previous draft from Nicole Adler.

number of different airports. As a consequence, the size of some of the markets between the studied airport and other airports is so small that the continuity assumption typically employed in economics analysis is not relevant. And fourth, larger airports are often congested, thus augmenting the natural monopoly problem to include also the appropriate use of existing, and construction of new capacity.

The services provided by an airport include four components; the passenger and the freight terminals – often separate buildings – the air traffic control system and the runways. Emphasis in the present paper is on runway capacity but the discussion readily bears over also to other parts of capacity if this is relevant.

The costs for holding infrastructure available to users comprise two components, spending in order to construct the necessary facilities and the cost for maintaining them and keeping them available for use by airlines. We think of the average cost for using a facility as the sum of these (appropriately discounted) components over the lifetime of the facility divided by the total number of users over this period of time. In contrast, the marginal cost signals the cost implications of admitting additional take-offs or landings. This may comprise some impact on runway wear of additional movements. But primarily the marginal cost refers to the impact on other users from the marginal flight, in particular what consequences it has on their respective businesses.

In one specific situation one could believe that the operator of an airport is capable to allocate capacity between different flights efficiently; this would happen if the industry was vertically integrated in the same way as was the case for railways until a few years ago. In particular, only one operator would use the facilities. The integrated organisation would have reason to take joint care of interests both of its airline operations and the costs for holding infrastructure available. If an airport instead is run by a profit maximising (natural) monopoly organisation and used by competing operators, it is less obvious that the monopolist's priorities would establish an efficient price level. The standard projection is rather that prices then would be too high, flights too few and incentives for capacity expansion poor.

Avoiding this is possibly one reason for that airports in many countries are operated by some type of national or regional agency. Prices are designed in order to make operations break even rather than to maximise profits. Airport agencies are not left with the responsibility to allocate access to runways. Rather, this is primarily the responsibility of the users; the details of how these user conferences settle the allocation problem are explained in section 4 below.

Although parts of the industry now are deregulated, in particular with respect to inter-European operations, it is not subjected to the same kind of competitive forces that for example characterises the road haulage industry. One major impediment to competition is that scarce runway capacity makes entry difficult. Airport congestion is therefore not only of interest from the point of view of the airport itself and the efficiency in using its facilities; it is of equal importance to the functioning of the air transport industry at large. While the main question in focus here is allocative efficiency, i.e. how to ensure that airport infrastructure capacity is efficiently provided and used, this issue must also be seen in a wider perspective. A proper way to deal with airport capacity is as a means to improve the functioning of the entire air transport sector.

### 3 Relevant objectives

The criteria to be applied for the evaluation of various means to organise and deal with airport activities are the same as when processes for the allocation of scarce economic goods are considered. The criteria are adopted in recognition of the fact that airport capacity is a critical resource in the operation of an airline. Based on Grether et al. (1989) the following standards can be used to evaluate the performance of alternative processes for capacity allocation.

Equity, here understood to mean that flight services to small communities are given due consideration by the allocation method.

- (Static) efficiency, ultimately measured in terms of the value delivered to the consuming public from a resource base. Since an airport slot is critical to the operation of air service, the efficiency of the air transportation system is dictated almost directly by the efficiency of the air transportation system the process engenders. Such a measurement at both the airport or independent market level and the air transport systems level is natural. It comprises the following components.
- Carrier expansion and contraction: Within a given market, carriers with efficient marginal operations (relative to the marginal operations of other carriers) should expand. Since a slot is a critical resource for flight, those carriers which have relatively high profit opportunities for additional slots should be allocated additional slots by the system. The argument bears over also to a potential entrant since it is similar to an expanding firm. As a corollary, the slots for expanding carriers should come from carriers whose marginal operations are the least efficient; marginally unprofitable services, and less successful operators, should therefore be induced to exit from the business by the selected mechanism.
- Coordination: Carriers have some time-frame latitudes within which operations can be shifted. This means that gains from exchange can be achieved through proper coordination of carrier services. Any system of slot allocation should foster such coordination of service and capacity use patterns.
- System-level efficiency: There exist interdependencies among carrier operations at different airports since any flight must involve (at least) two airports. A system of slot allocations operative at many different airports must be capable of capturing the efficiency gains and reductions in the overall costs which can result from proper coordination among airports.
- Susceptibility to monopoly and/or collusion. Since slots are a critical resource, their allocation dictates the pattern of competition. In any market the control of slots could provide a key for the development and enforcement of anticompetitive practices. Therefore the process of allocating slots should have adequate safeguards to prevent these possibilities.
- Dynamic efficiency or long-run industry growth. Without additional capacity the industry cannot expand. Yet, capacity expansion necessarily absorbs valuable resources. One measure of the need for capacity expansion is the value created by additional slots. If such values, when integrated over time, exceed the cost of expansion, then capacity expansions are in order. The allocation procedure should be supportive to this end. In situations where there are absolute barriers to capacity

expansion, the need to ascertain efficiency in the allocation of the slots that after all are available is even more acute.

- Transaction costs. The allocation procedure should be reasonably cheap to administer, including the oversight that is necessary to minimise the risk for collusion etc.

#### **4 Current procedures for capacity allocation**

To focus the discussion, we think about each airport as being a separate accounting unit, in principle required to break even.<sup>2</sup> This means that we pay no attention to the large number of regional airports that receive different types of financial support to handle their business. Often, the airport is owned by some public sector agency or the local community. There are also important examples of privately owned airports.

To raise revenue, the airport authority designs a more or less complex charging scheme for use of the facilities. In Sweden, the scheme comprises five components. A Landing Charge is collected per ton of an aircraft's maximum weight at start. A Terminal Navigation Charge is paid for navigation services at each airport. A charge per departing passenger is supposed to recover costs for common areas in airport terminals and all sorts of handling of luggage etc. A Security Charge is also levied per departing passenger in order to care for costs for safety controls. Finally, based on an aircraft's starting weight, an Underway Charge is levied relative to distance flown by the plane.

It has already been noted that charges typically are designed in order for the airport authority to break even. In this, parking areas, licence agreements with sales activities located in the terminals etc. are also important sources of revenue. The charging schemes typically include flat charges meaning that there is no price difference to make use of airport facilities during the day, the week or over the season. One exception seems to be the British Airport Authority – being under regulatory control by the Civil Aviation Authority – which makes use of a more elaborate differentiation scheme (Bruzelius, 1997).

In Sweden, airport charges amount to about SKr 175–200 (€ 19–21) for an international ticket. International comparisons indicate that aeronautical charges relating to ground activities on average amount to about 7 percent of airlines' costs, and that en route charges add some 5 percent to this.

Any charge for infrastructure use will have implications for total demand for capacity. Although they comprise only part of the airlines' total cost, charges for using airport facilities obviously have an impact on ticket prices for travellers and freight consignors. Consequently, they also have an impact on the extent of travel, and as a result also for the extent of take-offs and landings at each airport; although price elasticity may be low, the higher is the charge, the lower is demand.

Standard charges are, however, not at a level that automatically clears the market on the major airports in focus here. These typically witness shorter or longer periods of the day with demand in excess of available capacity. Airlines don't get access to the number of rights to take off or land – to the number of slots

---

<sup>2</sup> The fact that for instance Britain has collected its main airports in one organisation does not affect the present line of reasoning.

– that they demand at existing price levels. The degree of this problem differs for different airports but at least five of Europe's airports witness a high level of demand during much of the day.

Since charges don't clear the market, as would be the case in the delivery of many "normal" private goods such as hotel facilities etc., the industry has developed alternative means to handle the conflicts of interest. The procedures date back to a system created by IATA in 1947 which over the years have been gradually updated. The system is administrated by airport authorities, but it is the airlines themselves that have the last word about the allocation of available slots. The IATA rules have subsequently been codified by a EC directive (Council Regulation (EEC) No 95/93). In Sweden, the rules are applicable primarily for one airport, namely Stockholm/Arlanda.

The directive states that arrival and departure slots shall be allocated by a coordinator for an airport when demand exceeds available supply. The coordinator is instructed to handle the task in a neutral, impartial and transparent way. The coordinator is supposed to participate at the international time-tabling conferences that take place twice per year, arranged by IATA. Except for one representative of each airport, spokespersons for all airlines take part in the meeting; a recent gathering comprised delegates from some 260 airlines and the total number of delegates was about 1 100.

Before these meetings, airlines submit their (confidential) demand for slots to the coordinator. Using the principles established by the Directive, these wishes are compiled and a draft proposal about the number and timing of slots allocated is sent back to each operator before the opening of the conference. A basic principle that steers this first allotment is that existing slots are grand-fathered, meaning that an airline that flies in a certain way during the current timetable is given priority to the same slot in the next time-table, if this is what it wants.

A slot has to be used at least 80 percent of the time during the period in question, if the historical right to the slot during the equivalent next season is to be protected (the use-it-or-lose-it rule). A grandfathered slot is therefore not the property of its incumbent owner. The Directive also provides for the creation of a slot pool into which newly created slots through increase in hourly scheduling limits, slots returned either voluntarily or under the use-it-or-lose-it rule and slots otherwise unclaimed by anyone are placed. At least 50 percent of all slots in the pool shall be earmarked for new entrants.

During the conference, the initial proposals are discussed between operators in order to establish a final solution to the shortage problem. It is therefore not the coordinator but the operators that have the last word about the allocation. This will of course in turn shape the behaviour of the coordinator in the creation of the draft proposal. After that an allocation has been established there are also ongoing bilateral negotiations between airlines that seek to establish mutually beneficial exchanges of slots.

## 5 Implications of current allocation procedures

The procedure to allocate slots between competing uses that has developed over the years can be understood in the light of the economic theory of clubs. The generic definition of a club is a voluntary group of individuals who derive mutual benefit from sharing production costs, the member's characteristics or a good characterized by excludable benefits. A *club good* is the sharing of an excludable (rivalrous) public good.

