Integrated train scheduling and dynamic pricing with endogenous demand under uncertainty: A machine learning-enhanced Benders decomposition approach

Yahan Lu*[1], Dongyang Xia[1], and Shadi Sharif Azadeh[1]

[1]Department of Transport & Planning, Delft University of Technology, Netherlands

Short summary

This study develops an integrated optimization framework for mitigating crowding in urban rail transit systems through the joint design of train scheduling, dynamic pricing, passenger arrival time shifting, and passenger flow control under demand uncertainty. Unlike conventional studies that treat time-varying demand as exogenous, we explicitly model fare-dependent passenger arrival time choice. The problem is formulated as a stochastic program that minimizes total passengers' travel time subject to train capacity limits, service priority for reserved passengers, and a government subsidy budget. To solve the model, we develop an improved Benders decomposition method further enhanced by machine learning. Numerical results show that the proposed approach effectively reduces passengers' waiting time and improves computational efficiency.
Keywords: Integrated learning and optimization; Urban rail transit; Arrival time choice behavior; Stochastic optimization; Machine learning; Benders decomposition

1 Introduction

In rail transit systems such as railway and metro, capacity is substantially more flexible in the sense that train capacity can often be adjusted during the sales horizon through operational measures such as carriage reallocation or modular train composition. This feature distinguishes rail transit from traditional revenue management sectors such as airlines, hotels, and fashion retail, where supply is usually fixed once capacity is released to the market. As a result, in rail transit systems, pricing decisions and supply-side resource allocation are inherently intertwined, and their joint optimization becomes of genuine practical and theoretical importance.

Several rail transit systems worldwide have piloted dynamic pricing policies through fare discounts. For example, since January 2017, the Beijing metro has implemented off-peak discounts at 24 stations on three lines that are particularly crowded during the morning peak hours (Beijing Subway, 2017). Specifically, entries before 7:00 am on weekdays at these stations receive a 50% discount, with the destination unrestricted anywhere in the network. The explicit purpose is to encourage travelers departing from these bottleneck origins to shift their trips from the peak period to the early off-peak period. Dutch domestic rail on dense short-distance corridors functions in many respects similarly to metro systems such as the Beijing metro. Services are frequent, intercity travel times are short, and ticketing is largely non-reserved, so ticket purchase does not guarantee a seat. For instance, the rail connection from Delft to Rotterdam takes 11 minutes, and 12 trains operate between 7:00 and 8:00 AM, indicating substantial commuter demand on such corridors. In this context, dynamic pricing mainly serves to influence passengers' arrival time choices at their origin stations and shift part of the peak demand toward less crowded off-peak intervals. Reflecting this rationale, starting from 1 May 2025, the Dutch national rail operator Nederlandse Spoorwegen (NS) introduced the time-varying *NS PriceTime Deal* to actively smooth demand over the day. One-way e-tickets receive discounts of up to 60% depending on the selected time slot, expected crowding, day of travel, and advance purchase interval (AmsterdamTips, 2024; Nederlandse Spoorwegen , 2025).

In the literature, pricing and supply-side resource allocation in transportation and logistics are still most often studied separately. Studies that jointly optimize pricing and supply-side decisions remain limited, especially in rail transit. Among the few existing contributions, Manchiraju et al. (2024) study pricing and capacity allocation on a single railway line by comparing dynamic versus static pricing and fixed versus flexible train capacity policies. Their results show that when demand

is high and services are likely to be fully utilized, dynamic pricing is particularly effective, whereas when demand is relatively low, flexibly adjusting train capacity becomes more valuable and the additional benefit of dynamic pricing is more limited. Although insightful, that study is based on an exogenously given timetable and treats each origin-destination (OD) pair independently. In practice, however, when train timetables can be adjusted and train capacity is limited, congestion and passenger spillover create strong interactions across OD pairs and across time. In such a setting, pricing not only changes ticket revenue, but may also alter the temporal distribution of demand, which in turn affects in-vehicle crowding and passengers' waiting times. This raises a fundamental question: if train schedules and demand management can be jointly coordinated, how should pricing be designed when passenger demand over time is itself endogenous?

