Can The “Invisible Hand” Draw The Railroad Timetable?

Patricia Perennes, Centre d’économie de la Sorbonne (CES), Université Paris 1, Réseau Férré de France

patricia.perennes.cnrs@exterieur.rff.fr

1- Introduction

The railroad industry is currently being liberalized in Europe under the influence of the European Commission. The goal of this liberalization is to reduce costs and prices in the industry, thereby making it more competitive in intermodal competition: freight transportation by train can re-earn market shares over trucks, passenger fast train transportation over airplanes and commuter trains over personal cars. This improvement derives from basic results of microeconomic theory: (i) if a market is characterized by a monopoly, prices are higher and less quantity is produced, (ii) a firm that does not face competition has no incentive to be efficient. Since only infrastructure (railroad network) is a natural monopoly, the industry has to be split between infrastructure and operations (train services) in order to introduce competition on operations.

Once the industry has been split, a question emerges: how can the infrastructure manager (IM) efficiently allocate capacity between different train operating companies (TOCs)? This question has been subject to a vigorous debate between economists at the beginning of the Swedish railroad liberalization (see Brewer and Plott (1996)). At the heart of this debate was the possibility for the IM to use market mechanisms to allocate train paths.¹ Can market mechanisms deal with the complexity induced by technical specificities of railway capacity or should this capacity be allocated by a central planner? This debate is not settled today. Most of the current economic literature focuses on capacity allocation through auctions but some articles (see Quinet (2003) or Gibson (2003)) still underline that it is difficult, if not impossible, to use market mechanisms to allocate rail capacity. In addition, the industry is still

¹ In this article I will indistinctly use the word “slot” and “path” to designate the same object: the right for a TOCs to use a portion of the network (between two stations) at a specific moment of time. For example, it is the right to run a train between Paris and Lyon using the high speed line between 6.30pm and 8.30pm. I am aware that this definition is debatable. One may consider that my example does not constitute a single slot but is vector of slots: one slot each time the train passes a station (in this case there would be two slots Paris – Le Creusot and Le Creusot-Lyon). In the economic literature both terms are used quite indifferently.
reluctant to use market based mechanisms to allocate capacity. The best proof thereof being the fact that auctions processes have not been used to date to allocate capacity.

As far as I know, this debate between auctions opponents and proponents never mentioned Ronald Coase’s landmark paper “The Nature of the firm”, whereas Coase’s article boils down to a confrontation between market mechanisms vs. decisions taken by an integrated firm. This article fills this gap. Based on Ronald Coase’s approach, it clarifies the reason why auctions processes are not used today: it is maybe not the result of the railroad industry archaism or of the lobbying of TOCs but simply of the costs and uncertainties that come along an auction process. It then suggests an intermediate way to attribute capacity, taking into account the contribution and the shortcomings of the auction-oriented literature. This intermediate solution is based on the distinction between slot allocation and timetable drawing processes.

It is organized as follows. First, it explains more precisely how capacity is allocated in the railroad industry (2). Then, it reviews the existing literature (3) and compares this theoretical literature with a real world example (4). As shown by a brief description of the costs involved by market mechanisms utilization, the long and centralized timetabling process carried out by IM may be justified by “price discovering” costs (5). To conclude, it introduces an intermediate way to allocate capacity and highlights the fact that a timetabling process reducing transaction costs may be inconsistent with the liberalization process (6).

2- Capacity allocation in the railroad industry

The following section explains how capacity is divided between different trains and/or TOCs. It first describes how capacity is concretely allocated, then underlines why the railroad industry faces specific issues compared to other industries.

a) Attribution of capacity in the railway industry

The IM uses timetable to allocate capacities between activities (freight/passengers) and companies. Concretely, it means that the IM looks at its infrastructure, see how much capacity is available, calculates how many trains can circulate on its network and literally draw a
graph\(^2\) that shows which train is going to be on which track. The best way to illustrate this definition is to give an example:

**Fig 1: Imaginary example of a timetable**

![Diagram of train timetable](image)

The abscissa axe is the timeline, whereas the ordinate axe represents the main stop-off points. The red and blue lines are trains. I use the French train terminology: in this example, there are high speed trains (TGV) and regional trains (TER).