When club decisions can be represented as a cooperative action, the resulting outcome will be a Pareto optimum for the club members; members belong since they perceive a net benefit from membership. If not, they would exit. If the club decisions are to account for the well being both of those within the club and those outside, then the outcome will be a Pareto optimum for the economy as a whole. Even if there is no such concerted action to take account of all members of a society, a core – i.e. a Pareto efficient – solution can result if the population is partitioned among a set of clubs so that each individual is a member of an optimally configured club in terms of membership size and provision. (Cf. further part IV in Cornes & Sandler, 1996.)

Airports represent an example of a multiproduct club where commercial and non-commercial travellers and freight operators utilize the terminal, the air traffic control system, the runways and specially designed freight terminals.<sup>3</sup> Commercial travellers rely on the first three club goods, private flight operations on air traffic control and runways only and freight airlines use the last three products of the club. The question is then what the equilibrium outcome from this specific club could be expected to be. A first observation is that the possibility to start new clubs if supply is saturated at an existing facility is limited. Alternatives can at least be considered to be inferior to the preferred alternative which typically is an international hub (London/Heathrow, Amsterdam/Schipol, etc.). We could therefore not expect that efficiency can be achieved by way of saturating demand by expanding capacity in this way.

If we consider the decision making of a club at one specific airport in isolation, a further concern is whether or not it could be conceived of as a (fully) cooperative solution. If we partition members only into two categories – large and small, or incumbent and entrant airlines – it is not obvious that they are handled as equals when decisions are to be made. The small or entrant firm can choose to accept the allocation that is suggested by the other club members, not since they conceive of it as overall beneficial but since the alternative – not to receive any slot at all – could be worse. Not least the feature of grandfathering provides incumbent or large operators with more clout in the committee meetings than they would have if all would start from scratch and try to find a solution to the benefit of them all. This is so since the behaviour of club members is strongly driven by the rule in force if members can not arrive at a common conclusion, which here is the grandfathering principle.

It is therefore not clear that the cooperative solution could be thought of as a (fully) Pareto optimal allocation between operators that actually belong to the club. The club solution is even more socially problematic if the problem of becoming a member is considered; the position of (traditional) flag carriers at national hubs is very strong compared to that of prospective entrants. Moreover,

---

<sup>3</sup> Lipsman & Sandler (1996) handle the multiproduct aspect of transport infrastructure pricing.

the haggling at international gatherings provides incumbents with a long history in the business with even more clout relative to an upstart. In short, there are few reasons to believe that the scheduling committees of major airports allocate slots in ways promote efficiency in the way discussed above.

The relative position of incumbents and newcomers may furthermore jeopardize the dynamic efficiency of the market as a whole. There are reasons to believe that there are scale and scope economies in the business, not least generated by the extensive hub-and-spoke production systems (Horner & O'Kelley, 2001, Adler & Berechman, 2001). This does not ascertain that a hub-and-spoke system necessarily is the fully optimal way of providing these services (Pels et al. 2000). But irrespective of this, it is not straightforward to disentangle the actual powers of these economies from the importance of having a first mover position at international hubs. Put in another way, in spite of possible cost economies for incumbent operators it could become easier for entrants to challenge them if access to scarce airport capacity was allocated on different grounds. To the extent that companies with aspirations to enter or expand at a given airport do not have the opportunity to compete for these resources, or are denied the use of them on the same bases as are other established companies, one of the major sources of competitive pressure cannot be operative in this market.

The ongoing consolidation in the industry and the creation of alliances could also be explained, at least partially, by the difficulty to get access to airport capacity. A maverick operator who runs an odd type of service with poor scale economies could after all have a business concept that is worth testing on the market. The possibility to do so is severely hampered by the problems with getting access to airport facilities under present allocation techniques. This will then hold back the ability of market innovations and further strengthen the position of incumbent behemoths.

From the perspective of intertemporal efficiency, the club solution has an additional drawback. If a market would be set up for airport capacity, it is reason to believe that the price for airport slots would differ over the day. It would be higher at peak than at off peak. A high peak price means that at least part of the demand is induced to use airport capacity at off peak, i.e. demand is rescheduled over the day. The very size of the peak premium that is charged does also work as an investment signal; at high enough price levels the operator earns even more by expanding capacity.<sup>4</sup> This is not so with the administrative procedure. Rather than providing incentives to reallocate demand over the day, operators have reason to ask for additional capacity. (Demand for) over investment is therefore a complementary drawback of this approach.

In Sweden, this means that the construction of a third runway at Arlanda, which is to be opened in the summer of 2002, could have been postponed if a pricing device would have been put in operation. The discussion is also to start already today to project a fourth runway which – it is estimated – would have to be built within another ten years time, at least if current slot allocation procedures are maintained.

---

<sup>4</sup> This is at least so if the market is competitive since an overly high charge would otherwise induce entry into the business. Airports represent a less competitive activity so an uncontrolled monopoly supplier of the services may have incentives to under invest in order to hold up prices and earn supranormal profits.

## 6 Using a market mechanism

A first and most obvious approach to handle airport congestion would be to use the standard procedure of a differentiated price scheme, posted well before an allocation decision is to be taken. To simplify, we can think about a two-level structure with a higher price for peak, and a lower for off peak periods; the concept of course generalises up to any number of periods. A first scheme proposal would be posted and the operators would react accordingly. The airport authority could then see whether or not excess demand has been eliminated. If not, the authority could raise the peak price for the subsequent time-table period; if the price is so high so that capacity is not fully used the price would have to be reduced.<sup>5</sup>

This is the standard way for a Walrasian tâtonnement equilibrium process to work. In this particular instance it could be cumbersome; since there are no more than two time-table periods per year, it could take long to establish an equilibrium solution. Meanwhile, administrative allocation decisions would still have to be taken for the excess demand situations. Alternatively, the process could be conceived of as taking place before each time-table period. But it would then be more straightforward to ask operators to quote a price for the flight(s) they want to operate, rather than having the administrator test different price levels. We have then moved over to an auction-type mechanism.

Research into the economic theory of auctions has virtually exploded over the last few decades. It is still reason to go back here to the seminal work relating to slot allocation at airports. In 1979 the US Civil Aeronautics Board and the Federal Aviation Administration commissioned a study to find appropriate ways to deal with shortages in airport capacity. The subsequent report was not made generally available until 10 years later; Grether et al. (1989) was suitably published in a series referred to as *Underground Classics in Economics*.

The main recommendations of that report will be summarised in section 6.1. We then add some observations that have bearings to the more recent literature in the field (6.2).

### 6.1 A benchmark auction mechanism

The suggested mechanism comprises the following important features. We think about a primary market for slots organised as a *sealed-bid, one price auction*. Each potential buyer submits a bid for each unit desired indicating the maximum price the buyer is committed to pay. The bids are then arrayed from highest to lowest. If  $x$  units are to be auctioned, then the highest  $x$  bids are accepted. The price paid by each of the winning bidders is the value of the lowest accepted bid.

The simple logic of Figure 1 provides the intuition for how the mechanism works. We think about the bidding for a carefully specified part of a day at an airport; slots could be allocated for this day only or for a typical Friday (etc.) during the whole scheduling period or whatever measure of supply period-of-time found appropriate. It is assumed that carrier A submits a single bid for 450; carrier B submits three bids of 3000, 700 and 400; carrier C submits two bids of 550 and 425; and carrier D submits three bids of 1 500, 500 and 350.

---

<sup>5</sup> Jones et al. (1993) seem to advocate a solution of this nature.

Suppose that six slots are available for this period. The slot that was last accepted would then go to bidder A while bidders B, C and D would each get one slot less than they want. All six slots that are allotted would go at the price 450, even though (for instance) carrier B submitted the bid 3 000.

This particular set of rules has the feature that the optimum bidding strategy is for each buyer to bid close to the maximum that he/she is willing to pay. This value is of course directly related to the profits the flight will generate. As a result, the economic circumstances are reflected immediately and accurately in the market; slots are allocated to those airlines that value them highest. For carriers it means that the profits from their most profitable flights are protected – they will not be dissipated for slot acquisitions; the highest bids do not determine price. Price is determined by the lowest accepted bid and therefore by the least profitable flight in the market.

The original proposal suggested that the sealed-bid auction would be supplemented by an *aftermarket*. The reason was that the sealed-bid auctions could be applied only to one airport at a time. The aftermarket would then handle the coordination between airports. Since much has happened in this field, we will return to the issue of coordination between airports below.

Many feel that service to *small communities* will be terminated if slots are allocated by a market process. Markets can, however, be organised in ways which will prevent this from happening. One way to do so is to establish a restricted market for small communities. Out of a total number of slots, a certain number would be restricted for this use. Only buyers with special status could participate in the bidding for these slots. The special status could alternatively be based upon the origin or destination of the flight, passenger classification of aircrafts (commuter vs. general aviation etc.) or the size of the aircraft could be used as a proxy.

Both the primary and the secondary markets should be organised in an identical way. The consequent outcome would be two different prices. Classified services that are believed to have a value for society except for what is revealed by market demand would come away cheaper than other aviation services.

A general comment is in order about how markets could be expected to work, not least relative to the possibility of small operators, communities etc. to get access to slots. A common belief seems to be that the market price of slots will be bid up by the higher density routes and that smaller-scale services will be squeezed out. This can clearly happen but in assessing its likelihood two aspects must be kept in mind.

For a small city to lose its link to a major hub the **most** profitable service that could be scheduled from the small city to the hub must be **less** profitable than the **least** profitable services along the higher density routes. The relevant statistics to assess the probability for this to happen is to compare marginal rather than average profits from the respective markets.

In deregulated markets, rates can be adjusted to reflect higher costs per passenger. Since demand elasticity from small cities is likely to be low, it is not at all clear that they will lose in a market competition for slots. They may, on the other hand, have to pay higher ticket prices for the possibility to retain services.

