
Maite Peña-Alcaraz  
(corresponding author)  
Ph.D. Candidate  
Engineering Systems Division  
Massachusetts Institute of Technology  
Cambridge, MA  
Email: maitepa@mit.edu

Mort Webster  
Associate Professor of Energy and Mineral Engineering  
Energy and Mineral Engineering  
Pennsylvania State University  
University Park, PA  
Email: mdw18@psu.edu

Joseph M. Sussman  
JR East Professor of Civil and Environmental Engineering and Engineering Systems  
Engineering Systems Division  
Massachusetts Institute of Technology  
Cambridge, MA  
Email: sussman@mit.edu
Recently, governments worldwide have started promoting the use of shared railway systems as a way to take advantage of the existing capital-intensive railway infrastructure. Until 1988, all major railways both managed the infrastructure and operated the trains. In contrast, in shared railway systems, multiple train operators utilize the same infrastructure. Such systems can achieve high utilization, but also require coordination between the infrastructure manager and the train operators. Such coordination, in turn, requires capacity planning and regulation that determines which trains can access the infrastructure at each time, capacity allocation, and the access price they need to pay, capacity pricing. The need to establish capacity pricing and allocation mechanisms in the railway system is relatively new and the comparative performance of alternative mechanisms to price and allocate capacity is still a matter of study. This paper proposes a framework to analyze the performance of shared railway systems under alternative capacity pricing and allocation mechanisms. The paper focuses on how the introduction of price-based and capacity-based mechanisms affect the train operators’ ability to access the infrastructure capacity in the context of the Northeast Corridor in the US. The results of this paper suggest that there are trade-offs associated with each mechanism and none of them is superior to the other on all dimensions. As a result, Northeast Corridor stakeholders should carefully analyze the implications of alternative pricing and allocation mechanisms before locking the system into one of them.

Key words— Capacity pricing and allocation mechanisms, shared railway systems, railway system performance

1. INTRODUCTION

Recently, governments worldwide have started promoting the use of shared railway systems as a way to take advantage of the existing capital-intensive railway infrastructure. Until 1988, all major railways both managed the infrastructure and operated the trains, i.e., they were vertically integrated (Drew, 2006). In contrast, shared railway systems propose some level of vertical unbundling to allow multiple train operators (TOs) to utilize the same infrastructure. Examples of shared railway systems are the Northeast Corridor (NEC) in the US and the railway system in most European countries. Several countries in Asia and Africa are also opening access to their railway systems. Shared railway systems can achieve high infrastructure utilization, but also require coordination between the infrastructure manager (IM) and the TOs (Gomez-Ibanez,
especially when the IM has still some ties with a TO. Such coordination, in turn, requires capacity planning mechanisms that determine which trains can access the infrastructure at each time (capacity allocation) and the access price they have to pay (capacity pricing).

This paper focuses on the main spine of the NEC that stretches from Boston, MA to Washington, DC. This segment has four infrastructure owners and nine passenger TOs. Several freight TOs also operate trains in the system. Until now, capacity pricing and allocation in the corridor is managed via bi-lateral contracts negotiated between the IMs and the TOs. The price that each TO pays to access the infrastructure and schedule their trains depends mostly on how much capacity was available when the contract was signed (Gadner, 2013). This imposes two challenges in today’s operations: 1) the IMs recover a very small percentage of the revenues they would need to simply maintain the infrastructure, and 2) the introduction of new services is extremely complicated. Even if some train schedules could be shifted to make room to schedule new trains, rescheduling those trains would require the renegotiation of the contracts. As a result, the Federal Railroad Administration required Amtrak and the rest of the commuters and freight railway companies to agree on a new capacity pricing and allocation mechanism (PRIIA, 2008).

There are three main types of mechanisms to price and allocate capacity: negotiation-based, administrative-based, and market-based mechanisms. The use of market-based mechanisms for capacity pricing and allocation is preferred in systems like the NEC characterized by capacity scarcity (congestion) and conflicting demand (Perennes, 2014; PRIIA, 2008). According to (Gibson, 2003), the two main types of market-based mechanisms for capacity pricing and allocation for shared railway systems are 1) price-based and 2) capacity based. Price-based mechanisms are those that determine the price at which capacity will be offered, and let TOs decide whether they are willing to access the infrastructure or not. An example of a price-based mechanism would be a cost-allocation mechanism that allocates the cost of infrastructure among the trains that use it (Crozet, 2004; Nash, 2005; Lopez-Pita, 2014; Texeira and Prodan, 2014). Cost-allocation mechanisms are typically complemented with priority rules that allow the IM to decide which train to schedule when there are conflicts (multiple TOs willing to pay the predetermined access charges). Capacity-based mechanisms are those that determine the amount of capacity that will be offered, and let the TOs reveal the price that they are willing to pay to use that capacity, e.g. an auction (Affuso, 2003; McDaniel, 2003; Newbury, 2003; Perennes, 2014; Stern and Turvey, 2003).

So far, different countries have adopted different mechanisms to price and allocate railway capacity. (Nash, 2005; Lopez-Pita, 2014; Texeira and Prodan, 2014) provide an overview of the differences in pricing used in each European country, concluding that charging mechanisms are getting more heterogeneous. Furthermore, alternative mechanisms have been designed and implemented with different objectives and their comparative performance is still unclear (Drew and Nash, 2011; Nash, 2010). According to Nash (2003), “it is important to recognize that the concept of multiple operators may be relatively new for railroads: This means that the institutional framework has not been developed, and the intellectual understanding may not be in place, to facilitate planning and operating the shared-use system.” The authors warn against moving ahead quickly with the design of pricing and allocation mechanisms before understanding the implications of such mechanisms for all stakeholders.

In addition, the design, assessment, and implementation of capacity pricing and allocation mechanisms at the planning level are tightly coupled with the design of capacity operations at the tactical level (Krueger et al., 1999; Pouryousef and Lautala, 2015). In other words, the analysis of the performance of a pricing and allocation mechanisms cannot be determined in the absence
of operations because available railway capacity depends on how the infrastructure is operated. The understanding of the implications of pricing and allocation mechanisms thus requires the design of the train timetable to determine the arrival and departure time at every station of all trains scheduled.

