Morning commute in congested urban rail transit system: a macroscopic model for equilibrium distribution of passenger arrivals

Jiahua Zhang\textsuperscript{a}, Kentaro Wada\textsuperscript{b} and Takashi Oguchi\textsuperscript{a}

\textsuperscript{a}Institute of Industrial Science, The University of Tokyo, Tokyo, Japan; \textsuperscript{b}Faculty of Engineering, Information and Systems, University of Tsukuba, Tsukuba, Ibaraki, Japan

ABSTRACT
This paper proposes a macroscopic model to describe the equilibrium distribution of passenger arrivals for the morning commute problem in a congested urban rail transit system. We use a macroscopic train operation sub-model developed by Seo, Wada, and Fukuda to express the interaction between the dynamics of passengers and trains in a simplified manner while maintaining their essential physical relations. The equilibrium conditions of the proposed model are derived and a solution method is provided. The characteristics of the equilibrium are then examined both analytically and numerically. As an application of the proposed model, we analyze a simple time-dependent timetable optimization problem with equilibrium constraints and reveal that a ‘capacity increasing paradox’, in which a higher dispatch frequency increases the equilibrium cost, exists. Furthermore, insights into the design of the timetable are obtained and its influence on passengers’ equilibrium travel costs is evaluated.

1. Introduction
Urban rail transit systems, with its high capacity and punctuality, serves as a typical solution to commuters’ travel demand during rush hour in most metropolises worldwide (Vuchic 2005). However, severe congestion and unexpected delays frequently degrade the travel experience of commuting. In many metropolises, the congestion and delay of rail transit have brought about tremendous psychological stress to commuters and considerable economic loss to society. For example, according to MLIT (2020), on an average, train delays (more than 5 min) were observed for 45 railway lines in the Tokyo metropolitan area in 11.7 days of 20 weekdays in a month, and more than half of the short delays (within 10 min) were caused by extended dwell time. Kariyazaki, Hibino, and Morichi (2015) estimated that in Japan, train delays resulted in social cost in excess of 1.8 billion dollars per year.

In a high-frequency operated rail transit system during rush hour, scheduled dwell times at stations tend to be extended due to demand concentration (temporal high demand). As illustrated in Figure 1, this passenger-related delay would force the subsequent trains to decelerate or stop between stations to maintain a safety clearance, which is a so-called ‘knock-on delay’ on the rail track (Carey and Kwieciński 1994). Conversely, this on-track congestion would induce the dwell time extension: each train is required to carry more passengers due to the increase in the train headway (i.e. the reduction of the train frequency). This is a typical dynamic interaction of demand concentration, the dwell...
time extension, and on-track congestion during rush hour (Kato, Kaneko, and Soyama 2012; Tirachini, Hensher, and Rose 2013; Kariyazaki, Hibino, and Morichi 2015; Li et al. 2017). To investigate strategic-level countermeasures such as dynamic demand management strategies and timetable planning for the commuting problem in urban rail transit systems, it is essential to understand how the complex interaction determines the whole system’s dynamics.

To model the dynamic congestion during morning rush hour due to demand concentration, Vickrey (1969) proposed a departure-time choice equilibrium problem (morning commute problem) at a single bottleneck road network. Various extensions of the basic model have been proposed in transportation planning and demand management literature (see Li, Huang, and Yang 2020, for a comprehensive review). However, the models for road traffic may not be readily applicable to rail transit because the mechanisms of congestion and delay differ considerably between these two systems. Several studies have addressed the problem in public transit systems (e.g. Kraus and Yoshida 2002; Tian, Huang, and Yang 2007; de Palma, Kilani, and Proost 2015; de Palma, Lindsey, and Monchambert 2017; Yang and Tang 2018; Zhang, Lindsey, and Yang 2018). These studies analyzed travel decisions of transit users and fare optimization issues under the assumption that the in-vehicle crowding and/or oversaturated waiting time at stations is the primary congestion cost of travelling.1 However, no studies have addressed passengers’ departure time choice behaviour (demand concentration) considering the aforementioned interaction. In other words, the existing studies assume that the in-vehicle travel time is constant and no studies consider the situation that the time increases considerably because of longer dwell time at stations (Zorn, Sall, and Wu 2012) and longer running time on the track (Kariyazaki, Hibino, and Morichi 2015; Li et al. 2017). A comparison of the main features of this study with other studies is summarized in Table 1.

For the modelling the train operation, there are several detailed models that capture the vehicle dynamics and railway signalling (see, for example, Barrena et al. 2014; Robenek et al. 2016; Shi et al. 2018; Xu et al. 2019). Recently, there are also plenty of studies on the optimization of the train timetabling considering the passenger flows, short-turning strategy, or route plan of trains (see, for example, Hao, Song, and He 2022; Yuan et al. 2022; Zhu et al. 2022; Yang et al. 2022). However, the sophisticated microscopic modelling of congested transportation systems is typically complex and its solution relies heavily on a numerical analysis. Although they are suitable for case-specific optimizations and evaluations of operational-level countermeasures, it is generally difficult to obtain general or strategic policy implications for which we aim. Therefore, we turn our attention to macroscopic modelling of extracting an essential feature of such a complex system. Specifically, we focus on a macroscopic fundamental diagram (MFD) (Daganzo 2007; Geroliminis and Daganzo 2007) that has

![Figure 1.](image-url) The interaction of on-track congestion and dwell time extension.
emerged as a promising modelling methodology for congested road networks. The MFD provides simple relations among aggregate traffic variables in homogeneously congested neighbourhoods, and is regarded to be useful for obtaining general insights on demand management and control strategies (e.g. Fosgerau and Small 2013; Geroliminis, Haddad, and Ramezani 2013). The rail transit system resembles neighbourhood-size road networks such that the average density or accumulation larger than some critical values degrades the throughput of the system, which corresponds to the on-track congestion of trains. However, rail transit differs from road networks in terms of travel time because it is extended not only by the congestion of trains (vehicles), but also by the boarding and alighting of passengers. To simultaneously describe these two features of the congested rail transit, Seo, Wada, and Fukuda (2017, 2022) proposed a macroscopic and tractable train operation model (train-FD model), which enables the analysis of the morning commute problem in this study.