Certain types of aircraft use more “capacity” than do other aircraft. An increase in the share of “heavies” would then result in a loss of airport capacity. This could be handled by a flexible way to deal with *slot definitions*. Operations that have characteristics with place disproportional demand on capacity require

more slots than other operations and it would become more costly to bid for them. The auction could handle this by making it necessary to bid on two slots for large aircraft or to make the precise definition part of the mechanism.

*Disposition of funds.* Funds generated by the sale of slots should be used to defray the cost of removing the binding airport capacity constraints. Many possibilities exist, including the establishment of satellite airports, but almost all of them require funding. The sale of slots provides a natural and economically efficient way of recovering the costs.

The importance of this “use of funds” provision needs to be emphasized. The agents who receive the funds might be able to act as monopolists, restricting capacity in order to drive up the price of slots. The agent could then enjoy the benefits of the revenues by diverting them to other uses. Such behaviour would substantially hamper the efficiency gains from the policy proposed here.

In the event that the funds cannot be used to expand capacity, they should be used to encourage off-peak traffic. This can be done through a negative lottery of the sealed-bid, one-price type. Carriers are then allowed to bid negatively, thereby indicating the subsidy it would take to entice them to provide off-peak services. This process is similar from the sealed **offer** process described above; each carrier operating at a subsidised hour would receive a subsidy equal to the lowest accepted (negative) bid. The number of such bids accepted would be governed by the availability of funds.

*Antimonopoly policies:* It is difficult to see how a carrier successfully could utilize an auction process to monopolise an airport. Even collusion is difficult in an auction since the auction rules could be designed so that neither winners nor bids are announced. The large number of airlines that normally operate at an airport would per se jeopardize the success of any collusion attempt. A possible aftermarket for trading could also be organised so that neither buyer nor seller of slots need to know the identity of each other. But monopoly or predatory behaviour is especially difficult since the act of driving up slot prices to prevent competition necessarily uses up all the presumed monopoly profits. Furthermore, with the above structure, the funds would be destined for capacity expansion which would further undermine any monopolistic tendencies.

Monopolistic tendencies, if they occur, could also be managed by the use of complementary rules. A monopoly would be effective only if it could withhold supply. In the case of airports this would mean that large proportions of slots go unused or that they are used for operations which do not involve many passengers. Revenues from several of the operations would then not cover the price paid for slots. The existing use-it-or-lose-it rule could make attempts to pre-empt competitors very costly.

These proposals were accompanied by meticulous experimental testing. An economic experiment, which today typically is run in a computerised environment, is basically a way to test (economic) hypotheses in a semi-real environment before the rules under scrutiny are put to large-scale use. It makes use of experiment subjects, typically students, which represent the actors of the transactions that are to be evaluated. Subjects receive real revenue relative to their performance during the experiments, thus mimicking the incentives of real actors. Much more detail about procedures and results of this testing is available in the original Grether et al (1989) report.

In this particular case, repeated experiments were organised to test the performance of the market scheme described above relative to a negotiated

committee solution corresponding to what was described in a previous section. Needless to say, the committee process was inferior to the recommended process on all economic criteria that have been enumerated. From the carrier point of view the committee process might, however, be preferable. A resource which they have been receiving for free would become costly under the recommended process. Carriers that perceive themselves as having a strong political base may also prefer the committee process since they would be reasonably certain to get slots in this way. For carriers that perceive themselves as being marginal within the industry, the administrative procedure might be preferable since they understand that a pricing procedure would push them out of business. This provides ample illustration of the motives for the industry to resist the introduction of any market-like process.

## **6.2 Other aspects**

The market mechanism was never implemented, the main reason being – as described in the preface to the Grether et al report – the fierce resistance from airlines. From April 1986, a different type of market based mechanism, known as the “buy-sell-rule”, has however been in use at four major US airports (Kennedy, LaGuardia, O’Hare and National). The baseline way to allocate capacity is still via committee procedures based on grandfathered rights. The new rule lays down that, in addition, any person is authorised to purchase, sell, trade or lease slots. It is thus feasible not only to buy and sell the right to a particular slot permanently but also to lease them on a temporary basis.

The market is overseen by the Federal Aviation Administration (FAA) which lays down certain conditions governing the way the rule operates. The most important of these are the “ring fences” that mean air carrier slots can be traded without restriction while commuter slots cannot be bought by the larger carriers and all other slots, including those for international flights, are excluded from the trading market. Slots not used for a stipulated minimum of time in a two-month period are to be returned to the FAA; carriers must “use-or-lose” their slot. Surrendered slots, or others becoming available, are assigned to a pool and reallocated using a lottery but with 25 percent initially offered to new entrants. More detail about the rules is available in Starkie (1994).

Grether et al (1989) was one among a few pioneering contributions to the use of economic experiments in assessing the qualities of specific allocation mechanisms. In the same way as in the field of auction theory, and often in a complementary manner, research into the field has exploded. Much of this material is collected in two volumes edited by Klemperer (2000).

With one exception (cf. below), little research has been done in the field of airport slot allocation after the Grether et al report. The flood of research in the field of auction or mechanism design at large however opens a wide array of possibilities, should a market-near solution be considered again. It is not the purpose of the present paper to draft the details of such work, but it is reason to mention a couple of developments that may have implications for the design as compared to the 30 year old mechanism suggested by Grether et al.<sup>6</sup>

---

6 Additional aspects of this issue are available in DotEcon (2001), a report commissioned by HM Treasury together with the Department of the Environment, Transport and the Regions. DotEcon – a consultancy – recommends the British government to consider auctioning of airport slots and also suggests a design for it.

Commuter traffic to regional airports is widely considered to have a value to the regions beyond the bids that operators could be conceived to submit. This motivated the separation of slots for commuter traffic from other slots in the Grether et al. model; standard commercial aviation could in this way not bid commuter services out of the market. An alternative way to deal with social aspects relative to the items sold was used in the US sales of radio wave frequencies that took off in 1993 and subsequently have allocated hundreds of licenses, generating hundreds of billion US\$ in revenue. Bidders that were thought to represent certain pinpointed communities in society were given “preferential treatment” in these auctions, meaning that their bids were given an extra weight. This increased the chance that the group could win a license at a non-prohibitive cost; cf. further McMillan (1994). A corresponding technique could of course also be applied in the slot auction context.

An auction for radio wave frequencies included the parallel sale of several similar while not identical items, i.e. one and the same frequency for several regions. In order for prospective buyers to make use of possible scale or scope economies, or to allow them to bid on one frequency when the cost of another had climbed up too high, a simultaneous ascending auction format was used. In contrast to the Grether et al. mechanism which was based on a one-shot, sealed bid procedure, subjects could successively raise their bids, and all markets were kept open until bidding had ceased on each and every of the licenses; cf. also McAfee & McMillan (1996) for further detail.

A similar approach could be used in the context of slot allocation at several airports. The international gatherings could thus be replaced by an international data grid, comprising all bidders on all airports connected over the Internet (which of course was the technique used in the airwave auctions, i.e. with bidders submitting and revising bids from their home offices). The possibility of augmenting the one-airport mechanism with a technique that handles also the fact that both departure and arrival airports are congested was analysed in a paper by Rassenti et al. (1982). More recent experiments with time-tabling of railway operations have also demonstrated that optimisation techniques are available to deal with the extremely complex problem of allocating track capacity (Nilsson, 2000). In comparison, an international grid comprising all congested airports in the world would be a technically “small” problem to handle.

But there are, of course, other barriers to implementing this type of scheme, except for the technical complexity of an auctioning scheme. In particular, much international air travel is still regulated by strict, bilateral agreements between countries. This, on the other hand, has less to do with technicalities, more with the interest from governments in open skies or – which often seems to be the other side of the same coin – the eagerness to support flag carriers.

## **7 Conclusions**

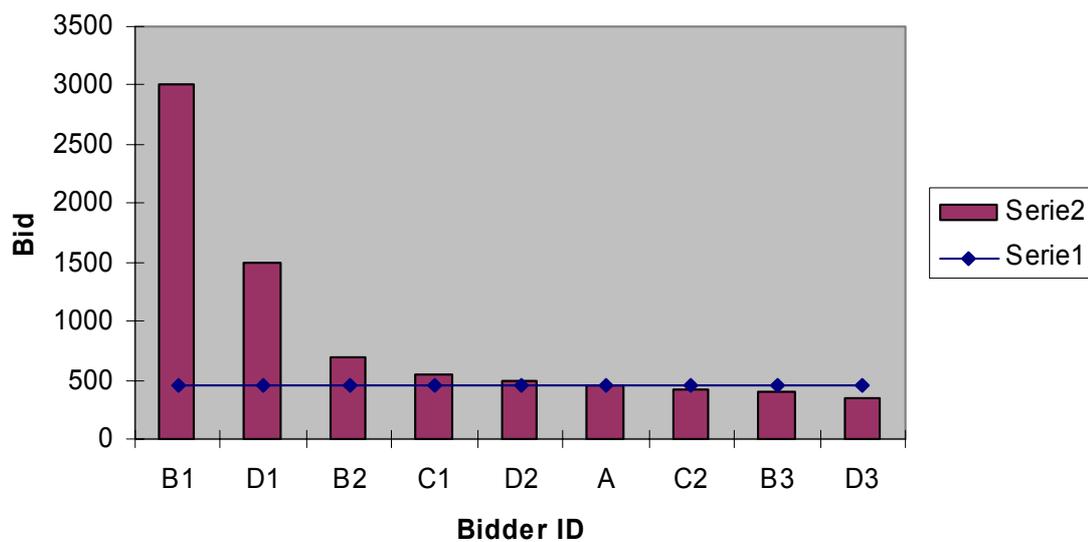
The present paper has argued strongly in favour of an active attitude towards pricing airport congestion at major airports. In particular, we have advocated the use of auctioning techniques as the means to establish the price of take-off and landing slots. The main argument behind the proposal comes from within the industry: In order to keep the airline industry reasonably competitive, it is of strategic importance to provide access to scarce airport capacity on equal grounds for all airlines, large and small, incumbents and entrants.