This representation allows the IM to make sure that two trains are not at the same time on the same tracks in which case they would collide with each other.\(^3\) This may sound self-evident, but one has to keep in mind that this so called “self-evidence” is not true for other industries, such as telecommunication where different phone calls are sent on same phone lines without any trouble. The extreme rigidity of capacity repartition\(^4\) in railroad industry makes timetables

\(^2\) More precisely, one graph per railway line.

\(^3\) In my example, TGV2 passes TER2 in a station, so there is no collision as long as there are enough platforms.

\(^4\) This “rigidity” manifests itself on various levels: (i) a train cannot pass another one except when the line is doubled (for example in major stations), (ii) for several reasons, such as passenger expectations and difference in trains’ speeds, it is impossible to organize capacity allocation on a spontaneous basis (i. e. TOCs do not book capacity in advance but just use capacity on a day-to-day basis as long as some capacity is available, like trucks on roads). Once again, these facts may sound self-evident, but they are not verified for other industries: there is no « passing difficulty » in the airline industry.
an indispensable tool: they allow the IM to optimize crossing and trains repartition on the network.

A distinction can be drawn between timetabling and slot allocation. Slot allocation is the decision to attribute capacity for a specific train. To go back to my graphic example, it means allocating one of the “lines” to a company. To sum up, timetable drawing means constructing the whole “tree” when slots allocation is giving the “branches” one by one.

In real life, these two processes are of course linked. Timetables are usually drawn in the following manner: the timetable of the previous year is used as a basis for discussion. The companies already entitled to capacity announce if they still need this capacity. Newcomers also ask for capacities. If the demands of the newcomers can be added in the graph, new slots are drawn. In the case of conflicting demands, the IM can try to accommodate them, moving backward/forward in time different slots so they can all fit in the timetable. It can also use priority rules to decide which company is entitled to capacity.

I will describe more precisely the overall process in France in the fourth part of this article, but this brief description enables to understand the issue: it is hard to draw a clear line between what belongs to timetabling and what belongs to the slot allocation process. Nevertheless, this distinction may have a theoretical relevance, as I will show in the fifth part of this article.

\[ b) \text{ Difficulties encountered in railroad capacity allocation} \]

The timetable drawing process described above is specific to the railroad industry. Gas transportation requires booking injection and withdrawal capacities, but the exact location of the gas after injection is not important. Airports schedule slots only encompass taking off and landings, but not path planning. As noted by Borndörfer et al. (2005).

“\[ \text{The analogy between railway and airport allocation is limited. Indeed, for air planes, only take-off and landing slots have to be planned in advance, while routing in air takes place spontaneously. In railways, in contrast, issues like track capacities, overtaking, and signaling systems are essential. Therefore, the efficient use of a railway network hinges not only on departures and arrivals, but also on the combination of routes and speeds taken by individual trains in the network. (...) Constraints such as these make railway capacity allocation much more complex than that of airport slots.} \]"
It is also worth noticing that liberalization makes the timetable drawing process and the slot allocation process more complex. As Brewer and Plott (1996) write “Historically, scheduling has been seen primarily as a technical problem and not as an economic/political problem”. Previously, these processes were totally internalized by the incumbent company that was at the same time the IM and the unique candidate for capacity. The decision to allocate capacity to different activities was an internal managerial choice. Once network and operations are split the IM faces a fundamental information problem. The TOCs have no incitation to reveal the true valuation of the slots. So the IM has to find incentive mechanisms (or a reliable source of information regarding the costs and benefits of TOCs) to price slots and allocate capacities. Therefore, auctions may seem a natural way to set the relevant price and attribute capacities at the same time, all the more so as auctions processes were and are used in other network industries to allocate capacities. For example, in 1993 the US Congress authorized the use of auctions as a spectrum assignment mechanism. Why not replicate the auction pattern used in other industries for railroad?

Rail capacity has key features that hinder the application of simple auctioning methods:
- As explained above, capacity planning is much more complex in railroad than in other industry, such as airports or energy, because of its exclusionary property.
- Slots are not homogenous goods: a train path drawn for high speed trains cannot be used by slower trains providing regional service (or vice versa).
- The value of one train path is contingent to the overall pattern of service. The value of a slot can increase or decrease depending on which company runs a service on which part of the tracks. It decreases when many direct competitors provide the same services (have been entitled to similar slots) and increases if slots are allocated in a way that allows connections.