We address this question by making the following contributions. (i) We propose an integrated optimization framework that jointly optimizes train scheduling, dynamic pricing, arrival time shifting, and passenger flow control under demand uncertainty. (ii) We formulate a stochastic optimization model, where the temporal distribution of demand is endogenized by explicitly modeling arrival time shifting, rather than treating time-varying demand as a fixed input. (iii) We incorporate passengers' arrival time choice behavior using a multinomial logit formulation with fare-dependent choice probabilities, which allows pricing decisions to endogenously shape the temporal distribution of demand. Building on the structural properties of the choice model, we derive tractable constraints to behaviorally bound feasible demand shifts. (iv) An improved Benders decomposition framework for the resulting stochastic optimization problem is developed. (v) We further integrate machine learning techniques into the Benders decomposition algorithm to accelerate computation while preserving the exactness of the proposed algorithm. It is general and can be applied to other stochastic programs.

2 Methodology

We study an integrated optimization problem for mitigating crowding in urban rail transit systems through the joint design of train scheduling, dynamic pricing, passenger arrival time shifting, and passenger flow control under demand uncertainty. The key methodological feature of our framework is that the temporal distribution of passenger demand is treated as endogenous rather than fixed. Specifically, fare discounts offered in earlier off-peak time intervals may induce a portion of peak-period passengers to shift their arrival times, thereby reshaping the demand pattern faced by the operator.

To capture this mechanism, we incorporate a fare-dependent passenger arrival time choice model into the optimization framework. A multinomial logit model is used to describe passengers' preferences over alternative arrival intervals, where the choice probabilities depend explicitly on ticket prices and schedule-related disutility. Based on the structural properties of the choice model, we derive a set of tractable linear constraints that bound the amount of peak demand that can be shifted to each earlier off-peak interval. In this way, the behavioral response of passengers is embedded into the optimization model without introducing a fully nonlinear choice-based assignment problem.

Building on this behavioral component, we formulate a two-stage stochastic optimization model. In the first stage, the operator determines train departure times and time-dependent fare discounts before the realization of uncertain demand. In the second stage, after the demand scenario is observed, passenger flow control and arrival time shifting decisions are adjusted in a scenario-dependent manner. The model aims to minimize total passenger travel time while satisfying train capacity limits, flow conservation conditions, and a government subsidy budget constraint. In addition, passengers with reservations are given service priority, while non-reserved passengers are managed through controlled boarding and temporal demand shifting.

The resulting model is computationally challenging due to the strong coupling among pricing, scheduling, shifting, and flow control decisions, as well as the presence of multiple demand scenarios. To solve it efficiently, we first develop an improved Benders decomposition approach. We further integrate machine learning into the decomposition framework to reduce unnecessary subproblem solves while preserving exactness.

3 Results and discussion

We examine the impact of the relative behavioral sensitivity parameters in the proposed fare-dependent systematic utility function on total passengers' waiting time and the proportion of passengers shifting away from the peak period. The fare-dependent systematic utility $U_{svt't}(\tilde{f}_{svt'})$ of choosing arrival time $t' \in \mathcal{A}(t)$, conditional on the originally planned arrival time t, is specified as

$$U_{svt't}(\tilde{f}_{svt'}) = -\beta_P \, \tilde{f}_{svt'} - \beta_E \, \text{EARLY}_{t't} - \beta_K \, \mathbf{1}\{t' \in \hat{\mathcal{T}}\}\text{CROWD}_{t'} \; \forall s, v \in \mathcal{S}, v \geq s+1, t \in \mathcal{T}, t' \in \mathcal{A}(t), \tag{1}$$