Auctioning of capacity can not be seen in isolation from the organisation of the industry in other dimensions. It is difficult to envisage how an auctioning scheme would work in practice without further considering the way in which airports are organised at large; should they be owned by one type or another of public agency, or should they be private?; what role should profit maximising play for the organisation?; and should some sort of price regulation scheme be considered for a profit maximising monopolist? It is beyond the purpose of the present paper to address such issues. The interested reader is, however, advised to consult further the proposals forwarded under the common heading of the New Regulatory Economics, for instance Laffont (1994) or Laffont & Tirole (1993).

## References

- Adler, N & Berechman, J (2001): **Evaluating Optimal Multi-Hub Networks in a Deregulated Aviation Market with an Application to Western Europe**, *Transportation Research A*, 35 (5), pp. 1-18.
- Bruzelius, N (1997): **The Airport Problem. An Economic Analysis of Scarce Runway Capacity**. KFB-report 1997:6.
- Cornes, R & Sandler, T (1998): **The Theory of Externalities, Public Goods and Club Goods**. Cambridge University Press.
- DotEcon (2001): **Auctioning Airport Slots**. A report for the HM Treasury and the Department of the Environment, Transport and the Regions. Available at [www.dotecon.com](http://www.dotecon.com)
- Grether, D & Isaac, M & Plott, C (1989): **The Allocation of Scarce Resources. Experimental Economics and the Problem of Allocating Airport Slots**. *Underground Classics in Economics*, Westview Press.
- Horner, M W & O'Kelly, M E (2001): **Embedding economies of scale concepts for hub network design**. *Journal of Transport Geography* 9, 255-265.
- Jones, I & Viehoff, I & Marks, P (1993): **The Economics of Airport Slots**. *Fiscal Studies*, vol. 14, no. 4, pp. 37-57.
- Klemperer, P (2000): **The Economic Theory of Auctions**. Vol I and II. Edvard Elgar Publishing Ltd.
- Laffont, J-J & Tirole, J (1993): **A theory of incentives in procurement and regulation**. MIT Press.
- Laffont, J-J (1994): **The New Economics of Regulation Ten Years After**. *Econometrica* 62(3), May, pages 507-37.
- Lipsman, M & Sandler, T (1996): **A Multiproduct Club Approach to Transportation Infrastructure Pricing**. *Public Finance* 51(4), pages 453-72.
- McAfee, P & McMillan, J (1996): **Analyzing the Airwaves Auction**. *Journal of Economic Perspectives* 10(1), Winter 1996, pages 159-75.
- McMillan, J (1994): **Selling Spectrum Rights**. *Journal of Economic Perspectives*, Summer 1994, 8, 145-162.
- Nilsson, J-E (2000): **Towards a Welfare Enhancing Process to Manage Railway Infrastructure Access**. Forthcoming in *Transportation Research*.
- Pels, E., Nijkamp, P & Rietveld, P. (2000). **A note on the optimality of airline networks**. *Economics Letters* 69, pp. 429-434.
- Rassenti, S & Smith, V & Bulfin, R (1982): **A Combinational Auction Mechanism for Airport Time Slot Allocation**. *Bell Journal of Economics* 13(2), Autumn 1982, pages 402-17.
- Starkie, D (1994): **Developments in Transport Policy: The U.S. Market in Airport Slots**. *Journal of Transport Economics and Policy*; 28(3), September 1994, pages 325-29.

Bilaga 1  
Sid 1 (1)



# Pricing the use of Sweden's railways; Are charges in line with marginal costs? \*

## 1 Introduction

In 1988, Sweden was the first country in the world to vertically separate its railway sector.<sup>7</sup> The state owned monopolist was broken up into two parts; *Banverket* – the Swedish National Rail Administration – with responsibility for infrastructure, and *Statens Järnvägar* (SJ) running railway services, still under a monopoly franchise. Subsequent events, some of which will be reviewed below, have almost totally deregulated the freight market while entry still is restricted for (most) interregional passenger services.

The purpose of this paper is to present the current structure of charges for infrastructure use and to contrast this against the norm of marginal cost pricing; how, and to what extent does Sweden's charges differ from the norm. This is done against a background presentation of the original separation and the subsequent chain of events (section 2). Emphasis is on previous, and in particular the existing pricing schemes in section 3, in particular of whether or not these charges coincide with marginal costs. Section 4 presents work-in-progress regarding the possibility to charge for track capacity shortages, a component of marginal costs often discussed but rarely handled in reality. Section 5 concludes.

## 2 The Original Reorganisation and the Subsequent Deregulation

By the late 1950ties, Sweden's railways were essentially state owned and operated by one single company, *Statens Järnvägar*. Government control over activities was strict. The operator could not decide about rate changes, rates had to be publicly announced, also for freight operations, and line closures had to acquire consent from the Parliament. The 1963 Transport Policy Act started a series of management changes that by the mid-1980ties had transferred SJ into a corporation essentially operating on commercial conditions, while still state owned.

The organisational reform in 1988 was prompted by the worsening performance of the organisation. After that state subsidies were first introduced in the early 1960ties, the need for support had escalated over the years. When a major financial reconstruction in 1985 shortly afterwards proved to be insufficient to turn around performance, the patience of the government was over, and a more radical transformation was drafted.

The 1988 reform, part of a comprehensive Transport Policy Act of that year, was given three specific motives. One was to put railways on an equal footing with roads by organisationally separating infrastructure from service operations. This is the reason for making *Banverket* a government agency, operated in the same way as the National Road Administration.

---

\* I am grateful for comments on a previous draft from Bryan Matthews.

<sup>7</sup> Strictly speaking, this may not be so; there is reason to believe that several operators made joint use of a common infrastructure during the first decades of the industry, but evidence is not available.

Since railways were considered a uniquely safe and environmentally friendly means of transport the Parliament – secondly – voted for continued financial support so that these special benefits could be fully realized. The third given reason was to transfer the responsibility for commercially unviable traffic over secondary, low-density lines to regional transport authorities. This would then be a means to carry on with operations for regional policy reasons. Nilsson (1995) describes the reform in further detail.

The post-1988 system makes use of two subsidy techniques. The bulk of the support is directed through *Banverket*. While it charges operators for track use, the revenue raised in this way was never intended to fully recover costs for maintaining the system. Technically, the government allocates an annual appropriation to the agency, also taking the expected revenue from track user charges into account. Except for maintenance, appropriations have also come to include huge allocations for investment purposes. This originates from a perception at the time of the reform that railway infrastructure had been severely neglected over the extended period of deficit years and now needed upgrading in order to be on equal footing with roads.

The second means for allocating subsidies is through the regional transport authorities. Since the late 1970ties, one authority in each county has had the monopoly franchise to run subsidised local and intra-regional bus transport, using in-house resources. During the 1980ties bus services gradually came to be subject to competitive procurement. In exchange for a lump sum transfer guaranteed for a 10-year period, the 1988 Act transferred the responsibility also for regional rail transport to the relevant regional body that subsequently has had to procure railway services on the market. In addition, the government has come to subsidise some interregional passenger services that otherwise would have been closed down. To this end a separate state agency, *Rikstrafiken*, was set up in 1998.

The 1988 reform was substantially an attempt to deal with a recurring financial deficit by way of organisational restructuring. The preparations for the reforms were meagre and no background documents analysing alternative strategies etc. was presented. In particular, the possibility of deregulating the industry was never discussed in the Transport Policy Act – which after all deregulated both taxi services and domestic air traffic. This also means that alternative organisational structures never were discussed. The only organised opposition came from SJ which from the beginning of the new system and for several years onwards still favoured a unified industry.

The first move towards market entry was, however, not far away.<sup>8</sup> Already in 1989 the first competitive procurement of regional train services resulted in a four-year contract being awarded to a private operator. *BK Tåg*, at that time an operator of couch services, submitted the lowest bid and could start operations in one of the regions in 1990. This was then in effect the first example of competition **for** the tracks where one operator is given the full control of a certain class of services for some pre-determined period of time.

The first years of the 1990ties also saw the establishment of several small-scale freight operators. Most of these are sub-contractors to SJ on peripheral parts of the network. From July 1996, anyone “fit, willing and able” can run freight services over the network, competing with SJ for contracts with consignors. This

---

<sup>8</sup> The chain of events taking place during the 1990ties is thoroughly reviewed, and a series of penetrating case studies is presented in Alexandersson et al (2000).

change, and the other gradual changes that have taken place in the sector after the 1988 reform, have not been part of any grand scheme; the post-reform development has, never the less, been better prepared than the original change was. When this is written in early 2002, only long-distance passenger services run by SJ on a commercial basis operate on a monopoly franchise. This is the last heritage of the previously monopolised railway market.

During the 1990ties several private firms other than *BK Tåg* have competed for and won contracts for non-commercial services. The most notable transfer took place in 1998 when the contract for commuter transport in the greater Stockholm area was awarded to an international consortium comprising *BK Tåg*, the French Via GTI and the British Go Ahead Group. When the new operator commenced traffic in January 2000, it had had problems to recruit previous SJ drivers, and it took several months until services were running satisfactorily again. From this day, SJ is no longer the major operator of short-distance passenger services in the country and the consortium accounts for about 15 percent of the total number of passenger kilometres in Sweden.