Therefore, a specific approach has to be developed to tackle these problems. I will now review the economic literature that devotes itself to this question.

3- Literature review
The early economic research devoted to rail capacity allocation appeared in the mid-90 and was inspired by the reform taking place at that time in Sweden (Brewer and Plott (1996), Nilsson (1999)).

In 1996 the article of Brewer and Plott set the context of the discussion that has dominated the economic literature on railway capacity since then: can auction be used as a tool to allocate
capacity in the railroad industry? Brewer and Plott’s article was written in response to a consulting report\(^5\) that stated that it is impossible to allocate slots by decentralized mechanisms because of the very technical nature of the industry. The arguments developed by this report leaned on the features of railroad capacity mentioned above.

“These train paths cannot be treated as independent units, since they are not interchangeable, and depend on the specification of all other paths in the integrated timetable. There is therefore no common unit of capacity on a mixed-use railway which can be allocated to owners, prices and traded among a number of buyers and sellers.

“However, a simple free auction cannot be used for railway capacity since there are no independent units of capacity to bid for. The viability of every bid to operate a train service depends on the specification of every other train service which has been bid for.”

To counter these arguments, Brewer and Plott used a testbed experimental environment to prove that it is possible to create an auction process that successfully allocates rail capacity. Nevertheless, their experiment concern only a limited number of slots and was based on a simplifying assumption\(^6\). As the authors conclude “the issues raised by the deregulation effort in Sweden do not end with this paper. More complex environments must be pursued and the BICAP might become modified in order to accommodate the problem they present”.

Following this landmark article, numerous articles were written improving each time the auction modeling.

They either introduce new hypothesis to come closer to “real world” situation or test new mathematical/computer tools to support resolution. Nilsson (2002) introduces the idea that slots can have a complementary value and suggests bidding for bundles of trains, Parkes and Ungar (2001) introduce double-track segment that allow train to pass each other, Brännmund et al. use an integer program (Lagrangian relaxation techniques), etc. An excellent review of this literature can be found in Borndörfer et al. (2005).

\(^5\) This report was prepared by the consulting firm Coopers and Lybrand and was published as an appendix to a larger report for the Committee “Increased Competition Within the Railway Sector – Review of Proposals” January 1993. Unfortunately, this report and its appendixes are not available today. I was not able to review them.

\(^6\) Borndörfer et al. (2005) explains that in Brewer and Plott model “slots have the binary exclusion property, that is, a set of slots is consistent if any two of its elements are. This condition, however, does not hold for a general network of tracks with alternative connections available between origin-destination pairs.”
Nevertheless, this literature was not able to give a realistic process that can be used under real-world conditions, i.e. for large scale network (and not separated lines), taking into accounts all the features of rail capacity. To quote Borndörfer et al. (2005)

“Overall, the question of whether auctions can be used for efficient rail path allocation, and how an implementation would need to work, can’t be answered from the existing literature. As far as we know, the literature does not propose models and solution concepts for track allocation problems of practically relevant sizes that would adequately reflect the complexity of a real-world railway network. In fact, most of the literature considers simplified, non-branching lines, or even single segments.”

Borndörfer and al. (2005) article is no exception to this rule. It uses linear programming to solve the bid maximisation, but as it acknowledges “it remains to be shown that our approach can be scaled to larger scenarios”. Articles written since then face same kinds of problem: they try different mathematical programming technique or optimization through iterative improvement of a candidate solution (Ho et al. (2012)), but they either face problems of implementability (Borndörfer et al.(2005)) or may lead to non-feasible solutions (Ho et al.2012).