The objective of this paper is to identify some of the trade-offs involved in the choice among alternative capacity pricing and allocation mechanisms for shared railway systems in the context of the NEC. This paper focuses on how the introduction of alternative pricing and allocation mechanisms impacts the ability of intercity and commuter TOs to compete for the access to infrastructure capacity. With over 2,000 commuter trains and 150 intercity trains scheduled in the NEC per day (Gadner, 2013), the ability of commuter and intercity TOs to access the infrastructure has a direct impact on the NEC passengers. We are particularly interested in evaluating these issues under two alternative capacity pricing and allocation mechanisms: a cost-allocation and priority-rule mechanism proposed by Amtrak (Gadner, 2013) versus a capacity-based mechanism. Auction mechanisms are widely proposed in the railway economic literature (Affuso, 2003; Perennes, 2014) but are not yet implemented in any country.

To analyze alternative capacity pricing and allocation mechanisms, we develop a framework to evaluate the performance of shared railway systems considering both planning and operational aspects. This framework consists of two models: a train operator model (TO Model) and an infrastructure manager model (IM Model). The TO Model simulates the behavior of the TOs to determine their demand to use the infrastructure, their willingness to pay for access, and the fares they would charge to the end users. The IM Model optimizes the timetable design considering the demand from TOs and all the technical constraints from the infrastructure. The results obtained are the demand to schedule trains, the access charges (capacity pricing), and the final train timetable (capacity allocation: set of trains scheduled and their timetable). We use this information to analyze the performance of both capacity pricing and allocation mechanisms from the perspective of the IM (cost recovery), the TOs (access charges, trains scheduled), and the end users (number of services, fares).

The results of this paper show that, in the context of the NEC, an auction mechanism could result in almost 20% more IM cost recovery and trains scheduled as compared to a cost-allocation and priority-rule mechanism. However, it also results in lower profits for the TOs. An auction mechanism also requires that the IM is able to solve the train timetabling problem to determine the set of trains to schedule. A cost-allocation and priority-rule mechanism, on the other hand, ensures higher profits for the TOs, making the TOs more resilient to uncertainty in end-users transportation demand. This mechanism is not very resilient to uncertainty in infrastructure capacity availability. Under an auction mechanism, intercity TOs are in better position than commuter TOs to access the tracks with current NEC levels of service. The priority level of each TO is a design choice in price-based mechanisms. This choice, however, has important implications for NEC commuter and intercity passengers and TOs, especially if the IM does not have access to sophisticated methods to solve the train timetabling problem.

Note that both the TO Model and the IM Model used are simple in nature. The value of this effort is two-fold. First of all, these models rely on information readily available for regulators. More detailed models would require that regulators have access to extensive information about the TOs and IMs. Second, this paper represents a first effort to compare price-based and capacity-based mechanisms in the same shared railway system. These models, although simple, already capture the main essence of the interaction of intercity and commuter TOs in shared railway system and allow us to compare both pricing and allocation mechanisms. An improved
understanding of how these mechanisms perform in the NEC allows NEC stakeholders to choose the mechanism that captures best the overarching objectives.

Although this paper focuses on the interaction between intercity and commuter TOs, the framework proposed is valid to analyze other aspects of the performance of shared railway systems as well.

The rest of the paper is structured as follows: Section 2 describes the framework used to evaluate the performance of shared railway systems under alternative capacity pricing and allocation mechanisms. Section 3 presents the results obtained using that framework to evaluate the performance of the NEC under both price-based and capacity-based mechanisms. Section 4 summarizes the main conclusions of the paper and identifies lines of future research.

2. SHARED RAILWAY SYSTEM PERFORMANCE EVALUATION FRAMEWORK

This section describes the framework we use to evaluate the performance of shared railway systems under alternative capacity pricing and allocation mechanisms. We first describe the TO Model that simulates the behavior of the TOs to determine their demand to use the infrastructure, their willingness to pay for access, and the fares they would charge to the end users. We then describe the IM Model that optimizes the timetable design considering the demand from TOs and all the technical and operational constraints. The outputs of these two models, the demand to schedule trains, the access charges (capacity pricing), and the final train timetable (capacity allocation: set of trains scheduled and their timetable) are then used in the next section to evaluate the performance of the system.

2.1. Train Operator (TO) Model

The objective of this subsection is to develop a model to anticipate how rational TOs would operate under a capacity pricing and allocation mechanism, given the institutional and regulatory framework and the technical characteristics of the specific type of railway service that they provide. For the purpose of this research, we are interested in three main TOs’ operational decisions: 1) the number of trains operated, 2) the access charges paid to the IM to access the infrastructure, and 3) the fare or shipping rate charged to end users. The TOs’ level of control over these three decisions depends on the context.

This model is based on three assumptions: 1) different types of services are not substitutable, 2) each type of TO serves a single origin-destination (OD) pair, and 3) TOs are profit maximizing agents. In other words, we first assumes that the demand of intercity services does not depend on the demand for commuter or freight services, neither on the commuter or freight level of service. The same applies for the other types of services. We consider commuter services in different metropolitan areas as distinct services. We also assume that each type of TO is serving a single OD pair. We assume that intercity and freight TOs offer services between Boston and Washington, and commuter TOs offer services to/from downtown to the suburbs. Although this assumption does not hold in reality in the NEC, it is useful to determine average number of trains across services offered by a TO, average fares, and average access charges paid to the IM. This provides a basis to determine the order of magnitude of TOs operational decisions using aggregated cost and revenues information provided by TOs. In other words, the average fares and ridership capture the value that the Boston-Washington service provides by serving intermediate markets (such as Boston-New York or Philadelphia-Washington), but it does not provide with a detailed model of how the capacity of the train is filled up with passengers travelling to different OD pairs. Future research will analyze the nuances for each service segment in terms of train capacity and fares and ridership per market. Finally, the third
assumption implies that rational TOs determine the operational decisions with the objective of maximizing profits. If there are several operational decisions that yield on the same profit, the TOs prioritize the ones with maximum number of services.