The gradual transformation of the industry has also led to further reconstruction of the financially battered, previous monopolist. During this period, SJ has been run as *affärsverk*, a state business administration. This is not a separate judicial body but formally part of the public sector, administering assets on behalf of the government, but doing so based on commercial principles. From 2001, three independent limited liability corporations have been formed, one running passenger services (SJ AB), one in charge of freight transport (SJ Green Cargo AB) and the third – AB Swedecarrier – being a holding company for real estate assets, heavy maintenance, etc. All stock is still owned by the government, but the new organisation can now be sold to other investors, wholly or partially. The new organisational format inter alia means that the previous monopolist has been turned into two separate claimants for track capacity that in 2001/02 has to compete for track access with another 20 firms.<sup>9</sup>

### **3 Marginal costs and present track user charges**

Although persistent deficits were one reason for the 1988 reform, policy-makers realised that substantial parts of the network would have to be abandoned if the industry was forced to fully recover its own cost. (Full) cost recovery was therefore of secondary importance in the 1988 organisational separation. This sentiment – although implicit – still permeates the political attitude towards charges for railway use.

In this context, a simplistic version of the (short-run) marginal cost pricing paradigm fits well in with the common way of thinking around railways. The model suggests that if the industry's costs fall with increasing train volumes, if there is no shortage of capacity and if there is no need to charge for external effects, efficiency-enhancing charges for using existing infrastructure should be set below average costs. This version of the reasoning legitimises financial deficits as a means to ascertain that sunk investment costs are efficiently used. The vertical separation of the previously monolithic organisation is then a convenient tool for subsidising infrastructure which is believed to be the main

---

<sup>9</sup> An assessment of the 1988 reform and the subsequent deregulation is made in Nilsson (2002).

reason for the decreasing-cost structure. This then legitimises that railway infrastructure is administered according to the same logic that since long has been in use for the road sector.

The first version of the pricing regime was constructed as a multi-part tariff. A fixed annual charge was levied per vehicle – different for different vehicles – and in addition several variable components of the tariff generated revenue relative to gross ton km or train km run, etc.; cf. Hansson & Nilsson (1991) for more detail. The projection was that revenue and infrastructure costs would add up to SEK 890 and 1 854 million, respectively (prop. 1987/88:50 p. 133), the difference being the projected need for subsidies. Except for this, a promise was made to invest massively in infrastructure, and to finance these costs over the public budget.

Table 1 (cf. back of paper) summarises spending on infrastructure and revenue from infrastructure use from the beginning of the new structure. From the beginning, the industry's deficit has come to be larger than originally projected. Primarily, this is due to the expansion of investment spending. Much of the reason for this lies outside the railway sector and could be seen as political compromises made by minority governments.

The charging system was revised in February 1999. One part of the transformation was to abandon the fixed charges per vehicle. This would offset the consequences of a previous change of road user charges, i.e. it would tilt the balance between modes back. Today, charges therefore only include components related to infrastructure use.

Table 2 provides an overview of the current structure of infrastructure charges. All charges apply equally across the network and across type of rolling stock; no allowance is therefore made to take possible geographical or vehicle-related cost differences into account. The rest of this section makes a systematic comparison of these charges and a system that would fully take social marginal costs into account. Wear and tear costs are discussed in sub-section 3.1, the (pigouvian) charges for sector external costs in 3.2 while 3.3 concludes.

### **3.1 Wear & tear**

A trackage fee (*spåravgift*) of 0.28 öre per gross ton km is paid by freight trains. The charging level is based on thorough analyses of the impact of rail traffic on the costs for maintaining the infrastructure, both in Sweden and in Finland. The idea in this work has been to relate track maintenance costs to traffic, measured as gross ton, and a number of technical parameters (track length, number of switches, tunnels and bridges, etc.) in order to establish whether or not there is any variation in spending with variations in traffic load. The analysis makes use of *Banverket* records where each piece of information is attributable to a track unit, which is a well-specified section of the network. Observations from some 160 track units from each of the years 1994–96 have been compiled into one comprehensive data set and analysed econometrically. The data set does not include costs that are reported to be common for several track units; the district, region or main office costs are not expected to vary with track usage. Since it is not obvious how to measure traffic load in station areas, costs related to usage of this part of the network are not part of the analysis.

*Table 2 Charges for using Swedish tracks per January 2000. SFS 1998:1827<sup>10</sup>*

| Type of charge                           | SEK                     |
|--|-------------------------|
| Trackage fee, passenger                  | 0,0086 per gross ton km |
| Trackage fee, freight                    | 0,0028 per gross ton km |
| Information fee, passenger               | 0,002 per gross ton km  |
| Accident charge, passenger               | 1,10 per train km       |
| Accident charge, freight                 | 0,55 per train km       |
| Diesel charge                            | 0,31 per litre diesel   |
| Shunting charge                          | 4 per shunted car       |
| Charge for using Öresund bridge, freight | 2 325 per passage       |

Initial results indicated that spending on tracks was linearly (Cobb-Douglas) related to train traffic and that the elasticity was only 0.2. A 10-percent increase of traffic would therefore boost costs with about 2 percent; cf. further Johansson & Nilsson (1998). Marginal costs were derived from this primary observation and the results from this study formed the basis for introducing the charge of SEK 0,0028 in 1999. In view of the elasticity estimate this charge would only recover some 20 percent of those maintenance costs that have been included in the analysis.

An updated version of the analysis has aggregated the data for the three years into one single dataset. It is then demonstrated that a trans-log function better describes the link between track use and costs. The curvature is, however, small and the resulting elasticity is 0.17. The marginal cost is, however, this time estimated to be SEK 0.0012 and we can not even be certain that the marginal cost differs from zero; cf. Johansson & Nilsson (2001). The paper also reports of a parallel analysis of Finnish data of a nature very similar to the Swedish material. The results basically corroborate the findings based on Swedish numbers.

We can draw at least two conclusions from this brief review. The first is that marginal wear & tear costs from train traffic is low, at least if Nordic data are representative in a larger context. The second is that the current charge for infrastructure wear in Sweden is not way off the mark.

The trackage fee for passenger trains is of 0.86 öre per gross ton km, three times higher than for freight trains. The difference goes back to the Swedish-Danish deal over the Öresund Bridge. In this, it was stipulated that train traffic was to pay an annual lump sum of SEK 50 million for using the bridge. When it was to be opened in 1999, SJ declared that its passenger services could not bear this cost; the incumbent would not run commuter or long-distance services over the bridge if it would have to pay its share. The government did, however, not accept to foot the bill for passenger services. The solution was to charge freight services SEK 2 325 per train and passage over the bridge but to bake the cost for passenger services into the trackage fee. All passenger services in the country therefore now contribute to generate the revenue required by the agreement. This eloquently illustrates the blend of motives behind the charging structure.

So far we have addressed the immediate consequences of additional train traffic for track maintenance and its implications for the charging scheme. A complementary while more implicit aspect is the extent of variation in reinvestment spending with variations in traffic load. Although empirical analyses

---

<sup>10</sup> €1 corresponds to about SEK 9.

of railway sector data on this issue don't exist, the logic of the problem is straightforward. Any (road- or) railway that is to be used over long periods of time has to be renewed with certain time intervals. The infrastructure deteriorates because of climate conditions and because of use. Variations in use – i.e. in traffic load – will therefore also have implications for the timing of future reinvestment spending. The present-value cost for advancing or delaying future spending dates is therefore a marginal cost that should be made part of the charging scheme.

Lindberg (2002) discusses this cost aspect for a road sector application. The discussion leads to the familiar expression  $MC_{\text{average}} = -\varepsilon AC$  where  $-1 < -\varepsilon \leq 0$ , by Newbery (1988) referred to as the *fundamental theorem*.  $MC_{\text{average}}$  is the marginal reinvestment cost for an average section of the network and  $\varepsilon$  is the elasticity of the reinvestment date for variations in traffic volume; if all deterioration is due to traffic  $\varepsilon = -1$  and the marginal cost is equal to the average cost. If not – that is, if there is an independent ageing factor that affects reinvestment time – the marginal is below the average cost.

Sweden does not levy any charge which accounts for marginal reinvestment costs. On the other hand, all passenger operations pay an information charge (*trafikantinformationsavgift*) of 0.2 öre per gross ton km. This has nothing to do with marginal costs but is a way to recover costs for information services to passengers, an activity that was transferred from SJ to *Banverket* in 1999. Considering the 1999 traffic volume, this charge would generate about SEK 30 million that year; in contrast, the spending on reinvestment was SEK 520 million this same year. Even if less than half of this expenditure was to be included as a marginal cost charge, current charging would still be far below the efficient level.

Infrastructure charges also include a fee of SEK 4 per car shunted at one of the twenty-something major shunting yards. The logic behind this fee is marginal costs, the suggestion being that the handling of cars on a marshalling yard in order to set together new, or to split up trains generates wear on switches and other installations. Since this equipment is part of the infrastructure it is also relevant to be included in an efficiency enhancing payment system. The empirical substance behind the charging level is, however, poor. The quality of cost data, about what installations that are available at each yard and in particular information about the number of cars handled is incomplete and partially incorrect. Moreover, the number of yards is low, making any regression statistically difficult to perform.

### 3.2 External effects

The original 1988 accident charge was based on a cost allocation principle, the total (external) accident cost<sup>11</sup> being averaged over total number of train kilometres driven. In a 1997 proposal submitted to the government, *Banverket* took the position that the costs for road/rail accidents were irrelevant for charging, the reason being the judicial principle that a road vehicle is culpable in any incidents with railway vehicles. The agency also advocated that accidents occurring since people (illegally) had been walking on the tracks should not be priced as an accident externality. As a consequence of this proposal, the accident charge was substantially reduced. The differentiation between freight and

---

<sup>11</sup> Total cost is computed by multiplying the social cost per average accident, including a willingness-to-pay component for reducing accident risks, with the number of injuries and casualties over a year.

passenger vehicles, the former paying twice as much as the latter per train km, is based on that freight services conduct a larger proportion of their transports on those parts of the network that are more exposed to accidents.