where the decision variable $\tilde{f}_{svt'}$ denotes the fare paid when the passenger arrives at the origin station in time interval t'. In addition, the component $\text{EARLY}_{t't}$ measures the number of intervals by which the chosen arrival interval precedes the originally intended one, i.e., $\text{EARLY}_{t't} := \max\{0, t - t'\}$. The function $\mathbf{1}\{t' \in \hat{\mathcal{T}}\}$ equals 1 if t' belongs to the peak period, thereby capturing residual peak-hour disutility (e.g., crowding or inconvenience) beyond fare and arrival-time adjustment. The parameter $\text{CROWD}_{t'}$ denotes the penalty of peak-hour disutility. The parameters satisfy $\beta_P > 0$, $\beta_E > 0$, and $\beta_K > 0$.

To do so, we construct two sets of experiments. In the first set, we fix the coefficient associated with peak-period disutility, β_K, at 10, and vary the coefficients associated with fare, β_P, and the penalty for earlier arrival at the destination, β_E. In the second set of experiments, we fix β_P and β_E and vary β_K. In both sets, the peak-period inconvenience penalty $\text{CROWD}_{t'}$ is fixed as 10. The maximum allowable subsidy ratio is set as 0.5, i.e., $\beta = 0.5$. In these two sets of experiments, all instances are solved to optimality.

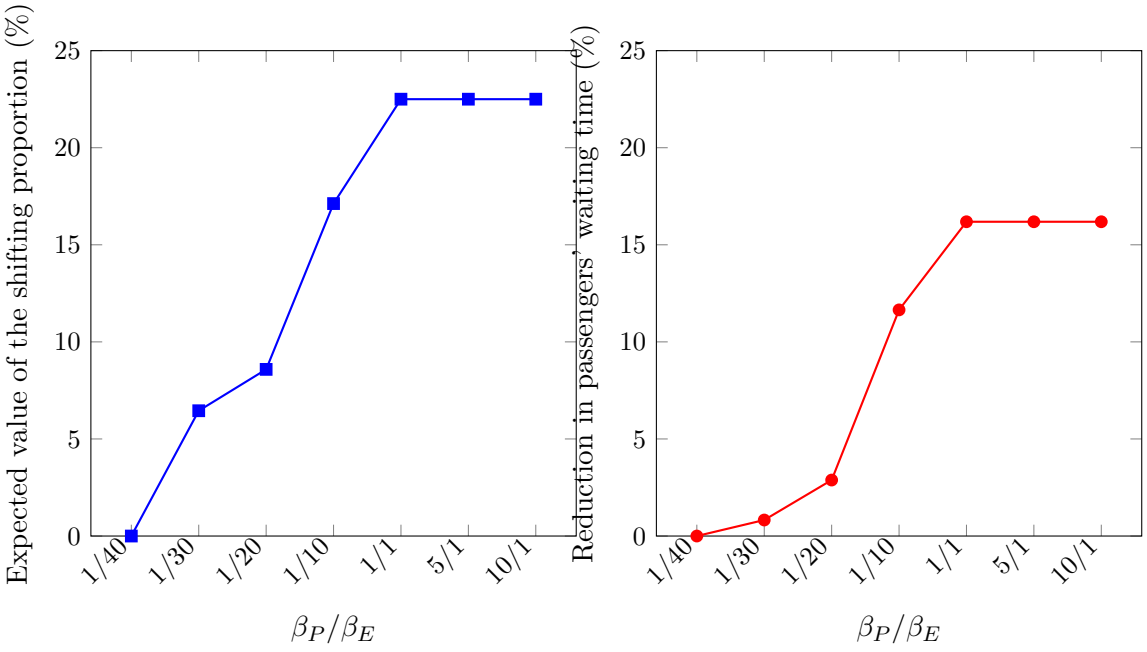
FINDING. *The greater the relative sensitivity to fare, relative to the disutility of earlier arrival, the more likely peak-period passengers are to shift to the off-peak period. In addition, the arrival time shifting strategy is effective in reducing the objective value, namely total passengers' waiting time, including both initial waiting time and secondary waiting time caused by being stranded.*