TO profits can be determined by analyzing TO revenues and costs for a given number of trains, \( n \). There are three main types of costs that TOs face: the cost of accessing the tracks, \( ac(n) \) or track-access charges; fixed costs, \( fc \), such as the cost of buildings and the cost of purchasing cars and locomotives in the medium term; and variable costs of operating trains, \( vc \cdot n \), such as fuel, personnel, train maintenance, and train lease, if trains are being leased. The two main sources of revenue come from transporting users (passenger or freight) and from the government (subsidies). The revenues obtained from transporting users can be determined by multiplying the fare or shipping rate (\( f \)) by the demand transported. The demand transported is limited by either the capacity (reduced by a reasonable average loading factor) of the trains \( (c \cdot n) \) or by user demand \( (d) \). According to literature, user transportation demand depends fundamentally on the fare \( (f) \), the frequency of the service (proportional to \( \frac{1}{n} \) if we assume uniform services during the time period), and the travel time (\( tt \)) (Bebiano et al., 2014). While intercity passengers are typically more sensitive to the fare and the travel time, commuter passengers are typically more sensitive to the fare, the frequency, and on-time performance, and freight users tend to be sensitive to the fare. Government subsidies, \( s \), depend in general on the demand transported too.

Summarizing, the costs and revenues of a TO can be determined using the following formulas:

\[
\text{Cost}(n, ac) = fc + vc \cdot n + ac(n) \tag{1}
\]

\[
\text{Revenues}(f, n) = s(f, n) + f \cdot \min(d(f, n, tt), n \cdot c) \tag{2}
\]

where bold letters are used to denote the three main operational decisions. We will call them \textit{TO decision variables}. Note that some of these decision variables may be pre-determined or conditioned by the context. For instance, the fare of commuter services is typically set by the government. Likewise, access charges under cost-allocation and priority-rule mechanisms are fixed inputs for TOs.

We can use now equation (1) and (2) to determine the operational decisions to maximize profits:

\[
\max_{n,f} \text{ profits}(f, n, ac), \text{ i.e. to max}_{n,f} \text{ revenues}(f, n) - \text{ costs}(n, ac):
\]

\[
\max_{n,f} s(f, n) + f \cdot \min(d(f, n, tt), c \cdot n) - fc - vc \cdot n - ac(n) \tag{3}
\]

Fixed costs do not depend directly on the number of trains operated or the fare. If the subsidy does not depend directly on the number of trains or the fare then equation (3) is equivalent to:

\[
\max_{n,f} f \cdot \min(d(f, n, tt), c \cdot n) - vc \cdot n - ac(n). \text{ We can use this equation to determine the fare and the number of services as a function of the access charges.}
\]

Since the number of trains and the fares depend on the access charges, the access charges can be determined implicitly using sensitivity analysis. In the medium term, the TO faces fixed costs \( fc \) independently of the operational decision. However, if the TO decides not to operate any trains, the TO will not have any variable costs and it will not have to pay to access the infrastructure. Similarly, it will not receive any revenues from operations. As a result, we know that the minimum profit that the TO would accept is \( s - fc \), otherwise it would be better off simply not operating any trains. Assuming that \( ac(n) \neq 0 \) can be written as \( ac_f + ac_v \cdot n \) (fixed plus variable charges), the TO maximum willingness to pay to access the infrastructure can be determined using equation (4).

\[
s(f, n) + f \cdot \min(d(f, n, tt), c \cdot n) - fc - vc \cdot n - ac_f - ac_v \cdot n \geq s(0) - fc \tag{4}
\]
The implications of equations (3) and (4) depend on the context in which TOs operate. The context is determined by both the institutional and regulatory environment, and the technical characteristics of the type of railway service that the TO provide. In section 3, we solve equations (3) and (4) analytically to determine the operational decisions of TOs similar to Amtrak and the MBTA in the NEC, operating under two alternative capacity pricing and allocation mechanisms. See (Levy et al., 2015) for further details.

2.2. Infrastructure Manager (IM) Model

The TO Model developed in the previous subsection informs us of the desired operational decisions of each type of TO. However, each type of TO determines how many trains to schedule independently of other types of TOs. When the infrastructure is congested, as it is in the NEC, the demands of different types of TOs will likely be conflicting. In this subsection, we formulate an IM Model to analyze the final timetable under alternative capacity pricing and allocation mechanisms (train timetabling problem).

The model presented here determines the optimal set of trains that the IM can accommodate, assuming both a price-based (cost-allocation and priority-rule) mechanism and a quantity-based (auction) mechanism. For the cost-allocation and priority-rules mechanism we assume that the IM determines the access charges for each type of service. The set of trains than can actually be scheduled and their timetables is computed maximizing the number of trains scheduled multiplied by the priority level of each train. For the auction mechanism we assume that, at some predetermined frequency, the TOs will have the opportunity to submit bids. Each bid will consist of a list of trains that the TO wants to schedule on the infrastructure, the desired timetable for each train, and the access charges they are willing to pay to schedule each train. The IM will then determine the set of trains that can actually be scheduled, their timetable, and the access charges that the TOs will pay. We assume that the IM’s objective is to maximize revenue and cannot restrict access to the infrastructure beyond the infrastructure constraints (e.g., safety, infrastructure maintenance plans).

The differences between the IM models for each mechanism affect mainly the definition of the parameters and the choice of the objective function. The constraints however are related to the physical characteristics of the infrastructure and remain unchanged across mechanisms. The model formulation is discussed below.

2.2.1. Sets

\( i, i' \in \{1,...,I\} \) set of trains that the TOs would like to schedule.

\( j, j' \in \{1,2,...,J\} \) railway system stations.

2.2.2. Parameters

We use lower-case letters to denote parameters. The information that is provided as inputs for every train \( i \) is:

\( a_i \), the access charge for train \( i \) or the maximum access charge that the TO is willing to pay if train \( i \) is scheduled in an auction. It is important to note that the TO will only operate a train if that access charge is less than or equal to its willingness to pay for access, as determined using the TO Model.