There are two problems related to the current level of the accident charge. The first is that the rate is based on average rather than marginal costs. There are indications of that additional train traffic – controlling for the extent of safety installations – actually reduces accident risk (Lindberg, 2002). If this is correct, the current charge is – *ceteris paribus* – too high. But secondly, and as described above, the current charge does not include certain classes of accidents. The legalistic motive given for this has poor relevance for calculating social marginal costs; irrespective of if people have been culpable in the judicial sense, the presence of railway traffic does *per se* give rise to an accident. This would then, *ceteris paribus*, motivate a rise of the charge. The net consequences of these offsetting aspects are not yet clear, but it is not probable that the optimal charge would be considerably higher than today.

Except for the above fees, diesel trains – which accounts for about 10 percent of total train kilometres driven – pay SEK 0.31 per litre fuel used, i.e. a payment in addition to the price the operator has to pay for the fuel *per se*. This is a means to internalise otherwise external environmental costs. The current charge has not been adjusted since it was first introduced in 1988. To ascertain that emission costs should be priced in the same way as when environmental benefits of investment projects are calculated in cost benefit analyses a 1997 *Banverket* report demonstrated that the rate should be raised between five and ten times. The charges in place from January 1999 were never adjusted to take this into account. One reason may have been regional policy concerns; much diesel traffic is operated in remote parts of the country and upward adjustments of the charging level might impact negatively on the possibility to retain these services.

Subsequent work has demonstrated that there may even be reason to consider raising the charge even further. The primary uncertainty around the appropriate level of the charge concerns the costs for reducing CO<sub>2</sub> emissions to levels that accord with the Kyoto protocol. Irrespective of which, the current emission charge on diesel trains is way too low.

An additional concern of second best nature has been raised with respect to the railways' use of electricity. The marginal production of electric energy seems to emanate from power stations with coal burners, where the environmental consequences of burning coal are not fully internalised. There may therefore be an argument to levy a supplementary charge on the railway's electricity use in order to compensate for this inability, at least during the winter season when these burners are being fully used. SIKÅ (2002) argues along these lines.

### **3.3 Conclusions**

The review suggests that current charges for infrastructure use are lower than they should be to enhance an efficient use of railway infrastructure. The financial net, presented in Table 1, certainly demonstrate a deficit, but this is not the reason for the conclusion. Rather, it is first and foremost the failure to include reinvestment costs into the charging regime that warrants the conclusion. The present review does not make it feasible to say whether or not this omission in absolute numbers is larger or smaller than the surcharge on passenger services emanating from costs for using the Öresund Bridge.

The surcharge on diesel use also seems to be below the efficient level. This is less a concern for the industry's activities at large since the charge would, even if raised, mean little in the aggregate; a small share of traffic is run by diesel locomotives and the diesel cost is a small share of the operators' total operating costs. A price hike might, however, have implications for the way in which shunting operations are organised and in particular for services on non-electrified secondary railways. We may (also) for this reason have more operations on secondary lines than would in principle be efficient. The final conclusion on the efficiency implications of this issue would also have to account for whether trucks on adjacent rural roads are fully charged in this respect (Nilsson 1992 analyses second best investment and pricing policies in the road and railway sectors when the use of road infrastructure is not appropriately charged for).

By and large, charges seem to be lower than would be efficient. The counterfactual question – what would the development in the sector have been if charges had been higher – is of course not feasible to answer. Figure 1 provides an image of how railway traffic has developed over a longer period. The peak year for traffic – the year in which traffic reached its highest level – has been chosen as norm for the comparison. The figure demonstrates that both passenger (passenger km) and freight (net ton km) transport today is at or above its overall peak; in this perspective railways is in a healthy mode. This is done with some 20 percent less train km than at the peak year, i.e. the industry transports the same or larger volumes with a smaller production apparatus. It is, however, not feasible to detect whether the 1988 separation has made a difference for traffic growth, or – which is what matters – what traffic volumes would have been under some other organisational regime.

For the assessment to be even-handed, it should however be reiterated that efficiency neither has been, nor is today the sole concern of successive governments; concerns over regional policy as well as safety and environmental benefits seems to be of similar relevance. The evaluation should therefore try to address the question whether the chosen pricing strategy has been efficient, given the constraint that railway services if anything should be expanded and that the network should not be allowed to shrink. This issue is not further dealt with in the present paper.

Over the years, the domestic debate over the marginal cost pricing concept has been poor. One reason may be that the policy has served well to provide legitimacy to the post-1988 policy of keeping track user charges low; one implication is that the average charge for using infrastructure in the late 1990ties was SEK 3.50 per train km in Sweden while Denmark charged SEK 31.50 and Germany SEK 48 per train km (SJ, 1998). The prime beneficiaries of the policy – the operator and the customers – have been pleased. The general public has had to foot the bill, but the picture that has been provided here of major financial deficits is not well known.

## 4 Scarcity

Like most countries, Sweden does not price its scarce track capacity. There are at least two efficiency implications of this. First, and most obvious, a scarcity charge would raise the average price level for infrastructure use, thus depressing the demand for track access. Second, and less obvious but probably more problematic, an inability to charge for capacity shortages generates risks for that traffic operations are prioritised in the wrong way.

A parallel with a holiday resort illustrates the argument. The price for a peak week at the hotel is typically (much) higher than the price charged during off season. Absent the higher peak charge, the management would have to make use of some administrative technique to ration demand, with the risk that guests who value the accommodation highest would not get priority.<sup>12</sup> In a similar way, current (administrative) techniques to allocate track capacity may fail to trade off wishes from different operators against each other in an appropriate way. The consequence is that the industry as well as the society at large would benefit from introducing smarter techniques for handling capacity shortages.

There are several reasons for why scarcity charges are not in use. The most obvious is, however, the lack of appropriate charging techniques. In particular, to handle the consequences of insufficient track capacity delicate tradeoffs have to be made between competing wishes, and a solution at one bottleneck may have implications for the way in which conflicts are solved elsewhere and at other times of the day. Averaging over scarcity in a larger network, which may be natural to do on congested roads, generates a risk to send incorrect signals to operators and to depress demand at large rather than to induce operators to adjust their wishes. The present section will present a research project that has addressed this problem (4.1) and describe some of the tests that have been made of the modelling approach (4.2) before drawing some conclusions (4.3).

As a matter of clarification it should be observed that scarcity is here distinguished from congestion. A congested situation may occur once a certain timetable has been established. One possible reason is that an operator's train is delayed and that other operators as a consequence may have also their trains delayed. This is a sort of externality which probably could be handled by introducing different sorts of carrot-and-stick systems which remunerates and/or punishes behaviour in order to make operators internalise all consequences of their actions. In contrast, we here deal with scarcity, technically solved by establishing a time table.

### 4.1 Modelling the scarcity problem

Adding one train to, or taking one train away from an existing time-table may have consequences for the way in which other trains can be run. Scarcity costs in the railway network are defined to be the deviation from individually ideal timetables caused by insufficient availability of infrastructure capacity. The ideal timetable is thus the departure-arrival pattern that would prevail if everyone could run their services in the way they want, without interruption from other operators.

---

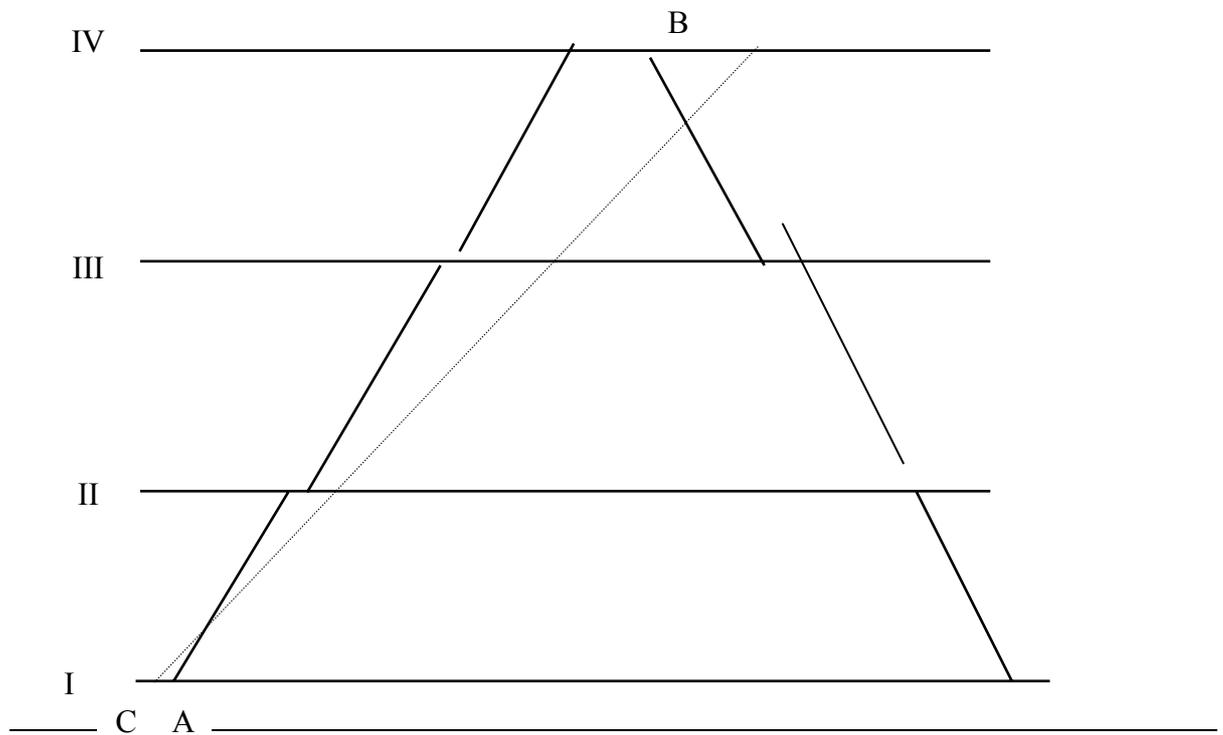
<sup>12</sup> The prime motive for price differentiation is to earn more than with a flat charge. This does not invalidate the present argument.