Figure 1(a) reports the expected value of the shifting proportion. The value on the vertical axis is computed as the number of passengers shifted from the peak period to the off-peak period divided by the total number of peak-period passengers before shifting, multiplied by 100 (%). Since all shifted passengers move from the peak period to the off-peak period, this measure can also be interpreted as the percentage reduction in the passenger demand during peak periods. Figure 1(b) reports the relative decrease in the objective value (i.e., passengers' total waiting time), compared with the benchmark in which no passenger shifts. Here, total waiting time includes both initial waiting time and secondary waiting time caused by being stranded.

It can be observed that the expected value of shifting proportion rises from 0 to about 22.6% as β_P/β_E increases from 1/40 to 10/1. This is because a larger value of β_P/β_E indicates that passengers are more responsive to fare discounts relative to the disutility of arriving early. A second observation is that the relative reduction in passengers' total waiting time follows a similar trend, increasing from 0 to about 16.2%. These results indicate that, when a larger share of peak-period demand is shifted to the off-peak period, a greater reduction in total waiting time can be achieved. This finding confirms that the shifting strategy is effective in reducing passengers' total waiting time by transferring part of the peak-period demand to the off-peak period. Moreover, both curves gradually level off when fare sensitivity becomes sufficiently strong relative to the disutility of earlier arrival. This suggests that, beyond a certain point, further strengthening passengers' responsiveness to fare discounts does not generate additional improvements.

4 Conclusions

This paper addresses the joint optimization of train scheduling, dynamic pricing, and multiple demand management plans in an urban rail transit system with endogenous and uncertain passenger demand. Extensive experiments on small-scale and real-world case studies have demonstrated the effectiveness of the proposed modeling and algorithmic framework. The research findings provide several key managerial and methodological insights. From a managerial perspective, the joint optimization of dynamic pricing, train schedules, and demand management can effectively smooth out temporal demand, facilitate a feasible shift from peak to off-peak hours, and better utilize existing train capacity, thereby reducing both initial waiting times and additional waiting times caused by stranded passengers. The results also indicate that passengers' arrival time choice behavior plays a critical role in designing realistic and effective pricing strategies. From a methodological



(a) Shifting proportion

(b) Expected value of total waiting time

Figure 1: Impact of sensitivity parameters related to fare β_P and earlier arrival at the destination β_E on the expected values of the shifting proportion from the peak to the off-peak periods and passengers' total waiting time under fixed sensitivity for peak-period disutility β_K of 10.

perspective, compared with GUROBI as well as the classic and improved Benders decomposition methods, the proposed machine learning-enhanced Benders decomposition algorithm can considerably improve computational efficiency. Due to space limitations, detailed numerical results and managerial insights will be presented at the conference.

References

AmsterdamTips. (2024, December). *How to get cheap train tickets on ns dutch railways.* Retrieved from `https://www.amsterdamtips.com/kruidvat-train-day-tickets` (accessed: 01-11-2025)

Beijing Subway. (2017, January). *Fare discount policy during off-peak hours.* Retrieved from `http://bjsubway.com/support/faq/2017-01-09/127569.html` (Accessed 01-11-2025)

Manchiraju, C., Dawande, M., Janakiraman, G., & Raghunathan, A. (2024). Dynamic pricing and capacity optimization in railways. *Manufacturing & Service Operations Management*, *26*(1), 350-369.

Nederlandse Spoorwegen . (2025, May). *Terms and conditions for NS PriceTime Deal.* Retrieved from `https://www.ns.nl/binaries/_ht_1746710900606/content/assets/ns-en/terms/terms-and-conditions-ns-pricetime-deal.pdf` (Accessed: 01-11-2025)