The purpose of this study is to develop a macroscopic model that describes the equilibrium distribution of passenger arrivals for the morning commute problem in a congested urban rail transit system. In this model, we use the train-FD model to express the interaction between the dynamics of passengers and trains in a simplified manner while maintaining their essential physical relations. We derive the equilibrium conditions of the proposed model and provide a solution method. The characteristics of the equilibrium are then examined both analytically and numerically. Finally, by applying the proposed model, we analyze a simple time-dependent timetable optimization problem with equilibrium constraints and reveal that a ‘capacity increasing paradox’, in which a higher dispatch frequency increases the equilibrium cost, exists. Furthermore, the design of timetables and their influence on passengers’ equilibrium travel costs are investigated. Owing to its simplicity and comprehensiveness, the proposed model can be a useful tool to evaluate management strategies of congested rail transit systems from both the demand and supply sides.

| Table 1. Comparison of typical studies on rail transit modelling with departure-time choice. |
| Publications | Main objectives | In-vehicle travel time | On-track congestion | In-vehicle crowding | Over-saturated waiting time | Other concern |
| Kraus and Yoshida (2002) | Optimal pricing, capacity and service frequency | constant | ✓ | ✓ | ✓ | Operation cost |
| Tian, Huang, and Yang (2007) | Equilibrium property in many-to-one system | constant | ✓ | ✓ | ✓ | - |
| Li et al. (2010) | Activity-based transit assignment model for timetabling problem | constant | ✓ | ✓ | ✓ | - |
| de Palma, Kilani, and Proost (2015) | Discomfort implication for scheduling and pricing | constant | ✓ | ✓ | ✓ | Fare |
| de Palma, Lindsey, and Monchambert (2017) | Equilibrium and optimum of crowding in rail transit | constant | ✓ | ✓ | ✓ | Fare |
| Yang and Tang (2018) | Fare-reward scheme design | constant | ✓ | ✓ | ✓ | Fare reward |
| Zhang, Lindsey, and Yang (2018) | Frequency and fares with heterogeneous users | constant | ✓ | ✓ | ✓ | Operator profit |
| Zhang et al. (2020) | Modelling and optimizing congested rail transit with heterogeneous users | constant | ✓ | ✓ | ✓ | Fare |
| Tang et al. (2020) | Fare scheme design with heterogeneous users | constant | ✓ | ✓ | ✓ | Fare |
| This study | Macroscopic modelling of equilibrium and timetable optimization with equilibrium constraint | time-dependent | ✓ | ✓ | ✓ | - |

*This study only considers the undersaturated and near-saturated condition, thus the waiting time (or travel delay) is included in the travel time.*
Note that the purposes and contributions of Seo, Wada, and Fukuda (2017, 2022) and this study are fundamentally different. The focus of Seo, Wada, and Fukuda (2017, 2022) was to model the train-passage interactions as the train-FD (see also Section 2.1) and apply the model to describe the dynamics of the train for an exogenously given (time-dependent) passenger demand. In contrast, the focus of this study is to combine the passenger demand (departure-time choice) model with the train-FD. In the combined model, not only the dynamics of the train but also the dynamics of the passenger arrival for the railway system are simultaneously and endogenously determined as an equilibrium state of the whole system.

The remainder of this paper is organized as follows. Section 2 introduces the model for the morning commute problem in rail transit. Section 3 derives the user equilibrium and provides a solution method for the proposed model. Section 4 describes the characteristics of the equilibrium through analytical discussion and numerical examples. Section 5 applies the proposed model to a simple time-dependent timetable optimization problem. Finally, conclusions and directions for future studies are discussed in Section 6.

2. Macroscopic model for morning commute problem in rail transit

In this section, we formulate a model for the morning commute problem in rail transit. In Section 2.1, we present an overview of the macroscopic train operation model proposed by Seo, Wada, and Fukuda (2017, 2022), which is a supply side sub-model of the proposed model. In Section 2.2, we describe behavioural assumptions of users’ departure time choice, which is a demand side sub-model.

2.1. Macroscopic rail transit operation model

Consider a railway system on a single-line track, where stations are homogeneously located along the line. All trains stop at every station and the first-in-first-out (FIFO) service is satisfied along the railway track. In the following, we first present the microscopic operation assumptions on passenger boarding and train cruising to obtain a macroscopic model.

Passenger boarding behaviour is described using a queuing model (Wada et al. 2012). Specifically, the train dwell time \( t_b \) at each station is expressed as follows:

\[
t_b = t_{b0} + \frac{a_p h}{\mu},
\]

where \( t_{b0} \) is the buffer time, including the time required for door opening and closing, \( \mu \) is the maximum flow rate of passenger boarding, \( a_p \) is the passengers’ arrival rate at the platform, and \( h \) is the headway of two succeeding trains. We assume that passengers can always board the next train here.

The cruising behaviour of trains is assumed to be described by Newell’s simplified car-following model (Newell 2002). In this model, a vehicle either travels at its desired speed or follows the preceding vehicle while maintaining safety clearance. Specifically, the position of train \( n \) at time \( t \) is described as follows:

\[
x_n(t) = \min\{x_{n-1}(t - \tau) + v_f \tau, x_{n-1}(t - \tau) - \delta\},
\]

where \( n-1 \) refers to the preceding train of train \( n \), \( \tau \) is the reaction time of the train, and \( \delta \) is the minimum spacing. The first term represents the free-flow regime, where the train cruises at its desired speed \( v_f \). The second term represents the congested regime, where the train decreases its speed to maintain minimum spacing.

A train fundamental diagram (train-FD), \( q = Q(k, a_p) \) describes the steady-state relation among train flow \( q \) (\( q = 1/h \)), train density \( k \), and passenger arrival rate \( a_p \) in a homogeneously congested transit system. Specifically, by formulating traffic variables based on their (generalized) definition and the operating principles (1) and (2) and by relating them, the train-FD can be analytically expressed as
follows (see Seo, Wada, and Fukuda 2017, 2022, for a derivation):

\[
Q(k, a_p) = \begin{cases} 
 kl - \frac{a_p}{\mu} & \text{if } k < k^*(a_p) \\
 \left(1 - \delta/\nu_f + \tau\right) l - \frac{a_p}{\mu} \frac{t_{b0} + l/\nu_f}{l} (k - k^*(a_p)) + q^*(a_p) & \text{if } k \geq k^*(a_p)
\end{cases},
\]

where \(l\) is the (average) distance between adjacent stations, and \(q^*(a_p)\) and \(k^*(a_p)\) are the critical train flow and train density, respectively:

\[
q^*(a_p) = \frac{1 - a_p/\mu}{t_{b0} + \delta/\nu_f + \tau},
\]

\[
k^*(a_p) = \frac{(1 - a_p/\mu) (t_{b0} + l/\nu_f)}{(t_{b0} + \delta/\nu_f + \tau) l} + \frac{a_p}{\mu l}.
\]

The train-FD was inspired by the MFD for road networks, and they are similar as follows. First, they both describe the traffic states in a homogeneously congested area using system-wide aggregate variables. Second, they both exhibit unimodal relations between the density (accumulation) and flow (throughput) of the system, which yields two regimes: the free-flow and congested regimes. An essential difference between the train-FD and MFD is that the train-FD has an additional dimension of passenger flow. Introducing this new dimension enables the simplified modelling of rail transit operations in which passenger concentration is considered.

To describe the rail transit system behaviour when the demand (i.e. passenger flow) and supply (i.e. train density) change dynamically, we consider the system as an input-output system with the train-FD, as illustrated in Figure 2. Two types of inputs exist: train inflow and passenger inflow (i.e. passenger arrival rate \(a_p\)). Accordingly, two outputs are considered: train outflow and passenger outflow. Within the system, the trains operate based on the rules in Equations (1) and (2), whereas passengers arrive (inflow) at the system based on their assessment of the travel cost introduced in the next subsection. As in existing MFD applications for morning commute problems (e.g. Geroliminis and Levinson 2009; Geroliminis, Haddad, and Ramezani 2013; Fosgerau 2015), it is expected that this simplified model can provide insight into the time-dependent characteristics of the rail transit system despite its inability to capture spatial dynamics or heterogeneity within the system.