Scarcity is manifested either in that trains have to depart at other times than they would prefer or in that they are completely closed off from operation.

In order to deal with the scarcity problem, the qualities of supply of, and demand for track capacity must be carefully modelled. Figure 2 illustrates the situation by way of a simplified string diagram for a single-track line. Each station is depicted by a horizontal line, a station being a place where trains may meet or overtake each other; here, there are four such stations (I- -IV). A horizontal move along the line (on the x-axle) represents a movement in time. A string between two stations can therefore be used to describe a train, and the steeper is the string, the faster is the train.

Train A leaves station I at some specific point of time (say 07:00). It stops at station II (say at 07:16), leaves again (at 07:18) and arrives at the terminal station IV (at 07:46). The movement of train A from stations I to IV represents a route comprising a set of consecutive blocks, a **block** being a section of a railway line that can be occupied by at most one train at a time. On the single-track line in our example this is the stretch between meeting-stations I and II, II and III etc. Train B describes another route of a train going in the other direction. If train B is the same physical rolling stock as train A the two represents two **legs** of a round trip. Using these specifications we can establish some features of track capacity **supply**:

- When one train makes use of a block, others cannot use it at the same time; there can be zero or one train at a time on each block.
- A block is much longer than the train. Even if demand for access to a common infrastructure emanates from a fairly small number of train operators, a capacity shortage may occur.
- Once built, supply-adjustments to accommodate higher demand are lumpy, costly and take a long time before being operational.



**Figure 2** A string diagram depicting the movement of three trains over a certain network.

**Demand** for track access is derived from the demand by passengers and freight customers to travel or to dispatch freight loads between nodes in the network. The following aspects of track access demand can be observed.

- Each departure time typically has (a large number of) substitutes, but each of these may have a different economic value: While a passenger service operator prefers to leave at 7:00 it may be feasible, although less valuable, to leave at 6:45 or 7:15 or sometimes in-between instead. Since train operators have to meet customer demand there are strict boundaries on this flexibility; passengers are simply not interested to leave earlier than, say, 6:30 and a set of freight cars would not be loaded and ready to leave before some given point of time.
- There may be cyclical variations in demand. This creates peak periods where available capacity is insufficient and slack periods with under-utilised capacity. In conjunction with the previous point, demand cyclicity can be used to overcome capacity shortages by way of inducing operators with larger flexibility than others to adjust their departure times.
- There may be complementarities in the demand for track access. Going from Here to There could be valuable only if it is possible to proceed to Elsewhere at an appropriate point of time. The allocation process must therefore be designed so that an operator can co-ordinate (different legs of) its services, and also to facilitate coordination between different operations.

- Services of different operators may also be each other's substitutes. Two firms running the same type of train over a section of the network may value track access differently depending on how close to each other the respective routes over the network are laid.

These demand and supply aspects must be taken into account when an efficient means to handle the time-tabling problem is to be developed. The presence of an operator of (the slower) train C in Figure 2 illustrates the nature of this problem. Train C wants access in a way which is not compatible with the wishes of the operator(s) of trains A and B; crossing strings outside stations indicate prohibitive conflicts. The question is whether train C or the combination of A and B should be given priority, or indeed if one or more of them could be induced to adjust their demand so that all three can be licensed.

In order to create a train schedule that gets the maximum benefits out of existing infrastructure, two separate but strongly related analytical challenges have to be addressed. First, the **optimisation problem** is concerned with the mathematical aspects inherent in the problem. Second, the need to make operators reveal their value of track access is referred to as the **incentive problem**. Knowledge about each party's value of access must, for instance, be available to strike the balance between trains A and B in Figure 2 on the one hand, and C on the other. If not, it may be technically feasible to establish solutions, but the actual value of the trade-offs that have (implicitly) been made could be way off the value-maximising resolution of the conflict.

Based on these considerations a research project has developed an auctioning mechanism for allocating track capacity.<sup>13</sup> Operators interested in running train services are asked to articulate their demand a bit more than they do with present techniques. In particular, they don't only have to describe the ideal departure-arrival pattern, but also define feasible deviations from the ideal departure for each and every service they want to operate. In addition, they need to specify a value or a bid function over these departure times. The ideal departure time from station A is perhaps 07:00 but it would be feasible to leave anytime between 06:40 and 07:20. The ideal departure may be worth 100, but this value could taper off to zero at the outer boundaries of the feasible schedule.

Demand data of this nature is fed into an optimisation programme that establishes the value-maximising solution to all wishes. Operators are informed about what departure time that they are allocated – if any – and what they would have to pay in order to run their train in this way; they may then revise their demand specification. The process goes on as long as anyone would want to make changes of their demand scheme.

The process provides a solution to the capacity allocation problem and at the same time generates a set of prices. This is the congestion charge that would make up a separate component of a comprehensive pricing scheme, in the same way as congestion charges for using scarce road capacity in major conurbations is

---

<sup>13</sup> The auctioning process and the possibilities to use different auctioning mechanisms have been analysed by way of techniques from experimental economics in Brewer & Plott (1996), Nilsson, 1999 and Isacsson & Nilsson, 2002. A comprehensive description of the process is given in Nilsson (2002) and the optimisation technique that has been used is presented in Brännlund et al. (1998).

indispensable to ascertain efficiency in road use. We can note a few qualities of this way to establish a price for capacity use:

- The charge is made operational by specifying the payment for each specific origin-destination pair.
- The level of the charge depends on the level of demand relative to available capacity in each specific part of the network. A certain A-to-B train may have to pay different prices during different parts of the day. Another train between B and C which in all other respects is identical to the first (type of rolling stock, length of trip etc.) may have to pay a different charge.
- A railway network, and the links between trains making use of it, is an intricate cobweb of interactions. Even if only very few operators would be using tracks it would be very difficult to coordinate demand in order to fool the system.

## 4.2 Testing

Neither in Sweden, nor in any place known of, a system of decentralised pricing for capacity shortages in a railway network has been put to use. The system described above has, however, been subject to a sequence of desk-top tests, the most recent during the winter of 2001/02.

A network with outer boundaries given by the stations Storvik–Borlänge–Frövi–Västerås–Uppsala has been programmed (cf. Figure 3; back of document). This includes a specification of the running time of a number of specific trains over precisely these lines, taking all speed-affecting idiosyncrasies of the lines into account. There are also rules and restrictions about compulsory stopping times etc. when trains meet. This programming is done to incorporate all concerns that are today handled manually during the time-tabling process into the programme.

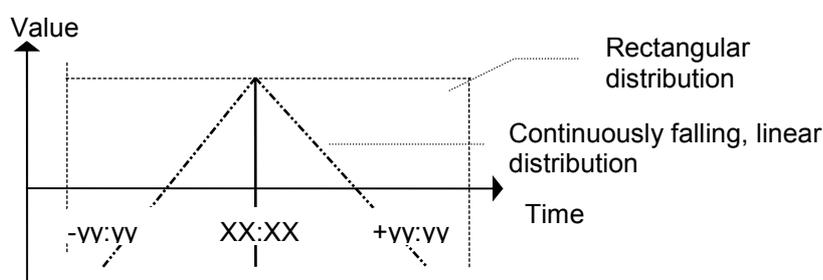
The particular feature of this test was that it included, for the first time, actual operators in the region. At an initial meeting, representatives of the firms were informed about the particulars of the model and the test. A week before Christmas, the demand for 223 departures, coming from eight different operators was submitted (cf. Table 3). The demand specification more precisely comprised the following material:

- Departure station
- Arrival station
- Type of train (specification of traction, no. and type of cars and/or weight).
- Time interval within which departure is of interest.
- A value function defined over this time period (cf. Figure 3 for examples).
- Maximum time acceptable to be sided after departure from origin.
- Compulsory stops.
- Minimum required turn-around time when defining two or more legs of a train.

**Table 3** Demand for and allocation of departures in desktop trial.

| Operator               | Demand for departures | No. of departures allocated after iteration 4 |
|------------------------|-----------------------|---|
| Green Cargo            | 110                   | 91  |
| TiB                    | 62                    | 24  |
| BK Tåg                 | 5                     | 1   |
| Inlandsgods            | 4                     | 4   |
| Uppsala-Borlänge       | 18                    | 4   |
| Uppsala-Sala-Västerås* | 12 (24)               | 4 (8)   |
| Tågkompaniet           | 2                     | 2   |
| TGoJ                   | 10                    | 8   |
| Total                  | 223                   | 138   |

\*The number in parenthesis represents all legs while the lower number is the demand that has specified for particular train-pairs.



**Figure 3** Examples of how value-of-access may vary with departure time.

The results from the first set of trials were sent back to operators after Christmas. This included a graphical timetable for the network, line by line, as well as a copy of the original demand specification where it was indicated whether the demand had been accommodated and, if so, what price the operator would have to pay for each departure. It was then feasible for them to adjust their demand specification and/or their bid. Four such iterations were made before the trial was ended.

Table 3 indicates that 85 out of 223 trains were not allocated a departure slot. The prime reason for that a large share of the demanded departures were not included was that the demand specifications were extremely narrow. Since several operators specified that their train(s) would have to leave within a time-slot of only a few minutes, there were small possibilities to adjust demand; these departures therefore had to be expelled. The traditional, manual way to handle this issue is typically to give the operator something else than is specified, but the present approach is based on the idea that such adjustments should not be done without the explicit consent of operators.

A complementary reason was that some operators did not allow for any slack in the time-table that was specified; the slack in the time-table that actually is operated is considerably larger than in the demand specifications. There were also omissions and pure misspecifications in the material that was submitted. The underlying reason for the relaxed position of several participants was, most probably, that the “bids” made were not for real which made the incentives to

participate weak. One of the major players also abstained from participating with more than the initial demand specification.