\( t_{ij} \), desired timetable of train \( i \), defined as arrival and departure time of train \( i \) at every station \( j \) in the path of train \( i \).

\( pr_i \), the priority level of each train \( i \) under a cost-allocation and priority-rule mechanism.
$\Delta t_{ij} p_i$, maximum acceptable changes in the desired timetable of train $i$ and per-unit penalty imposed by the TO if the IM modifies the desired timetable to schedule other trains in the infrastructure. The penalty specifies the reduced access price that the TO is willing to pay if the desired timetable changes.

The information about the topology of the line and the type of service is represented by different matrices that indicate the origin and destination of each type of service, as well as the path of each service along the infrastructure (intermediate stops).

In addition, the topology of the tracks and the signaling system will determine the minimum safe headway (time elapsed) between consecutive maneuvers at every station:

$h_{i_{i'}} j$ minimum headway between consecutive arrivals/departures ($i, i'$) to/from station $j$.

The IM can set large minimum headway to ensure the reliability of the timetable (including time-slab to recover delays in the system).

2.2.3. Variables

We use capital letters for variables. The decision variables of this problem are:

- $S_i$ binary variable that indicates whether train $i$ is scheduled.
- $T_{ij}$ final timetable, arrival and departure time of every train $i$ scheduled at every station $j$ in the path of the train.
- $\Delta T_{ij}$ final train $i$ changed in the desired timetable at each station $j$. Note that these variables can be determined knowing $T_{ij}$ and vice versa.
- $O_{i_{i'}} j$ is a binary disjunctive variable with value 1 if train $i$ departs before train $i'$ at station $j$ and value 0 otherwise.

2.2.4. Objective Function

As discussed before, the objective of the problem is to determine which trains should be scheduled and when, in order to 1) maximize the number of priority trains scheduled under a cost-allocation and priority rules mechanism:

$$
\max \sum_i p r_i S_i
$$

or, 2) maximize the IM’s revenue when an auction mechanism is implemented:

$$
\max \sum_i a_i S_i - p_i \sum_j \Delta T_{ij}
$$

2.2.5. Constraints

The first set of constraints establishes the relation between the desired timetable and the final timetable of every train scheduled:

The timetable at each station $j$ in the path of train $i$ can be determined as:

$$
T_{ij} = t_{ij} + \Delta T_{ij}
$$

Note that the set of stations for each train $i$ is determined using the topology matrices.

To ensure that the timetable is feasible, the total scheduled stopping and travel time between consecutive intermediate stations $j, j'$ must be greater than or equal to the total stopping and travel time in the desired timetable:

$$
T_{ij} - T_{ij} \geq t_{ij'} - t_{ij}
$$

The maximum change accepted in the timetable is bounded by the maximum change in the desired timetable of the train defined by the TO:
\[-\Delta t_{ij} \leq \Delta T_{ij} \leq \Delta t_{ij}, \forall i, j\]  

The TO may impose additional conditions to define acceptable desired-timetable changes. That happens when the TO is not interested in operating the train if the departure from or the arrival at one major station changes. In this case, additional constraints are included to ensure that the timetable respects the TO’s requests if the train is scheduled.

The final set of constraints ensures that the timetable proposed by the IM can be accommodated by the existing infrastructure. The IM must ensure that the difference between maneuvers $i, i'$ of every pair of trains scheduled is greater than or equal to the minimum safe headway, so at least one of the following equations must hold:

\[T_{ij} - T_{i'j} \geq h_{i'j}\]  
\[T_{i'j} - T_{ij} \geq h_{i'j}\]  

These conditions can be expressed using the following constraints:

\[T_{ij} - T_{i'j} \geq h_{i'j} - M(O_{i'j} + 2 - S_i - S_{i'})\]  
\[T_{i'j} - T_{ij} \geq h_{i'j} - M(3 - O_{i'j} - S_i - S_{i'})\]

In these equations $M$ is a “big enough” number to ensure that one and only one of the equations (10) and (11) holds. The binary variable $O_{i'j}$ is used to automatically activate one and only one of the constraints depending on the value of the other variables. $O_{i'j}$ has value 1 if train $i$ departs before train $i'$ at station $j$. This problem has on the order of $O(I^2J)$ binary variables and is very difficult to solve for large $I$ (number of trains) or $J$ (number of stations) due to a large integrality gap.

We use the QARLP algorithm, an algorithm based on a linear programming relaxation of dynamic programming to solve the problem. See (Pena-Alcaraz et al., 2015) for further details.

3. NORTHEAST CORRIDOR RESULTS

The NEC is one of the most interesting shared railway systems worldwide because of the number of TOs that share the system (1 intercity TO, 8 commuter TOs, and 4 active freight TOs) and the amount of traffic that it handles. With over 2,000 trains per day, the NEC is one of the most congested railway corridors in the US. As we mentioned before, the need to introduce a market-based capacity pricing and allocation mechanism by 2015 (PRIIA, 2008) requires a better understanding of the trade-offs associated with different capacity pricing and allocation mechanisms. In this section we use the framework proposed to evaluate and compare the performance of the NEC under two alternative capacity pricing and allocation mechanisms: a cost-allocation and priority-rule mechanism proposed by Amtrak (Crozet, 2004; Gadner, 2013; Nash, 2005; Lopez-Pita, 2014; Texeira and Prodan, 2014) and an auction mechanism widely proposed in the railway economic literature (Affuso, 2003; McDaniel, 2003; Newbury, 2003; Perennes, 2014).