2.2. Passenger travel cost

Consider a fixed number \(N_p\) of passengers that use the rail transit system during the morning rush hour. The length of their trip in the system is common for all passengers and is denoted by \(L\). Passengers
choose their departure time from home to minimize their travel costs. The travel time from home to the nearest station for any passenger is assumed to be constant; thus, without a loss of generality, it is set to zero. We also assume that the departure time from the system is the arrival time at the destination (i.e. the workplace). For clarity, if not particularly indicated, we refer to ‘passenger/train departure’ as the departure (or exit) from the rail transit system.

The travel cost (TC) is assumed to consist of the travel delay cost (TDC) in the rail transit system and schedule delay cost (SDC). Specifically, the TC of a passenger i departing from the system at time t is defined as follows:

\[ TC(t, t^*_i) = \alpha (T(t) - T_0) + s(t, t^*_i), \]

where \( t^*_i \) is the desired departure time (or work start time), \( \alpha \) is the time value for a travel delay, \( T(t) \) is the travel time for a passenger departing from the rail transit system at time \( t \), \( T_0 = \frac{1}{2} (t_{b0} + l/v_f) \) is the minimum travel time before the morning rush starts, and \( s(t, t^*_i) \) is the schedule delay cost. Here, we employ the following piecewise linear schedule delay cost function \( s(t, t^*_i) \) that has been widely used in previous studies (e.g. Hendrickson and Kocur 1981; Tian, Huang, and Yang 2007; Geroliminis and Levinson 2009; de Palma, Lindsey, and Monchambert 2017; Yang and Tang 2018).

\[
 s(t, t^*_i) = \begin{cases} 
 \beta (t_i^* - t) & \text{if } t < t_i^* \\
 \gamma (t - t_i^*) & \text{if } t \geq t_i^* 
\end{cases}
\]

where \( \beta \) and \( \gamma \) are the values of time for earliness and lateness, respectively. For simplicity, we assume that all passengers have the same cost parameters \( \alpha \), \( \beta \) and \( \gamma \). We specify the desired departure time distribution in a later section.

The travel time for a passenger departing from the rail transit system at time \( t \) is equal to that of a train departing from the system at the same time. Let \( n \) be train number\(^6 \) departing from the system at time \( t \), and \( T(n) = T(t) \) be its travel time. The service (or average travelling) speed of train \( n \) is \( L/T(n) \).

We denote the headway of train \( n \) when arriving at the system by \( h_a(n) \) and that when departing from the system by \( h_d(n) \). If we approximate the time-space trajectories of the trains as straight lines whose slopes are their service speeds (see Figure 3), the average spacing of train \( n \), \( \overline{s}(n) \), can be defined as follows:

\[
 \overline{s}(n) \equiv \frac{L}{T(n)} \overline{h}(n) \quad \text{where} \quad \overline{h}(n) = \frac{h_a(n) + h_d(n)}{2}.
\]

This is consistent with the generalized definition of traffic variables (Edie 1963).

Here, we introduce the main assumption in this study: the (average) train flow \( q(n) = 1/\overline{h}(n) \) and (average) train density \( k(n) = 1/\overline{s}(n) \) of the system with respect to train \( n \) satisfy the train-FD, that is,

\[
 q(n) = Q(k(n), a_p(n)) \Leftrightarrow \frac{1}{\overline{h}(n)} = Q \left( \frac{1}{\overline{s}(n)}, a_p(n) \right)
\]

where \( a_p(n) \) is the average passenger arrival rate for train \( n \) at the stations along the line. If the system is in a steady state, the relation in (9) must hold. Therefore, the relation in (9) approximately holds if the (average) values of the state variables vary gradually.

This assumption enables us to link the time-dependent (more precisely, train-dependent\(^7 \)) passenger demand \( \{a_p(n)\} \) to travel time \( \{T(n)\} \) in a simplified manner while maintaining their essential physical relationships. Specifically, as illustrated in the next section, the train traffic state variables are determined by the departure-time choice equilibrium conditions first, and the equilibrium passenger arrival rates \( \{a_p(n)\} \) can then be estimated using the relation in (9). Note that Equation (9) does not represent macroscopic train dynamics. The train dynamics using the train-FD (i.e. an exit-function model) can be found in Seo, Wada, and Fukuda (2017, 2022).
3. User equilibrium

Under the setting described in the previous section, the user equilibrium is defined as the state in which no transit user can reduce his/her travel cost by changing his/her departure time from the system unilaterally. In this section, we first derive the equilibrium conditions. Next, we present a solution method to the proposed model.

3.1. Equilibrium conditions

Since each passenger $i$ selects his/her departure time $t_i$ from the system to minimize the travel cost at equilibrium, the following condition is satisfied at time $t = t_i$:

$$\frac{\partial TC(t_i, t^*_i)}{\partial t} = \alpha \frac{dT(t_i)}{dt} + \frac{\partial s(t_i, t^*_i)}{\partial t} = 0.$$  \hspace{1cm} (10)

The derivative of the travel time $T(t)$ is obtained by substituting Equations (6) and (7) into Equation (10) as follows:

$$\frac{dT(t_i)}{dt} = \begin{cases} \frac{\beta}{\alpha} & \text{if } t_i < t^*_i \\ -\frac{\gamma}{\alpha} & \text{if } t_i \geq t^*_i \end{cases}.$$  \hspace{1cm} (11)

Furthermore, with the first-in-first-work assumption (Daganzo 1985), the travel time $T(t)$ is maximized when the schedule delay is zero (we refer to this time as $t_m$). Consequently, the travel time $T(t)$ under equilibrium is expressed as follows:

$$T(t) = \begin{cases} T_0 + \frac{\beta}{\alpha} (t - t_0) = T_e(t) & \text{if } t_0 \leq t < t_m \\ T_0 + \frac{\beta}{\alpha} (t_m - t_0) - \frac{\gamma}{\alpha} (t - t_m) = T_l(t) & \text{if } t_m \leq t \leq t_{ed} \end{cases},$$  \hspace{1cm} (12)

where $t_0$ and $t_{ed}$ represent the start and end of the morning rush hour, respectively.

As mentioned in the previous section, the equilibrium passenger arrivals are estimated using the train traffic state variables (i.e. $T(n), h(n), s(n)$). Among them, we have already specified the travel time under the equilibrium $T(n) = T(D^{-1}(n))$ as Equation (12). The remaining variables, the average headway and spacing, are derived as follows. Let $A(t)$ be the cumulative number of train arrivals at the
system at time $t$. Then, $A(\cdot)$ and the cumulative departures $D(\cdot)$ is related through the FIFO condition, $D(t) = A(t - T(t))$, or its derivative form:

$$d(t) = a(t - T(t)) \left(1 - \frac{dT(t)}{dt}\right),$$

(13)

where $a(t)$ and $d(t)$ are the inflow and outflow of the trains, respectively. Since $A(t)$ is the given information (i.e. timetable), $D(t)$ can be obtained from this FIFO condition. Also, from the definition, the arrival and departure headways $h_a(n)$ and $h_d(n)$ of train $n$ are obtained as follows:

$$h_a(n) = \frac{d(t - T(t))}{dn} = \frac{1}{a(t - T(t))}, \quad h_d(n) = \frac{dt}{dn} = \frac{1}{d(t)}.$$

(14)

We finally calculate the average headway $\bar{h}(n)$ and spacing $\bar{s}(n)$ from these variables.