The trial has, never the less, generated important insights. One is that the optimisation technique that has been developed works well. This means that (changes of) demand specifications generate the expected consequences for prices and priorities. Another is that the user interface, and the logic of the process, is simple to understand. One of the larger operators reacted on each new time-table within hours after that a preliminary result had been sent out, and the adjustments made were highly rational from the perspective of this operator. There is all reason to believe that the mechanism could be applied to increasingly more complex problems.

### 4.3 Conclusions

Current track user charges are below their efficient level since the Swedish network is unable to cater for the actual or potential demand that there is. But why has the rate of progress to develop and introduce congestion pricing been so slow?

It has previously been emphasised that a prime (implicit) concern when the charging system was first drafted, and also when it was revised in the late 1990ties, seems to have been to ascertain that railway traffic is not discontinued: A requirement for the industry to recover the full costs of the activity could mean that substantial parts of the network would no more be used. Congestion charges would increase total payments and the perceived risk may be that this would risk pushing operators closer to bankruptcy.<sup>14</sup> The need to ascertain an efficient use of existing infrastructure may weight easy compared to the desire to maintain services.

A second explanation to the poor interest in the topic of efficient track allocation is, however, that the conflicts of interest have not surfaced until fairly recently. The incumbent was in reality alone on the tracks until the mid-1990ties, and even when the number of operators started to increase towards the end of the century most operations were still handled by SJ. This was also the reason to leave the responsibility for track allocation issues with SJ when the split was made in 1988.<sup>15</sup>

The subsequent deregulation process has made the necessity to develop more transparent allocation techniques gradually more obvious. It was also the need to handle the allocation process in an even-handed way that transferred the Rail Traffic Administration to *Banverket* in 1999. The time-table operated from June, 2001 included more than 20 different operators. In particular, the split of the incumbent into separate passenger and freight companies has sharpened the conflict of interest.

A committee appointed in 2001 in order to draft a proposal for implementing the recent EC directive on infrastructure pricing etc. into the Swedish judicial practice also addresses the scarcity issue. One of the concerns of the committee has been whether or not it could suggest a procedure based on the above auctioning principles but which would still be agreeable with the directive. It still

---

<sup>14</sup> The same hesitation may lay behind the reluctance to introduce charges for marginal reinvestment costs.

<sup>15</sup> The 1988 Bill actually suggested that capacity allocation should be handled by the infrastructure agency. A committee assigned to design the details of the new organisation however – for the above reason – made the opposite recommendation, and the government choose this design.

remains to be seen what the committee actually will suggest and indeed what the political response to these suggestions will be.

A third and final motive for the low interest comes from inside what now is the Rail Traffic Administration. These people are highly regarded for their professional skills and control the process tightly. Demand specifications may be rehashed radically before being sent back to operators with a “take-it-or-leave-it” attitude. Time-tables that are being suggested are, by definition, operational and it is difficult for outsiders to realise whether or not it would be feasible to draft a schedule that in one way or another would be “better” than the proposal. The pressure from inside this community to replace current paper-pencil-ruler techniques with opaque optimisation procedures that would make much of the expert knowledge redundant has, to put it mildly, been low.

## **5 Conclusions**

The willingness to step away from a cost recovery perspective, which often seems to be an obstacle against marginal cost pricing, is not of the same concern in Sweden as it seems to be in several other countries. Sweden’s railway sector is rather organised so that it is feasible to introduce charges that comply with an efficiency-enhancing structure. The country’s problem in this context seems to be that charges have been cut by more than would be efficient.

The present paper has, thus, pointed to that the price for infrastructure use should be raised in order to take all social marginal costs into account. There are, in particular, two new cost components that should be lifted into the pricing scheme. One concerns the consequences of variations in traffic levels for the spending on reinvestment, i.e. on the date for replacing existing tracks. It is reasonable to believe that this expenditure is dependent of the traffic volume, but more work should be done in order to explore the optimal size of this charge.

Secondly, capacity shortages have to be priced. There are a number of reasons for doing so. At the core of the argument is that a scheme for scarcity pricing is the best way to ascertain that the existing infrastructure is efficiently used. In addition, the prices would signal at which spots investment spending is most warranted for capacity reasons; the higher the charge, the more difficult is it to admit all trains that would want access to the tracks.

We have also reported about an auctioning mechanism that could be used to implement scarcity pricing in practice. The next natural step in the development of such mechanisms would be to try it out in parallel with the traditional way to allocate capacity, albeit without actually charging operators in this trial.

An interesting feature of an auctioning mechanism is that – although it would increase operators’ costs for running trains – the burden of the new cost could not be passed over to passengers or freight customers. The reason is that the auctioned price would be a fixed cost for the operators, not part of their cost for running the system. The congestion price would therefore have to be paid out of operator profits.

## References

- Alexandersson, G, Hultén, S, Nordenlöw, L & Ehrling, G (2000): **Spåren efter avregleringen** (*Trailing the Deregulation*, in Swedish only). KFB-report 2000:25.
- Brewer, P J & Plott, C R (1996): **A Binary Conflict Ascending Price (BICAP) Mechanism for the Decentralized Allocation of the Right to Use Railroad Tracks**. International Journal of Industrial Organization, Vol. 14, No. 6, pp 857–886.
- Brännlund, U, Lindberg, P O, Nilsson, J-E & Nöu, A (1998): **Railway Timetabling Using Lagrangian relaxation**. Transportation Science, Vol. 32, No. 4, November.
- Hansson, L & Nilsson, J-E (1991): **A new Swedish Railroad Policy: Separation of Infrastructure and Traffic Production**, Transportation Research, January.
- Isacsson, G & Nilsson, J-E (2002): **Re-regulation of Previous State Monopolies. An Experimental Comparison of Allocation Mechanisms in the Railway Industry**. Working Paper (2000).
- Johansson, A-S & Helge, L (2001): **Avregleringen av den långväga buss-trafiken – en utvärdering**. C-uppsats i Nationalekonomi, Högskolan Dalarna.
- Johansson, P & Nilsson, J-E (1998): **Marginalkostnader för banslitage**. CTS Report 1998:7.
- Johansson, P & Nilsson, J-E (2002): **An Economic Analysis of Track Maintenance Costs**. Working Paper.
- Laffont, J-J & Tirole, J (1993): **A Theory of Incentives in Procurement and Regulation**. The MIT Press.
- Lindberg, G (2002): **Marginal Costs of Road Maintenance for Heavy Goods Vehicles on Swedish roads**. Working Paper for the UNITE project.
- Nilsson, J-E (1992): **Second Best Problems in Railroad Infrastructure Pricing and Investment**. Journal of Transport Economics and Policy, September.
- Nilsson, J-E (1995): **Swedish Railways Case Study**. Chapter 8 in Kopicki, R. & Thompson, L: Best Methods of Railway Restructuring and Privatisation. CFS Discussion Paper Series, No. 111, August.
- Thompson, L (1999): **Experimental Evidence on the Use of Priority Auctioning in the Railway Industry**. International Journal of Industrial Organisation no. 17, pp 1139–1162.
- Thompson, L (2002): **Towards a Welfare Enhancing Process to Manage Railway Infrastructure Access**. Transportation Research, Part A.
- Thompson, L (2002a): **Restructuring Sweden's Railways; the Deregulation that Wasn't**. Working Paper.
- Proposition 1987/88: **Trafikpolitiken inför 90-talet**.
- SFS, 1998:1827: **Förordning om avgifter för trafik på statens spåranläggningar**.
- SIKA (2002): **Nya banavgifter? Analys och förslag, SIKA/Banverket**. *New rail infrastructure charges?* SIKA-rapport 2002:2 available at [www.sika-institute.se](http://www.sika-institute.se) (Summary in English.).
- SJ (1998): Annual report.

**Table 1** *Banverkets revenue from track user charges and spending on investment and maintenance, million SEK. Source; Banverket data.*

|                           | 1989   | 1990   | 1991   | 1992   | 1993   | 1994    | 1995    | 1996   | 1997   | 1998   | 1999   | 2000   | 2001   |
|---------------------------|--------|--------|--------|--------|--------|---------|---------|--------|--------|--------|--------|--------|--------|
| Revenue                   |        |        |        |        |        |         |         |        |        |        |        |        |        |
| Fixed                     | 392    | 377    | 337    | 299    | 291    | 288     | 274     | 283    | 273    | 272    | 0      | 0      | 0      |
| Variable                  | 359    | 364    | 350    | 395    | 379    | 396     | 437     | 542    | 761    | 757    | 275    | 442    | 456    |
| Total revenue             | 751    | 740    | 687    | 693    | 670    | 684     | 711     | 824    | 1 034  | 1 029  | 275    | 442    | 456    |
| Costs                     |        |        |        |        |        |         |         |        |        |        |        |        |        |
| Maintenance               | 1 870  | 1 878  | 1 854  | 2 218  | 2 236  | 2 297   | 2 307   | 1 982  | 2 010  | 2 022  | 2 384  | 2 434  | 2 745  |
| Reinvestment              | 805    | 835    | 702    | 615    | 821    | 1709    | 1 822   | 1 354  | 1 043  | 859    | 520    | 827    | 858    |
| Investm., main lines      | 1 016  | 1 683  | 2 398  | 3 970  | 4 408  | 6 711   | 7 943   | 7 024  | 4 654  | 5 450  | 4 803  | 3 541  | 2 925  |
| Investm., secondary lines | 124    | 184    | 176    | 106    | 279    | 853     | 700     | 339    | 229    | 425    | 265    | 239    | 352    |
| Total costs               | 3 815  | 4 580  | 5 130  | 6 909  | 7 744  | 11 570  | 12 772  | 10 699 | 7 936  | 8 756  | 7 972  | 7 041  | 6 880  |
| Revenue - Cost            | -3 064 | -3 840 | -4 443 | -6 216 | -7 074 | -10 886 | -12 061 | -9 875 | -6 902 | -7 727 | -7 697 | -6 599 | -6 424 |

Figure 1: Relative change of traffic activities 1970-1999. Source: SIKÅ (1999)