To use the framework proposed we need information about the system to be able to use the TO and the IM Models. The information required for the models can be collected from the annual TOs’ financial reports and the IM’s network report. As we mentioned before, this is a design choice of both models. A model that allows regulators to anticipate the system reaction to a capacity pricing and allocation mechanism should not require extensive information about the railway system that only the TOs and the IM possess. The focus of this analysis is the relation between intercity and commuter services during peak hours.
According to (Amtrak, 2014) an intercity TO like Amtrak operating in the NEC faces fixed operational (direct) costs of $f_c = 281k$ per day and variable operational costs of $v_c = 3,425$ per train and day. In 2013, Amtrak’s average fare was equal to $f_0 = 96.5$, the number of trains was $n = 150$ trains per day in average, with a realized demand of $d_0 = 31,250$ passengers per day. The average train capacity was $c = 210$ passengers assuming a physical capacity of 250 seats with 85% load factor (Amtrak, 2014). No subsidies are required for the operations of intercity services in the NEC (Amtrak, 2010; Amtrak, 2012). We assume that the demand of intercity trains depends linearly and exclusively on the fare, with an elasticity of $-0.67$ (Morrison, 1990).

According to (MBTA, 2013a; MBTA, 2013b) a TO like the MBTA, the commuter TO in the Boston area, faces fixed operational (direct) costs of $f_c = 435.1k$ per day and variable operational costs of $v_c = 1,666$ per train and per day. The elasticity of the demand with respect to the headway (frequency) is estimated by (Lago et al., 1981) to be equal to $-0.41$. In 2013, MBTA’s average fare ranged from $f_0 = 7 - 25$ (average fare of $f_0 = 13$ are considered), the number of trains averaged $n_0 = 485$ trains per day, with a realized demand of $d_0 = 130,6k$ passengers per day. The train average capacity considered is $c = 350$ passengers, with 80% + load factor. Subsidies $s = 234k$ per day are considered following (MBTA, 2013a). Commuter TOs in the NEC are subjected to fare regulation, i.e. they cannot change the fares charged to the end users.

To capture the main infrastructure characteristics of the NEC, we consider the system presented in Figure 1. It consists of a double-track corridor with 12 stations. Stations 1 and 7 are terminal stations at both ends of the line (Boston and Washington DC respectively). Stations 2-12, 3-11, 4-10, 5-9 and 6-8 represent five stations along the corridor. We use a different station number for each traffic direction. Traffic moves in the direction of increasing station numbers in a dedicated track per direction. As a result, traffic traveling in different directions only interacts at the stations. Stations 1, 2 and 12 represent main stations in Boston metropolitan area, stations 3, 4, 5, 9, 10 and 11 are all in New York metropolitan area, and stations 6, 7 and 8 are in Washington DC metropolitan area. Five types of services can be considered: Boston commuter trains traveling around the Boston metropolitan area (stations 1, 2, and 12); New York commuter trains; DC commuter trains; and intercity and freight trains traveling between Boston and Washington DC. Intercity and freight trains may not stop at every station. Freight trains travel the line at much slower speeds than commuter and intercity trains. Intercity trains travel at higher speeds than commuter trains. The distance between Boston and Washington DC is approximately 450 miles, and the distance travel by commuter agencies around each city ranges from 40 to 70 miles per direction. Note again that although the infrastructure considered is simple and does not include many intermediate stations such as Philadelphia, New Haven, etc., it contains all the important elements to be able to compare how commuter trains will interact with intercity trains under both capacity pricing and allocation mechanisms.

![Figure 1. Detailed corridor infrastructure (source: authors)](image-url)
3.1. Cost-Allocation and Priority-Rule Based Mechanism

In this subsection we analyze the implications of a cost-allocation and priority-rule based mechanism. We first use the TO model to anticipate the number of trains that a TO would like to operate and fares the TO would charge to the end user for different values of variable access charges. We also determine the resulting TO profits. Figure 2 shows the results obtained for an intercity TO in the NEC with Amtrak’s cost structure. Figure 3 shows the results for a commuter TO in the NEC with MBTA’s cost structure. Note that we do not show the fares in Figure 3 because the TO cannot change them. We use a distance of 50 miles for commuter trains (Boston, MA to Cranston, RI) and a distance of 450 miles for intercity trains (Boston, MA to Washington, DC) to normalize the variable access charges per mile (Gadner, 2013).

![Figure 2](image)

**Figure 2.** NEC intercity TOs expected profits per day, number of trains per day, and fares for different variable access charges (source: authors)

Although both figures only show the response of TOs to variable access charges (per train), we can use them to determine the response of the TOs to fixed access charges too. Essentially, a fixed access charge in addition to the variable access charge would not change the TO operational decisions (number of trains and fares) as long as the resulting TO profits are greater than the profits it would obtain operating 0 trains. If the resulting profits were lower than the
profits operating no trains, a profit maximizing TO would decide not to get access to the infrastructure (and pay no access charges). Otherwise, the TO would see the fixed access charge as a fixed lump sum that would not change the optimality conditions in the profit equation and as a result, would not change its operational decisions either.

We first see that the number of trains that TO would like to schedule decreases as variable access charges increase. We can also use Figures 2 and 3 to determine the maximum access charges that both intercity and commuter TOs are able to pay. Table 1 summarizes this information. In particular, the maximum access-charge that an intercity TO like Amtrak would be able to pay is $102 per train-mile per day, which is equivalent to $46,000 per train per day for Boston to Washington. The maximum variable access charge that a commuter TO like MBTA would be able to pay is $51.55 per train-mile per day, which is equivalent to $2,578 per train per day. With higher variable access charges the TOs would be better-off not operating any trains. That means that an intercity TO is able to pay two (2) times as much as a commuter TO in a per mile basis, or almost eighteen (18) times as much as a commuter TO in an absolute basis.

We also determine the maximum sustainable access charges (access charges for which the TOs would have 0 profits after reimbursing capital at an adequate rate of return). That means that their finances allow them to continue operations over the medium term.