Now, we can estimate the passenger arrivals under equilibrium. For a given train density, the train-FD provides a one-to-one correspondence between the train and passenger flows, that is, $q = \hat{Q}(\alpha_p | k) = Q(\alpha_p, k)$. Therefore, from our main assumption (9), we have the following expression:

$$a_p(n) = \hat{Q}^{-1} \left(\frac{1}{\bar{h}(n)} \mid \frac{1}{\bar{s}(n)}\right),$$

(15)

where we use the following inverse function $a_p = \hat{Q}^{-1} (q | k)$.

To obtain a complete equilibrium solution (i.e. to determine $t_0$, $t_m$ and $t_{ed}$), the desired departure time distribution should be specified. In this study, we consider two types of distributions: Cases WT1 and WT2. For Case WT1, a fixed number $N_p$ of passengers has a common desired departure time $t^\ast$. For Case WT2, the cumulative number of passengers who want to depart by time $t$ is expressed by using a Z-shaped function, $W_p(t)$, with $N_p$ passengers and a positive constant slope (i.e. demand rate) (e.g. Gonzales and Daganzo 2012). An illustration of the cumulative curves of passengers for these two cases is shown in Figure 4.

For Case WT1, the first condition is $t_m = t^\ast$. The second condition is that the last user experiences the schedule delay cost only, i.e.

$$T(t_{ed}) = T_0.$$
Algorithm 1 Solution for Case WT1

Require: Operational parameters, $l, L, t_{B0}, \mu, v_f, \delta, \tau$; cost parameters, $\alpha, \beta, \gamma, t^s$; train inflow, $a(t)$, and total travel demand, $N_p$.

Ensure: Train flow $q(n)$, train density $k(n)$, and passenger arrival rate $a_p(n)$.

1: Set an initial $t_0$.
2: Calculate $T(t)$ and $t_{ed}$ by Equations (12) and (16).
3: Calculate $d(t)$ using Equation (13).
4: Calculate $a_p(n)$ by Equation (15), together with Equations (8) and (14).
5: Subtract $N_p$ from the left-hand side of the discrete version of Equation (17) (with unit $\Delta n$), denoted as an error.
6: if error $< -\epsilon_p$, then
7: \hspace{1cm} $t_0 = t_0 - \Delta t$, repeat lines 2–5.
8: else if error $> \epsilon_p$, then
9: \hspace{1cm} $t_0 = t_0 + \Delta t$, repeat lines 2–5.
10: else
11: \hspace{1cm} Calculation converges, $t_0$ and $t_{ed}$ are determined.
12: end if
13: Outputs are obtained from line 4 when the calculation converges.

The last condition is the conservation of the number of users:

$$D_p(t_{ed}) = \int_{D(t_0)}^{D(t_{ed})} a_p(n)\bar{n}(n)dn = N_p$$  \hspace{1cm} (17)

where $D_p(t)$ is the cumulative number of passengers departing from the system at time $t$, and $D_p(t_0) = 0$. By solving the last two conditions simultaneously, $t_0$ and $t_{ed}$ are determined.

For Case WT2, we assume that a unique time instant $t_m$ exists when the schedule delay becomes zero, as in the standard morning commute problem for road traffic (Smith 1984; Daganzo 1985). We then have the following expression:

$$D_p(t_m) = W_p(t_m).$$  \hspace{1cm} (18)

By solving the three conditions (16), (17) and (18) simultaneously, $t_0$, $t_m$ and $t_{ed}$ are determined.

A solution method for Case WT1 is presented in Algorithm 1, where $\Delta t$ is the step size of time, $\Delta n$ is the discrete unit of the train, and $\epsilon_p$ is the tolerance of error in the number of passengers. The solution method for Case WT2 is similar to Algorithm 1: another step is simply added to calculate $t_m$ that satisfies Equation (18) after line 1. Note that an equilibrium solution may not exist (i.e. the solution method can produce a physically infeasible result). This problem is addressed in the next section.

4. Characteristics of equilibrium

4.1. Analytical discussion

In this subsection, the characteristics of equilibrium are examined analytically. Specifically, we derive the analytical solution of the proposed model for some train operation patterns and discuss the equilibrium flow and cost structure. For clarity, we only consider Case WT1 hereinafter.
First, by using Equations (11)–(14), the average flow $q(n)$ and average density $k(n)$ for each train $n$ $(t = D^{-1}(n))$ can be expressed as follows:

$$q(n) = \frac{1}{h(n)} = \begin{cases} \frac{\zeta_1 a(t - T(t))}{\bar{T}(n)} & \text{if } t_0 \leq t < t^* \\ \frac{\zeta_2 a(t - T(t))}{\bar{T}(n)} & \text{if } t^* \leq t \leq t_{ed} \end{cases} \quad (19)$$

$$k(n) = \frac{1}{\bar{s}(n)} = \frac{T(n)}{\bar{T}(n)L} = \begin{cases} \frac{\zeta_1 a(t - T(t))}{\bar{T}(n)} \frac{T(t)}{L} & \text{if } t_0 \leq t < t^* \\ \frac{\zeta_2 a(t - T(t))}{\bar{T}(n)} \frac{T(t)}{L} & \text{if } t^* \leq t \leq t_{ed} \end{cases} \quad (20)$$

where

$$\zeta_1 \equiv \frac{2(\alpha - \beta)}{2\alpha - \beta} < 1, \quad \zeta_2 \equiv \frac{2(\alpha + \gamma)}{2\alpha + \gamma} > 1.$$

It can be understood from Equation (19) that the average flow $q(n)$ under equilibrium is determined only by the train inflow $a(t - T(t))$ and two time-value parameters, $\zeta_1$ and $\zeta_2$. For a constant train inflow $a(t) = a_c$, the constant average flow smaller than the inflow is realized before the desired departure time $t^*$ from the system, and the constant average flow larger than the inflow is realized after $t^*$. The average density $k(n)$ under equilibrium is the product of the time value parameter, train inflow, and piece-wise linear travel time. Thus, for a constant train inflow, the density linearly increases until $t^*$ and subsequently linearly decreases until the end of the equilibrium period $t_{ed}$.