Parallel to this coherent body of literature, some articles still dispute the fact that auctions can ever be a manageable way to allocate rail capacity. Quinet (2003) indicates that “a simple price per path may not be able to select the optimal solution (…). A possible solution to the path allocation problem [is] to solve it through central planning process”. Similarly Gibson (2003) writes that “given the key features of rail capacity, it would require a hugely complex mechanism to introduce an auction-based approach to allocating capacity, and there is very little appetite within the industry to do so.” The last remark of Gibson also points out an interesting fact: contrary to other industries (telecommunication, airports, energy) auctions based mechanisms were never implemented in the railroad industry partly because all the stakeholders (IM, TOCs, regulators) are reluctant to use such allocation processes. Moreover, market mechanisms are totally absent from the capacity allocation process in certain countries, as showed by the fourth part of this paper.
4- Example of a concrete allocation process: the “conciliation mechanism” in France

As underlined in the previous part, the economic literature has as early as 1997 ruled out the argument according to which decentralized market mechanisms are not able to deal with the complexity of rail capacity allocation. Nevertheless, the industry does not seem to have taken this result into account. The French rail capacity allocation process is a convincing example of this situation.

In France, the liberalization reform initiated by the European Commission has led to patrimonial separation of the infrastructure manager (IM) (Réseau Ferré de France or RFF) from the Société Nationale Des Chemins de Fer (SNCF). Competition was introduced for international freight transportation in 2003, national freight transportation in 2006 and international passengers at the end of 2009. National passenger transportation and commuter/regional trains are not to date open to competition.

The allocation process is totally coordinated by RFF (even if it subcontracts some tasks to the SNCF). Prices are set by RFF and reviewed by the railway regulator (Autorité des Activités Ferroviaires, ARAF). They are based on marginal cost of network usage. Profitable activities (high speed trains) also pay part of their total cost. Prices take congestion into account. Hence, TOCs are price takers.

Capacity repartition is a long process. RFF starts to collect information concerning the potential demand of TOCs five years to two years before the actual railway operation. RFF realizes prospective studies concerning potential traffic evolution. TOCs that want to launch a new service have to ask RFF to realize “feasibility studies” to make sure capacity is available. These studies have to be launched at least two year before the beginning of operations. RFF also organizes meetings with TOCS to discuss their future needs. It also discusses with the operator in charge of the network maintenance (SNCF infra) about its capacity needs. Based on these exchanges, RFF constructs a first timetable draft also called the “paths catalogue”. It submits it to TOCs. Each of them indicates the slots for which they will likely to ask.

---

7 Currently, only one competitor (Thello) of the SNCF offers train transportation between France and Italy.
8 The civil engineering branch of the former incumbent still has a legal monopoly on network maintenance.
The formal stage for capacity requests starts one year before railroad operations. TOCs submit requests for paths to RFF based on the slots catalogue. They are also allowed to ask for slots outside of the catalogue.

Then, RFF passes all the requests on to an independent branch of the SNCF (the DCF) which is RFF’s subcontractor for drawing the timetable. The DCF draw the graphs, starting by delimitating daily work time bands and drawing first the more complex slots (trains crossing the whole countries and international trains). In case of conflicting demands, it contacts the TOCs to see if their slots may be change a little to fit into the graph.

The French legislator created a process of “declaration of overloading” in case some conflicts are not solved by these conciliation mechanisms. If parts of the network are overloaded, RFF can move apart demands following priority rules. Nevertheless, RFF has to date never declared any lines overloaded and has always find a way to fit all the demands in the timetable.\(^9\)

Once the DCF has finished the timetable it gives it to RFF which releases it. After this, TOCs can still ask for capacity, but slots are only drawn in the remaining capacity, if any.

The French case is a good illustration of the capacity allocation processes across Europe. The goal of the IM is to conciliate as much as possible the conflicting demands and prices are generally not set by market mechanisms.\(^10\) Another interesting point is the fact that the different kinds of trains (high-speed, regional, freight) are not set on the same footing: capacities are delimited for each activities and the relative prices differ.

5- Lessons from “The Nature of the Firm”

How can we analyze this gap between the economic literature which elaborates highly complex auctions mechanisms and the actual allocation process in France or in other European countries which generally excludes any market mechanisms?

One may think that the reluctance of the industry can be justified by the lobbying of TOCs and the archaism of the railroad industry. As Nilsson (2002) explains

“Time-tableing staff are used to their rulers and pencils and may be sceptical towards automation of their precious skills; operators understand that they

\(^9\) The ARAF criticizes RFF behavior: in its Opinion n°2012-005 on the 2013 national rail network document, it notices that RFF should use overloading declaration whenever it is relevant.