<table>
<thead>
<tr>
<th>reference point</th>
<th>no access charges</th>
<th>maximum access charges</th>
<th>maximum sustainable access charges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercity</td>
<td>-0.28</td>
<td>0.00</td>
<td>116</td>
</tr>
<tr>
<td>Commuter</td>
<td>-0.20</td>
<td>0.00</td>
<td>450</td>
</tr>
</tbody>
</table>

Note again, that these numbers can be used to anticipate the response of the TOs to fixed access charges. The maximum fixed access charges that an intercity TO would be willing to pay would be $2.75 M per day. With this access charges, it would still operate 116 trains, and its profits would be $– 0.28 M ($2.47 M – $2.75 M). If the access charges increase at that point, the TO would only operate 0 trains. Otherwise the TO's profits would be smaller than the ones operating...
The TO Model anticipates the response of the TOs when each of them optimizes their operational decisions on their own. As a result, the revenues and profit presented in Figures 2 and 3 and Tables 1 and 2 assume that the TOs are able to schedule all the trains on the infrastructure. However, the NEC is very congested and scheduling all the trains may not always be possible. We use the IM Model to determine the optimal timetable for different variable access charges. The TO Model shows that as access charges increase, TOs’ demand to schedule trains would be smaller than the current level of operations. Consequently, the IM Model results show that the IM should be able to schedule all the trains in the system if TOs are willing to compromise their desired schedule to accommodate conflicting trains. However, under this mechanism some TOs are granted priority over others by the priority rules. As a result, priority TOs may not have incentives to be flexible on their trains scheduling preferences. This will have a direct impact on the other TOs ability to schedule trains in the system, their profits, and on the total revenues collected by the IM.
3.2. Auction Based Mechanism

In this subsection we analyze the implications of an auction mechanism. We start analyzing the optimal capacity allocation plan (train timetable) to determine how to coordinate different TOs’ conflicting demand to schedule trains. Figure 4 shows the time-space diagram for the optimal timetable designed using the IM Model in a case in which an intercity TO tries to schedule one train and three commuter TOs try to schedule commuter trains around Boston, New York, and Washington DC every 30 minutes. The y-axes represent distance in miles from station 1 and the x-axes represent time in minutes at which different trains are scheduled to pass through each point of the line (vs. desired scheduled in dashed line). There are no interactions between trains traveling in different directions. The IM Model proposes the final timetable analyzing the trade-off between eliminating trains and readjusting the desired schedules, according to the objective function in equation (6). We can use this information to determine how much intercity TOs will have to pay to be able to schedule services that conflict with commuter services.

For this example, we assume that each commuter TO pays 1 unit per train-mile to schedule a commuter service and gets a 5% discount from the original access charge for every minute that one of their train schedules is changed. To analyze the first case, we need to solve a train timetabling problem with 115 commuter trains and 1 intercity train. For clarity purposes, only the schedules of conflicting trains are shown in Figure 4.

As Figure 4 shows, the intercity train would initially conflict with 14 commuter trains. Rescheduling the commuter trains to accommodate the intercity service requires that the commuter TOs receive a discount of 150 units on their total access charges. As a result, the IM would only schedule the intercity train if it represents more than 150 units of revenue, i.e., if its bid is higher than 0.33 units per train-mile. This number does not change with the desired timetable of the intercity train. The same results are expected when the frequency of commuter trains is higher than 2 trains per hour, i.e., the number of conflicting trains do not depend on the
intercity train desired timetable. That means that we do not need to know the exact desired timetable of the trains to determine the relationship between intercity and commuter train bids in the NEC if the frequency of commuters is greater than 2 trains per hour.

If the frequency of commuter trains increases, for example to one commuter train every 15 minutes instead of every 30 minutes, the intercity train will initially conflict with 22 commuter trains and will only be scheduled if it represents more than 371 units of revenue for the IM (i.e., if the intercity TO bid is higher than 0.82 units per train-mile). Conversely, if the frequency of commuter trains decreases to one train every 60 minutes, the intercity train will be scheduled almost always (the IM would need to recover between 0 and 28 units of revenue depending on the desired timetable, what translates in bids higher than 0.00 or 0.06 units per train mile). The model can be used to quantify the trade-off between commuter and intercity trains for any other frequency of service (see Table 3).

Table 3 shows that, when the system is congested, an intercity TO may have to pay more than a commuter TO to schedule a train: the intercity TO has to pay between 0.82 and 5.86 times the access charges of commuter TOs per train-mile or between 5.27 and 37.67 times the commuter TOs’ access charges per train. **This minimum intercity access charge reflects the congestion rents.** The results show that greater cost recovery is expected in congested infrastructure. How can we know whether intercity TOs would be able to compete to access the infrastructure with the commuter TOs? So far we assume that commuter TOs pay 1 unit per train-mile. Is there a way to anticipate TOs bid? Fortunately, we can use the TO model (equation 4) to determine the maximum access charges that each TO would be willing to pay as a function of the number of trains they want to schedule.

**Table 3. Minimum intercity to commuter access-charge per train-mile bid ratio to ensure that their train is scheduled as a function of the commuter frequency (source: authors)**

<table>
<thead>
<tr>
<th>Commuter frequency [minute]</th>
<th>ratio [units per train-mi]</th>
<th>Commuter trains scheduled [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5.86</td>
<td>73%</td>
</tr>
<tr>
<td>10</td>
<td>0.84</td>
<td>100%</td>
</tr>
<tr>
<td>15</td>
<td>0.82</td>
<td>100%</td>
</tr>
<tr>
<td>30</td>
<td>0.33</td>
<td>100%</td>
</tr>
<tr>
<td>60</td>
<td>0.00-0.06</td>
<td>100%</td>
</tr>
</tbody>
</table>

Figure 5 summarizes these results, the maximum variable access charges that an intercity and a commuter TO with the cost structure of Amtrak and MBTA respectively could bid as a function of the number of trains to schedule. The results show again that the intercity TO ability to pay to access the infrastructure is almost double the one from the commuter TO counterparts. We need to make one adjustment before we use these results as inputs of the IM Model. The TO Model assumes that all the trains have the same OD pair. However, the 150 intercity services that (Gadner, 2013) mentions, include for instance Boston to New York services and New York to Washington DC services that we count in the IM Model as a single service. We use the following equivalences between frequency and number of trains: 118 intercity services in the TO Model are equivalent to 1 train per hour in the IM Model and 450 commuter services in the TO Model are equivalent to 6 commuter trains per hour in the IM Model (Amtrak, 2014; Amtrak, 2015; MBTA, 2013a; MBTA, 2015).