The passenger arrival rate $a_p(n)$ for train $n$ can also be expressed explicitly using Equation (15) and the aforementioned flow and density. If the train state $(q(n), k(n))$ is in the free-flow regime, $a_p(n)$ can be written as follows:

$$a_p(n) = \begin{cases} \frac{\mu L(T(t) - T_0)}{\bar{L}} a(t - T(t)) & \text{if } t_0 \leq t < t^* \\ \frac{\mu L(T(t) - T_0)}{\bar{L}} a(t - T(t)) & \text{if } t^* \leq t \leq t_{ed} \end{cases} \quad (21)$$

If the train state $(q(n), k(n))$ is in the congested regime, $a_p(n)$ can be written as:

$$a_p(n) = \begin{cases} \frac{\mu L}{1 - \delta} \left( 1 - \frac{\zeta_1 a(t - T(t))}{\bar{T}(n)} \frac{T(t)}{L} + \eta \right) & \text{if } t_0 \leq t < t^* \\ \frac{\mu L}{1 - \delta} \left( 1 - \frac{\zeta_2 a(t - T(t))}{\bar{T}(n)} \frac{T(t)}{L} + \eta \right) & \text{if } t^* \leq t \leq t_{ed} \end{cases} \quad (22)$$

where $\eta \equiv (l - \delta) t_{b0} + \tau l$.

Before completing the derivation of the analytical solution, let us show three feasibility conditions to ensure the existence of the equilibrium. The first condition ensures that the average flow $q(n)$ is positive, that is,

$$\zeta_1 > 0 \iff \alpha > \beta. \quad (23)$$

This condition is consistent with the sufficient condition for the existence of equilibria in the equilibrium models for road traffic (Smith 1984). The second condition ensures that the average flow $q(n)$ does not exceed the maximum train flow $q^*(0)$ in the train-FD (i.e. Equation (4)), which can be described as:

$$a(t - T(t)) \leq q^*(0)/\zeta_2. \quad (24)$$

This condition restricts the train inflow (or dispatch headway) that can be adopted by the operator. The third condition ensures that the passenger arrival rate $a_p(n)$ is non-negative. For a constant
train inflow, it can be observed from Equations (21) and (22) that the condition is equivalent to \( 1 - \zeta_2 a_c \left( \frac{\delta T(t^*)}{L} + \frac{\eta}{L} \right) \geq 0 \), that is,

\[
TC^e = \alpha (T(t^*) - T_0) \leq \alpha \left[ \frac{L}{\delta} \left( \frac{1}{\zeta_2 a_c} - \frac{\eta}{L} \right) - T_0 \right] = \overline{TC},
\]

where \( TC^e \) represents the equilibrium cost and \( \overline{TC} \) represents the upper limit of equilibrium cost to guarantee the feasibility of \( a_p(n) \). Since this condition depends on the equilibrium dynamics, we will discuss it in detail below.

Now, let us complete the solution by deriving an equilibrium cost analytically using Equations (19)–(22). We consider the train inflow as a given constant \( a(t) = a_c \) hereinafter. Since the train-FD has two regimes, the pattern of the railway dynamics under equilibrium can be classified by the combination and order of the two regimes. Specifically, all the possible patterns are FF, FCF, and FCCF, where F represents the free-flow regime and C represents the congested regime. The sequences of F and C indicate the order of occurrence of these two regimes. These three patterns are illustrated in Figure 5. All patterns form a counter-clockwise loop of the train state \((q(n), k(n))\) on the train-FD, as we can see from the above-mentioned properties of \( q(n) \) and \( k(n) \). Note that the essential difference between these three patterns lies in whether the upper-right and bottom-right corners of the loops, which correspond to time \( t^* \), are in the congested regime of the train-FD or not. In the following, we will discuss each of the patterns.

### 4.1.1. Pattern FF

In Pattern FF, the train state \((q(n), k(n))\) is in the free-flow regime of train-FD during the whole rush hour. In this case, according to Equation (21), the passenger arrival rate \( a_p(n) \) first linearly increases from 0 and maximizes at \( t^* \); subsequently it linearly decreases to 0 at the end of the rush hour.

The equilibrium cost can be derived analytically as follows. First, the conservation law in Equation (17) is rewritten by using Equations (11), (13), (14), and \( n = D(t) \):

\[
N_p = \int_{t_0}^{t_{ed}} a_p(n) \left( \frac{h_p(n) + h_d(n)}{2} \right) d(t) dt
= \left( 1 - \frac{\beta}{2\alpha} \right) \int_{t_0}^{t^*} a_p(n) dt + \left( 1 + \frac{\gamma}{2\alpha} \right) \int_{t^*}^{t_{ed}} a_p(n) dt
\]

(26)
Then, by substituting Equation (21) into Equation (26), the equilibrium cost $TCE$ is obtained.

$$
TCE = \sqrt{\frac{2\alpha LN_p}{\mu la_c \left( \frac{\beta}{\beta} + \frac{1}{\gamma} \right)}}. \quad (27)
$$

From this analytical solution, we see that the equilibrium cost increases with an increase in the demand $N_p$, time-value parameters $\alpha$, $\beta$, and $\gamma$; it decreases with the increase in the maximum flow rate of passenger boarding $\mu$ and dispatch frequency $ac$.

The occurrence condition for Pattern FF is given as the one in that $k(n)$ at time $t^*$ (the upper-right and bottom-right corners) is less than the critical density $k^* (ap(n))$. Specifically, this condition can be expressed as follows:

$$
\begin{align*}
\begin{cases}
\xi_1 ac \frac{TCE}{\alpha + T_0} L \leq \frac{1}{I} \left[ 1 + \xi_1 ac \left( \frac{l - \delta}{V_f} - \tau \right) \right] \\
\xi_2 ac \frac{TCE}{\alpha + T_0} L \leq \frac{1}{I} \left[ 1 + \xi_2 ac \left( \frac{l - \delta}{V_f} - \tau \right) \right].
\end{cases}
\end{align*} \quad (28)
$$

Since $\xi_2 > 1 > \xi_1$, the first line in Equation (28) is satisfied as long as the second line is satisfied. Thus, we have

$$
TCE \leq \frac{\alpha L \left[ 1 + \xi_2 ac \left( \frac{l - \delta}{V_f} - \tau \right) \right]}{\xi_2 ac} - \alpha T_0 = TC_{FF}. \quad (29)
$$

We denote the unique and maximum passenger demand $N_p$, satisfying Equation (29) as $N^{FF}_p$. Note that the feasibility condition (25) is satisfied because $TC_{FF} < \overline{TC}$.

### 4.1.2. Pattern FCF

When $N_p$ exceeds $N^{FF}_p$, the second line in Equation (28) is first violated, whereas the first line still holds. It means that all the trains exiting the railway system before $t^*$ operate in the free-flow regime, whereas a portion of trains exiting after $t^*$ are forced to operate in the congested regime. This operation pattern is Pattern FCF. In this pattern, the passenger arrival rate $ap(n)$ may exhibit one or two peaks during the rush hour depending on whether the value of the first line in Equation (21) is smaller than that of the second line in Equation (22) at time $t^*$. For the one-peak case, the peak occurs when the train state switches from the congested regime to the free-flow regime at time $t > t^*$. For the two-peak case, another peak occurs at time $t^*$ (see also numerical examples in Section 4.2).