\(^10\) A notable exception is the German regulation that allows the application of a “highest price procedure” in case of conflict between equally ranked slot requests. The Slovenian IM has the same kind of rule. These two countries seem to be exceptions among Europe.
would have to face higher costs because of an auction, and the same goes for freight customers and travellers; and since infrastructure around Europe is subsidised by tax money, the IM's incentives to forward an auction that compared to status quo – would transfer resources from railway operators to taxpayers are weak.”

Nevertheless, the truth might lie somewhere else. There are two competing visions of timetable drawing. One relies exclusively on market mechanisms. According to the other the timetable has to be drawn by an integrated firm. It is interesting that this debate never mentioned Ronald Coase’s landmark paper “The Nature of the firm”, whereas Coase’s article provides interesting food for thought for the timetable drawing process

“Within a firm, these market transactions are eliminated and in place of the complicated market structure with exchange transactions is substituted the entrepreneur co-ordinator, who directs production. It is clear that these are alternative methods of coordinating production”

He then explains the comparative advantages of one form of organization or the other. Using market mechanisms is the most efficient way to allocate productions factors. However, “there is a cost of using the price mechanism. The most obvious cost of “organising” production through the price mechanism is that of discovering what the relevant prices are”. This particularly fits timetable construction since the market mechanisms necessary to reveal prices and allocate capacity are extremely complex as the literature review demonstrated.

Nevertheless, it is hard to tell if the costs of using the price mechanism overcome the benefit of an efficient capacity allocation. Indeed, in the case of “The Nature of the Firm” competition among firms and market mechanisms lead “naturally” to one form of production coordination: a firm emerges when using price mechanisms is more costly than organizing production centrally. On the contrary, the choice of constructing the timetable through a centralized vs. a decentralized process is not done “naturally” through competition but results from a political choice. The government of a country, or the regulator if this choice has been delegated to it, decides if market mechanisms should be used to allocate capacity. Once this choice is done, its relevance is only tested by inter modal competition (trucks, individuals cars, airplanes, etc.). This competition is besides smoothed by subsidies, difference in the consumers’ preferences, etc.
In order to enlighten this choice, it is interesting to try to summarize the reason why costs of using price mechanisms to build the timetable may be high, and how to reduce them.

First, the fact that auctions are used in other industries is not an indubitable proof that they can be used in the railway industry. As explained in the second part of this article, rail capacity has distinctive features which make its efficient allocation by market mechanisms more complicated than capacity in other industry.

If a capacity allocation process based on market mechanism is implemented in the railroad industry, the costs induced by “price revelation” will not be borne primarily by the IM but by TOCs. One key feature of rail capacity is the price interdependency of different slots. To counter this, economists have introduced the idea that the TOCs should not bid for one single slot, but for vectors of slots (see for example Caillaud (2002)), or give a profit/valuation function (see for example Brännlund et al. (1998)). Therefore, TOCs internalized the links existing between different paths.

These solutions may seem attractive, but ask a real question about the ability of TOCs to formulate all the necessary price sets/functions in a reasonable amount of time and for reasonable costs. To illustrate this point clearly, let suppose the auctions process takes the form of bidding for sets. As Caillaud (2002) explains, if there are \( n \) slots that means each TOCs has to bid for \( 2^n - 1 \) configurations. Knowing that each year in France the numbers of slots requested approximate 36 000\(^{11}\) the number of configuration is astronomical.\(^{12}\) Defining proper valuation functions taking into account all the contingencies is even more complicated for a private firm.

So maybe the IM can deal with the number of data using a proper computer system and appropriate software, but the TOCs will not be able to formulate their offers. In addition, using such complex mechanisms for selling capacity raises the issue of the ability of TOCs to understand correctly the auctions process and to submit rational bids.\(^{13}\)

To reduce these costs, a simple solution may seem to be, following the ideas developed by Ronald Coase in “The Problem of Social cost” to allocate randomly the slots and to let the TOCs negotiate to reallocate them. Nevertheless, one the prerequisite of the so-called “Coase

\(^{11}\) This figure of 36000 slots is in facts minoring the number of slots that could be counted in an economic perspective. There is only one slot when a train is request for every day of one year. The slots also counts only once when a train crosses large distances, stopping in numerous stations (see footnote 1).