Table 4 shows the result of combining the information in Table 3 and Figure 5. The first three columns show the bids of each commuter TO as a function of the desired frequency (number of trains to schedule). The next three columns show the bid of an intercity TO that tries to schedule
1 train per hour. These results assume that each TO reveals its maximum willingness to pay to access the infrastructure on the bids. Section 3.3. discusses how these results would change if the TOs use their monopoly power. The last two columns determine how many trains of each type can be scheduled and compute the resulting revenues for the IM (again, assuming three commuter TOs).

Figure 5. NEC intercity and commuter TOs maximum willingness to pay for access (maximum variable access charges) as a function of the number of trains to schedule (source: authors)

Table 4. TOs’ demand to schedule trains for different variable access charges and resulting IM expected revenues per day assuming three commuter TOs (source: authors)

<table>
<thead>
<tr>
<th>Commuter frequency [minute]</th>
<th>N commuter TO [trains]</th>
<th>(ac_c) commuter TO [$ per train-mi]</th>
<th>(n) intercity TO [trains]</th>
<th>(ac_i) intercity TO [$ per train-mi]</th>
<th>(n) commuter n intercity TO [trains]</th>
<th>Revenues IM [$M]</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>900</td>
<td>11.54</td>
<td>81</td>
<td>68.05</td>
<td>657, 81</td>
<td>4.00</td>
</tr>
<tr>
<td>10</td>
<td>450</td>
<td>39.73</td>
<td>118</td>
<td>51.34</td>
<td>450, 118</td>
<td>6.30</td>
</tr>
<tr>
<td>15</td>
<td>300</td>
<td>51.25</td>
<td>118</td>
<td>51.34</td>
<td>300, 118</td>
<td>5.80</td>
</tr>
<tr>
<td>16</td>
<td>280</td>
<td>51.55</td>
<td>118</td>
<td>51.34</td>
<td>280, 118</td>
<td>5.61</td>
</tr>
<tr>
<td>28</td>
<td>160</td>
<td>2.16</td>
<td>118</td>
<td>51.34</td>
<td>160, 118</td>
<td>2.80</td>
</tr>
<tr>
<td>30</td>
<td>150</td>
<td>0</td>
<td>118</td>
<td>51.34</td>
<td>0, 118</td>
<td>2.73</td>
</tr>
</tbody>
</table>

Note that the intercity to commuter TO bid ratio exceed the ratio presented in Table 3 when the commuters frequency (headway) is bigger than 5 minutes. In other words, the intercity TO is almost always able to schedule all the intercity services. These results suggest that the intercity TO in the NEC is usually in better position to access the tracks than the commuter TO under
an auction mechanism with current levels of service. If the frequency of commuter trains where to increase around 85%, with 5 minutes headways, the intercity TO would not be able to schedule trains if it bid less than $67.62 per train-mile (5.86 x $11.54 per train-mile). Using the TO Model, we can determine that the intercity TO would still be able to bid over $67.62 per train-mile (specifically $68.05 per train-mile) when it tries to schedule 81 trains or less.

In this case, the commuter TO would only be able to schedule 657 trains (73% of 900). This equilibrium is stable because none of the TOs would want to schedule more trains. As Figure 5 shows, the commuter TO would be willing to pay higher access charges for 657 trains than for 900 trains. As a result, scheduling only 657 trains at the 900-train access charge level translates in extra profits for the TO. Similar results are obtained for the intercity TO bids and for all other commuter TO bids with more than 280 trains. Between 160 and 280 commuter trains the equilibrium is not stable because the demand for commuter services would significantly decrease due to the amount of service reduction, and also the commuter TO profits when not all trains are scheduled.

3.3. Comparison

The previous subsections discuss the operational decisions of intercity and commuter TOs under a price-based (cost-allocation and priority-rule) mechanism and a capacity-based (auction) mechanism in the NEC. The results are summarized in Tables 2 and 4 respectively.

Table 5 shows the number of trains that each TO scheduled and its profits for different access charges, together with the revenues raised by the IM under both mechanisms. Although there is not a one-to-one comparison between both mechanisms, we can compare both sides of the table.

Table 5. Distribution of profits and number of trains scheduled per day comparison (source: authors)

<table>
<thead>
<tr>
<th>Cost-allocation &amp; Priority-rule Mechanism</th>
<th>Auction Mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercity TO ac, n, profits [Str-mi, tr, SM]</td>
<td>commuter TO ac, n, profits [Str-mi, tr, SM]</td>
</tr>
<tr>
<td>0.0, 116, 2.5</td>
<td>0.0, 450, 0.7</td>
</tr>
<tr>
<td>25.0, 88, 1.3</td>
<td>25.0, 340, 0.5</td>
</tr>
<tr>
<td>50.0, 60, 0.5</td>
<td>50.0, 284, -0.2</td>
</tr>
<tr>
<td>51.3, 59, 0.4</td>
<td>51.3, 282, -0.2</td>
</tr>
<tr>
<td>51.6, 59, 0.4</td>
<td>51.6, 282, -0.2</td>
</tr>
<tr>
<td>53.3, 57, 0.4</td>
<td>53.3, 0, -0.2</td>
</tr>
</tbody>
</table>

These results show that the revenues collected by the IM under the auction mechanism studied are higher than the revenues collected when a cost-allocation and priority-rule with similar charges for intercity and commuter TOs is implemented. The auction mechanism does not only allows the IM to collect higher revenues (around 20% more than using the cost-allocation and priority-rule mechanism in the case highlighted), but also results in higher number of trains scheduled (20% more) and higher total welfare (also 20% higher) as compared to the cost-allocation and priority-rule mechanism case. Note however, that these advantages have a cost for the TOs, who retain much lower profits (negatives in most cases). As a result, an auction mechanism may require the design of a procedure to redistribute revenues to ensure that TOs can sustainably operate over the medium term.