By substituting the second line in Equation (22) together with the first line in Equation (21) into the conservation law (26), the analytical solution is obtained:

$$
TCE = \beta \left( \frac{-R + \sqrt{R^2 - 4U(S - N_p)}}{2U} \right), \quad (30)
$$

where

$$
\begin{align*}
R &= \frac{\mu \beta}{(l - \delta) \gamma} \left( 1 + \frac{\gamma}{2\alpha} \right) \left[ l - \eta \xi_2 ac - \delta \xi_2 ac \right] T_0 \\
S &= \mu \left[ 1 + \frac{\gamma}{2\alpha} \right] \frac{TC_{FF}}{\gamma} \left[ \frac{2\alpha L \xi_2 ac TC_{FF} \left( \frac{l - \delta}{l - \delta} \right) + \frac{\delta}{l - \delta} \xi_2 \frac{1}{2\alpha L} ac (2\alpha T_0 + TC_{FF})}{\eta \xi_2 ac} \right] \\
U &= \frac{\mu l}{2L} \xi_2 ac \left[ \frac{\beta}{\alpha} \left( 1 - \frac{\beta}{\alpha} \right) - \frac{\delta \beta^2}{(l - \delta) \alpha \gamma} \left( 1 + \frac{\gamma}{\alpha} \right) \right].
\end{align*}
$$
The sensitivity of $T_C^e$ to $N_p$ is also obtained as follows:

$$\frac{\partial T_C^e}{\partial N_p} = \frac{\beta}{\sqrt{R^2 + 4UN_p - 4US}}.$$  

(31)

The equilibrium cost $T_C^e$ monotonically increases with $N_p$ as long as the setting of parameters ensures that $R^2 + 4UN_p - 4US \geq 0$; its sensitivity to $N_p$ decreases with an increase in $N_p$ when $U > 0$ and increases with an increase in $N_p$ when $U < 0$. However, the sensitivities of $T_C^e$ to time value parameters $\alpha$, $\beta$, and $\gamma$ are tedious to derive from Equation (30). Therefore, we examine this issue numerically in the next subsection.

As in Pattern FF, from Equation (28) and feasibility condition (25), the occurrence condition for Pattern FCF is given as follows:

$$T_C^FF < T_C^e \leq \min \left\{ T_C^e, T_C^{FCF} \right\},$$  

(32)

where

$$T_C^{FCF} = \frac{\alpha L}{\zeta_1 l a_c} \left[ 1 + \zeta_1 a_c \left( \frac{l - \delta}{v_f} - \tau \right) \right] - \alpha T_0.$$

We denote the maximum demand $N_p$ satisfying Equation (32) as $N_p^{FCF}$. Note that, if $T_C^e < T_C^{FCF}$, the demand larger than $N_p^{FCF}$ cannot be accommodated at equilibrium because the feasibility condition is violated; otherwise, a new operation pattern FCCF appears which will be discussed in the next section. We can find that $T_C^e < T_C^{FCF}$ holds if

$$\alpha L \left[ \left( \frac{1}{\delta \zeta_2} - \frac{1}{l \zeta_1} \right) \frac{1}{a_c} - \left( \frac{1}{\delta} - \frac{1}{l} \right) (t_{b0} + \delta/v_f + \tau) \right] < 0.$$  

(33)

Since $\zeta_2 > 1 > \zeta_1 > 0$ and generally $l > \delta$, the condition implies that

$$\frac{1}{\delta \zeta_2} - \frac{1}{l \zeta_1} < \frac{1}{\delta} - \frac{1}{l}, \quad \frac{1}{a_c} > t_{b0} + \delta/v_f + \tau = \frac{1}{q^*(0)}.$$  

(34)

Thus, for a given rail transit system (i.e. $l$, $\delta$, $t_{b0}$, $v_f$, and $\tau$ are given), the railway system is more likely to reach the limit of the acceptable demand before realizing Pattern FCCF when the time value parameters $\zeta_1$ and $\zeta_2$ are far from 1 and/or the dispatch headway is close to the minimum headway.

4.1.3. Pattern FCCF

When the demand $N_p$ further exceeds $N_p^{FCF}$ and $T_C^e > T_C^{FCF}$, Pattern FCCF appears. In this pattern, one portion of trains exiting the railway system before $t^*$ already operates in the congested regime. Immediately after $t^*$, trains still operate in the congested regime because both flow and density suddenly increase. Therefore, a sudden decrease in $a_p(n)$ occurs, as can be understood from Equation (22). Also, two peaks of $a_p(n)$ happen for trains exiting before and after $t^*$. After $t^*$, the density $k(n)$ continuously decreases and the operation returns to the free-flow regime again in the end. We skip the tedious derivation of the analytical solution for this pattern. Note that the occurrence condition for Pattern FCCF is as follows:

$$T_C^{FCF} < T_C^e \leq T_C^e.$$  

(35)

4.2. Numerical examples

In this subsection, we confirm the characteristics of the equilibrium discussed in the previous subsection through several numerical examples. The basic parameter settings are listed in Table 2. For
**Table 2.** Parameter settings for numerical example.

| Parameter | Value   | Parameter | Value   |
|-----------|---------|-----------|---------|
| $l$       | 1.2 km  | $\alpha$ | 20 $/h |
| $L$       | 18 km   | $\beta$  | 8 $/h  |
| $v_f$     | 40 km/h | $\gamma$ | 25 $/h |
| $t_{00}$  | 20 sec  | $t^*$    | 240 min|
| $\mu$     | 36000 pax/h | $a(t)$ | 12 tr/h |
| $\delta$  | 0.4 km  | $W_p$    | 30000 pax/h|
| $\tau$    | 1.0 min | $N_p$    | 30,000 pax|
| $\Delta t$| 1.0 min | $\epsilon_p$ | 100 pax |
| $\Delta n$| 1 tr    |          |         |

Figure 6. Travel cost for two cases. (a) Case WT1. (b) Case WT2 with $t^* < t_m$.

Simplicity, the train inflow $a(t)$ was set as a constant. For Case WT1, the common desired departure time was set to 240 min, whereas for Case WT2, the slope of the Z-shaped function was set to $w_p = 30000$ pax/h, and the time period for the increase in $W_p(t)$ was [210, 270] min.

We first present the costs for Cases WT1 and WT2 in Figure 6, which reveals that the cost pattern is the same as the standard morning commute problem for road traffic with a piecewise linear schedule delay cost function. The dynamics of rail transit and passengers for Case WT1 are shown in Figure 7 (those for Case WT2 are almost the same). Figure 7(a) shows the cumulative arrival and departure curves of the trains and Figure 7(b) shows those of the passengers. We see that $D(t)$ first deviates from $A(t)$ during $[t_0, t^*]$ and again approaches $A(t)$ during $[t^*, t_{ed}]$. This rail transit system behaviour leads to an equilibrium in the travel cost.