\(^{12}\) And this is not only a play on words: knowing that the number of atoms in the universe is estimate to be \( 10^{80} \), that means that any combination of more than 265 slots would require TOCs to submit more auctions that there are atoms in the universe.

\(^{13}\) The learning by doing possibility is limited for rail capacity allocation, since capacity is sold only once a year. If the bids of a TOCs are over/under estimated it is likely to go bankrupt before being able to submit new bids.
Theorem” is a good definition of property rights to internalize the externalities incurred by reallocation. This does not hold for railroad capacity. If three TOCs A, B, C received slots and A and B agree to trade slots, they also need to negociate with C even if they do not directly need its slots. A reallocation of the slots they hold is going to change the value of the slots owned by C. This is especially true if A and B do not have the same activity, i.e. if one is a passengers trains company and the other a freight company. If they do not include C in their negotiations, this would be an infringement of C property rights and it can bring suit. One can easily understand that if negotiations may be manageable for three of four companies, it becomes harder if there is more than a handful companies. In addition, organizing negotiations involving all the TOCs is a problem on an anti-trust point of view.

So, are auctions absolutely impossible to use to sell railroad capacity? The arguments developed above points out the fact that the costs of using price mechanisms are high for railroad capacity because of the correlations between different slots’ prices and the uncertainties borne by the TOCs in case of latter reallocation of the slots. To make auctions possible, and to reduce the costs of price revelation, a feasible solution may be that the IM centrally draw the graph (or at least a sketch of the graph) but may sell individual slots using market mechanisms. In other words, to use the distinction made in the second part of my article and to reply to the title question, the invisible hand is not able to draw the timetable but it can sell the individual slots once the overall timetable is set.

The timetable construction process can therefore never be totally decentralized. The IM will have a role to play to collect information and to draw the timetable, before allocating paths to different activities and TOCs.

In this context, the question that the economic literature should try to answer is not to find a methodology to draw the entire timetable by a single auction but how much uncertainties and prices contingencies TOCs can deal with at a reasonable cost. Improvement of auctions techniques may of course enlarge the room for market mechanisms in the timetabling process, but as Borndörfer writes “The optimization of the entire network timetable in one single auction is not a practicable aim anyway”.

This approach will also solve another problem which was not raised in this article. Trains operating under a franchise contract (in particular regional trains) cannot bid in the same auction process that trains operating under open access competition. Indeed, when a TOC
operates a franchise its revenue is set at the beginning of the operation period defined by the contracts (in the railroad industry at least 5 years). If the price of the slots it used dramatically evolved during this period, it can face unbearable difficulty. Therefore, in some countries such as France, the transport authority exactly compensates the TOC for any increase in the slots prices. In these circumstances, franchise TOCs are theoretically able to pay an infinite price (see Dehornoy (2009)).

In the same way, the ability to pay of the different kinds of services (freight, high speed passengers, “normal” speed long distance passengers, regional transport) differs depending on the intermodal competition they face (trucks, airplanes, personal cars). Nevertheless, they share the same capacity in the timetable.

To put all activities on an equal footing, an obvious solution is to allocate dedicated capacity to each of them. This boils down to drawing first the timetable, then allocating slots to each activities, similarly to RFF “paths catalogue” approach.

6- Conclusion

By using Ronald Coase landmark article “the nature of the firm”, my article has filled the gap between economic literature regarding capacity allocation and allocation in real world conditions. There may be room for market mechanisms in the slot allocation process but the timetable drawing process (or at least some of this process) should be endorsed by a single entity. Therefore, some of the steps of the French IM rail capacity allocation process (construction of a slots catalogue, exchanges with the industry about its future needs) have an economic rationality and should perpetuate in the future.

This conclusion has strong consequences regarding the liberalization process. The fact that the timetable may be less costly designed through centralized planning than through market mechanisms questions the whole liberalization process. By organizing information transparency several years before actual slots utilization, it may constitute an entry barrier. A competitor that wishes to enter the market should be able to evaluate its potential capacity needs five to two years in advance. In addition, these discussions about capacity may provide the incumbent with information concerning the strategy of its future competitors. This may put back into question the possibility of a head to head competition “on” the network. The only viable competition may be “for” the network, once the centralized planner has divided the capacity utilization between activities and fringe competition.
Références