It is important to note that both the TO Model and the IM Model considered in this research are deterministic. They intend to capture the essence of a normal day of operations. However, the TOs and the IM face several sources of uncertainty when making their operational decisions. The
two most important sources of uncertainty in the NEC in the medium term are 1) the condition of the infrastructure, 2) the end users demand for transportation services.

The first source of uncertainty is particularly critical until the NEC reaches a state of good repair that ensures that the infrastructure is reliable. With today’s backlog of maintenance work, the need of last-minute maintenance and interventions has a direct impact on the capacity on the corridor and on the TOs ability to schedule trains. This uncertainty gets amplified under cost-allocation and priority-rule mechanisms. Any problems with infrastructure availability under price-based mechanisms would reduce the number of services operated by TOs that often operate the minimum number of services that allows them to be profitable. This lack of infrastructure capacity unevenly affects those TOs with lower priority assigned. Under an auction mechanism, the TOs will still make profits even if some trains are not scheduled due to infrastructure availability problems. The IM under an auction mechanism would have important incentives to avoid uncertainty on the infrastructure capacity availability, since fewer trains scheduled would lower its ability to recover infrastructure costs as compared to the deterministic case.

The uncertainty in the demand for transportation has a major impact on the expected revenues that the TOs would collect. The fact that the cost-allocation and priority-rule mechanism ensures high TO profits makes this mechanism more resilient to demand uncertainty than the auction mechanism, where the TOs operational decision will probably change if the TOs expect a very uncertain demand. Note that there is also uncertainty in the demand distribution. Passengers do not arrive homogeneously during the day; they instead concentrate around some particular times. As a result we may expect TOs scheduling some more trains than the ones that the model indicates. For example, while the commuter TO Model shows that the optimal number of trains to schedule is 397, MBTA currently runs 485 trains in the line. Although the model already considers a load factor of 80% to accommodate part of this demand, scheduling 485 commuter trains would result in an average load factor of 65% (industry benchmark for commuter services). This would result in a higher operational cost to accommodate the same demand. The need to offer 485 vs. 397 trains would depend on the exact distribution of the demand. The load factor of the model can be adjusted to consider this uncertainty. This uncertainty will propagate to the expected profits. As a result, we may expect to see a lower TOs’ willingness to pay to access the infrastructure even for the same number of trains to schedule, and hence a lower IM ability to recover infrastructure costs than in the deterministic case.

As we mentioned before, these results also assume that TOs would not take advantage of their position to game the mechanism to their interest. This is particularly important under auction mechanisms where the TOs have incentives to keep lowering their bids while their trains get scheduled to maximize their profits. Note however, that the framework proposed in this paper allows the regulator to anticipate the results of the auction and to investigate any variation with respect to these numbers.

4. CONCLUSIONS

The objective of this paper is to identify some of the trade-offs involved in the choice among alternative capacity pricing and allocation mechanisms for shared railway systems in the context of the NEC. In particular, this paper analyzes how the introduction of two alternative pricing and allocation mechanisms impact the ability of intercity and commuter TOs to compete for the access to infrastructure capacity. The two alternative capacity pricing and allocation mechanisms evaluated are a cost-allocation and priority-rule mechanism, which was proposed by Amtrak and
is currently being considered by the TOs in the NEC, and an auction mechanism, which are widely proposed in the literature but have not yet been implemented in any country.

This paper proposes a framework to evaluate the performance of shared railway systems considering both planning and operational aspects. This framework integrates a train operator (TO) model and an infrastructure manager (IM) model that allow us to anticipate the demand to schedule trains, set the access charges (capacity pricing), and set the final train timetable (capacity allocation: set of trains scheduled and their timetable). These models are simple by design choice. The main objective is to allow regulators anticipate the performance of both capacity pricing and allocation mechanisms from the perspective of the IM (cost recovery), the TOs (access charges, trains scheduled), and the end users (number of services, fares). More detailed models would rely on extensive information about the TOs and IMs that is not readily available for regulators. These models, although simple, already capture the main essence of the interaction of intercity and commuter TOs in shared railway system and allow us to compare both pricing and allocation mechanisms.

The results of the paper show that, an auction mechanism could result in almost 20% more IM cost recovery and trains scheduled as compared to a cost-allocation and priority-rule mechanism in the NEC. However, it also results in lower profits for the TOs. Under an auction mechanism, intercity TOs are in better position than commuter TOs to access the tracks with current NEC levels of service. A cost-allocation and priority-rule mechanism, on the other hand, ensures higher profits for the TOs, making the TOs more resilient to uncertainty in end-users transportation demand. Note again that this comes with a cost to users (that will have fewer trains) and to the infrastructure manager (that will obtain fewer revenues from access charges). Although TOs can easily understand priority rules and anticipate the number of trains that they will be able to schedule; this mechanism is not very resilient to uncertainty in infrastructure capacity availability. The priority level of each TO is a design choice in price-based mechanisms. This choice has important implications for NEC commuter and intercity passengers and TOs.

To the best of our knowledge, this is the first paper that compares the performance of price-based and capacity-based mechanisms in the same railway system. The results show that neither of these two mechanisms is superior to the other on all dimensions. A better understanding of these trade-offs though could help the regulator to design other mechanisms to price and allocate capacity. Note that ultimately, the design and choice of the most appropriate capacity pricing and allocation mechanism would depend on the overarching regulatory objectives.

Although this paper focuses on the interactions between intercity and commuter TOs, the framework proposed is valid to analyze other questions such as the implications of the mechanisms for freight TOs, for the end users, or for the whole system. We believe that the stakeholders in the NEC should carefully analyze the implications of alternative pricing and allocation mechanisms before locking the system into one of them. In particular, detailed studies that consider the variety of services that each TO offers in the NEC (services with different speeds and stops, serving different OD pairs) should be carried out to refine the understanding of the implications of alternative pricing and allocation mechanisms. Future research should also analyze how these results changed in the context of other congested and non-congested shared railway systems.

REFERENCES
Group: Washington, DC.
MBTA, (2013a). *Financial statements, required supplementary information and supplementary information*. Boston, MA.