As discussed in the previous subsection, depending on the total travel demand $N_p$, the passenger arrival rate may have one or two peaks in Pattern FCF. This phenomenon can be confirmed from the time evolution of the passenger arrival rate (i.e. the slope $dA_p(t)/dt$ in Figure 7(b)). A larger peak occurs just before $t^*$, and the other peak occurs near the end of rush hour. It can also be seen from the evolution of $k(n)$ and $q(n)$ on the train-FD in Figure 8. The black line exhibits that the evolution of $(k(n), q(n))$ for the demand $N_p = 30,000$ starts from the left boundary of the train-FD and moves along a counter-clockwise closed loop during rush hour. The dotted line indicates the sudden change in traffic states because of the discontinuity of travel time derivatives at $t_0$, $t^*$, and $t_{ed}$. The lower part of the loop ($q(n) < 12$ tr/h) represents the dynamics of trains departing from the system during $[t_0, t^*]$, whereas the upper part represents the dynamics during $[t^*, t_{ed}]$. The maximum of $a_p$ is reached at the lower-right corner of the loop, whereas the other peak occurs at the critical density in the upper part of the loop. This two-peak phenomenon is a new finding of the equilibrium distribution of passenger arrivals for a congested rail transit system, which should be verified through empirical observations.
Figure 7. Dynamics of the rail transit system. (a) Cumulative number of trains. (b) Cumulative number of passengers.

Figure 8. Dynamics of density and flow on train-FD.

Note that, when the travel demand is rather low and all trains operate in the free-flow regime (Pattern FF), such a phenomenon does not occur. The magenta line in Figure 8 depicts the evolution of \((k(n), q(n))\) for this low-demand case.

Next, we numerically evaluate the sensitivities of the equilibrium cost \(T_{C}^{e}\) to the time value parameters for Pattern FCF. Given the basic setting of time values in Table 2 and generally \(\gamma > \alpha > \beta\), we set the test ranges of time values as \(\alpha \in [9, 24] \$/h\), \(\beta \in [2, 18] \$/h\), and \(\gamma \in [21, 40] \$/h\). When a time parameter is tested, the values of other parameters are set as ones listed in Table 2. Additionally, the parameter settings leading to \(R^2 - 4U(S - N_p) < 0\) are eliminated. The results are shown in Figure 9(a). Although \(T_{C}^{e}\) increases with the increase in both \(\beta\) and \(\gamma\), it is more sensitive to the increase in \(\beta\), especially when \(\alpha - \beta\) approaches 0. Besides, \(T_{C}^{e}\) first decreases with an increase of \(\alpha\) but subsequently slowly increases with an increase in \(\alpha\).

Additionally, we show the relationship between \(N_p\) and \(T_{C}^{e}\) under the two train inflow settings in Figure 9(b). From this figure, we see that that \(T_{C}^{e}\) monotonically increases with an increase in \(N_p\) in both Pattern FF and FCF. When trains are dispatched with a higher frequency, on-track congestion easily occurs and thus, Pattern FCF starts from a smaller \(N_p\). When \(N_p\) is small (e.g. \(N_p < 1 \times 10^4\)), adopting a higher dispatch frequency can reduce the equilibrium cost even if some of the trains operate in
the congested regime. However, when \( N_p \) is sufficiently large (e.g. \( N_p > 2 \times 10^4 \)), a higher dispatch frequency conversely leads to a higher equilibrium cost. This result implies that adopting a timetable that dispatches trains as frequently as possible may not always be appropriate from the perspective of reducing the equilibrium cost when passenger demand is high.

Finally, we numerically compare the equilibrium characteristics between the two types of work start time distribution patterns (i.e. Cases WT1 and WT2 illustrated in Figure 4). As explained at the end of Section 3.1, the difference between WT1 and WT2 when solving the equilibrium solely lies in whether the time instant \( t_m \) when the schedule delay is zero is pre-given or calculated from the demand curve \( W_p(t) \). Therefore, the characteristics of train state \((q(n), k(n))\) and passenger arrival rates \( a_n(n) \) under equilibrium should remain the same between the two patterns except the \( t_m \) differs according to the setting of \( W_p(t) \). This statement can be confirmed by comparing the two TDC curves in Figure 6 and comparing the two cumulative curves of passengers in Figures 7(b) and 10. However, due to the change of demand curve \( W_p(t) \), the schedule delay for WT2 should considerably decrease as can be easily observed from Figure. 10. Therefore, the equilibrium cost \( T C_e \) derived from Equations (27) and (30) do not hold anymore.

Table 3 shows the numerical result of the change of equilibrium cost and its composition with the change of travel demand pattern. The time period for the increase in \( W_p(t) \) for WT2 is kept the same as [210, 270] min. It can be observed that the sum of SDC for WT2 is around half of those for WT1 although the sum of TDC is almost the same for Cases WT1 and WT2. Note that we use \( \sum T C/N_p \) instead of \( T C_e \) in Table 3 because the equilibrium cost for passengers with different work start times differs in Case WT1 although the sum of TDC is almost the same for Cases WT1 and WT2. Note that we use \( \sum T C/N_p \) instead of \( T C_e \) in Table 3 because the equilibrium cost for passengers with different work start times differs in Case WT2. \( \sum T C/N_p \) can be regarded as an average individual equilibrium cost in Case WT2 (\( \sum T C/N_p = T C_e \) in Case WT1 since all passengers share the same work start time). We can also find from the last column of Table 3 that the equilibrium patterns in Cases WT1 and WT2 are exactly the same under the same total travel demand \( N_p \). This numerical result confirmed that the work start time pattern does not affect the characteristics of the equilibrium in our model framework. Nevertheless, a distributed work start time undoubtly reduces the equilibrium costs of passengers by shortening their schedule delay.

### 5. A simple time-dependent timetable optimization

This section presents the optimization problem of a time-dependent timetable pattern as an application of the proposed model. The first subsection describes the problem setting, and the second subsection presents the results and provides their interpretations.
5.1. Problem setting

We consider the following simple time-dependent timetable pattern:

(i) Two dispatch frequencies (or train inflows) \(a_1\) and \(a_2\) \((a_1 \geq a_2)\) are used.
(ii) The train inflow is \(a_2\) initially; it becomes \(a_1\) from the beginning of the rush hour and lasts until the arrival time of the train departing from the system at \(t^*\), and then return to \(a_2\) (see Figure 11).

The condition (i) is widely adopted in practice. The timings of switching two dispatch frequencies indicated in the condition (ii) may be near-optimum because it is able to avoid degrading the rail transit system considerably under the equilibrium. Specifically, these conditions are aimed to prevent train outflow from becoming very low in the first half of the rush hour, and the traffic state from being in the congested regime in the second half. According to Figure 11, the ratio \(\omega \in (0, 1)\) of the duration
for $a_1$ to the rush hour is given as a constant:

$$\omega = \frac{\gamma (\alpha - \beta)}{\alpha (\beta + \gamma)}.$$  \hfill (36)

Thanks to the proper switching condition (ii) that exploits the characteristic of the equilibrium, the time-dependent timetable optimization problem to minimize the passengers’ total travel cost under the equilibrium can be expressed as the following concise mathematical problem with equilibrium constraints (MPEC) of determining $(a_1, a_2)$ only.

\[
\begin{align*}
\min_{a_1 \geq a_2 > 0} & \quad T_{C_e}(a_1, a_2 \mid N_p) \\
\text{subject to} & \quad \omega a_1 + (1 - \omega) a_2 \leq a_0
\end{align*}
\]  \hfill (37)

where $T_{C_e}(a_1, a_2 \mid N_p)$ is the equilibrium travel cost as a function of the decision variables. Equation (38) indicates the dispatch capacity constraint, where $a_0$ is the maximum available train inflow (fleet size) during rush hour. This problem can be easily solved by a brute-force search together with Algorithm 1. Note that, since the train inflow is time-dependent, the feasibility condition $a_p(n) \geq 0$ should be checked while evaluating the objective function.

### 5.2. Results

Before moving to the numerical results, we first show an optimal solution analytically to the problem subject to the free-flow train operation assumption. We expect this solution to be near-optimum because the operation in the congested regime vainly increases the travel time cost. By following almost the same manipulation as in Section 4.1.1, the equilibrium cost $T_{C_e}$ in Equation (37) can be obtained as follows:

$$T_{C_e}(a_1, a_2 \mid N_p) = \frac{2\alpha L N_p}{\mu l \left[ \left( \frac{1}{\beta} - \frac{1}{\alpha} \right) a_1 + \left( \frac{1}{\gamma} + \frac{1}{\alpha} \right) a_2 \right]}.$$  \hfill (39)
Moreover, if the dispatch capacity constraint (38) is inactive, it can be proved (see Appendix 2) that $TC_e$ is minimized when the following expression holds:

$$\frac{a_1}{a_2} = \frac{\zeta_1}{\zeta_2} \iff \frac{a_1}{a_2} = \frac{(\alpha + \gamma)(2\alpha - \beta)}{(\alpha - \beta)(2\alpha + \gamma)}.$$  \hspace{1cm} (40)

It is seen that $a_1/a_2$ increases with a decrease of $\alpha$ and with an increase of $\beta$ or $\gamma$. Besides, from Equation (19), we see that the average train flow remains constant (i.e. $q(n) = \zeta_1 a_1 = \zeta_2 a_2$) during the entire equilibrium period with this optimal setting.

Next, let us look at the numerical results (Figure 12) for Case WT1 under the parameter settings in Table 2 (and $a_0 = 18$ tr/h). The horizontal and vertical axes of this figure represent the high inflow rate $a_1$ and low inflow rate $a_2$, respectively; colour represents the value of the objective function. We evaluated the objective function every 0.1 tr/h for both dispatch frequencies $a_1$ and $a_2$. This figure reveals that the objective function is almost convex, and a unique optimal solution (S0) is obtained. The equilibrium cost of the optimal solution (39) plotted as $A_0$ ($a_1 = 15.5$ tr/h, $a_2 = 8.4$ tr/h, and $TC_e = 16.15 \$$) is close to that of S0 ($TC_e = 15.14 \$$). It confirms that the optimal timetable guaranteeing the free-flow operation is near-optimum.

The numerical results also indicate that too high or low dispatch frequencies can increase the equilibrium cost. To understand this phenomenon, the train dynamics for Scenario S0, S2, and S3 are shown in Figure 13. The ratio $a_1/a_2$ of Scenario S2 and S3 is the same as the optimal one S0, but with different average inflow rates. From the train cumulative curves in Figure 13(a), we see that the equilibrium period for S0 is shorter than that of S2 and S3. The reason can be understood from Figure 13(b). For S0, a high passenger arrival rate was achieved while maintaining an appropriately high average train flow. By contrast, for S2 and S3, trains cannot accommodate the high passenger arrival rate because of either too low or high average train flow. More specifically, for Scenario S2, the passenger arrival rate in Equation (21) is considerably reduced because of the insufficient dispatch frequency. As a result, a longer rush period is required to carry the same passenger demand $N_p$. For Scenario S3, the equilibrium period becomes larger because a considerable proportion of trains operate in the congested regime of the train-FD and a large average flow $a(t)$ in Equation (22) results in a small passenger arrival rate $a_p(n)$.

Also, Scenario 3 can be viewed as a particular type of ‘capacity increasing paradox’ (e.g. Braess 1968; Arnott, de Palma, and Lindsey 1993; Pas and Principio 1997; Zhang, Yang, and Liu 2014; Jiang and Szeto 2016) in the sense that the short-sighted supply-side improvement, which does not carefully
consider the demand-side feedback, makes the traffic state worsen than before. However, the precise mechanism is different from others because of the difference in the congestion mechanism. In our case, if the dispatch frequency is increased sufficiently to enter the congested regime, the increase in the travel time due to the train congestion around the desired arrival time disperses the passenger demand (i.e. the demand-side feedback), thereby increasing the rush hour period and equilibrium cost.

In addition, it is worth mentioning that the ratio $a_1/a_2$ in Scenario S0 is almost the same with Equation (40) even if a small number of trains operated in the congested regime. This may imply that flattening the train operation performance during the rush hour could be a useful strategy to reduce the equilibrium cost in a more general case other than the free-flow operation pattern.

Finally, Table 4 summarizes the travel costs for the timetable settings S0–S3 in Figure 12. Scenario S1 represents the case with the same average inflow as S0, but a smaller difference between $a_1$ and $a_2$. We can observe that the total travel cost $\sum TC (\sum TC = TC^C N_p)$ in Scenario S1–S3 are significantly higher than that of optimal one S0. Specifically, by comparing S0 and S1, the increase in the total schedule delay cost $\sum SDC (31\%)$ is greater than that of the total travel delay cost $\sum TDC (5\%)$. This suggests a primary deficiency of timetable patterns with a low $a_1/a_2$ ratio is that passengers cannot arrive at their desired arrival time $t^*$ sufficiently. A similar property can be obtained for a scenario with a redundant train supply (S3). However, when the train supply is insufficient (S2), passengers would suffer from a significantly longer travel delay ($\sum TDC$ increases by 27% compared with S0).

6. Conclusions

This study analyzed the equilibrium distribution of passenger arrivals for the morning commute problem in a congested urban rail transit system. We first developed a model for the problem in rail transit based on the train-FD and derived the equilibrium conditions. We then showed a solution method and
examined the characteristics of the equilibrium through both analytically and numerically. Finally, as an application of the proposed model, we analyzed a simple time-dependent timetable optimization problem with equilibrium constraints.

The proposed model is not only mathematically tractable but can also thoroughly consider the relations among passenger concentration, on-track congestion, and time-dependent timetable in a congested rail transit system. This enables us to investigate the characteristics of the equilibrium and optimal design of the timetable in a simple manner. The contributions of this study are summarized as follows:

- We revealed the evolution of rail transit flow and density under the equilibrium.
- We obtained the analytical solutions of the equilibrium in two of three patterns for a constant dispatch frequency.
- We further found several properties for the train operation under the equilibrium:
  (i) a ‘capacity increasing paradox’ exists in which a higher dispatch frequency can increase the equilibrium cost;
  (ii) an insufficient supply of rail transit mainly increases the total travel delay cost while redundant supply increases the total schedule delay cost;
  (iii) the average train flow maintains at an almost constant level under an optimal timetable setting.

The straightforward extensions of the proposed model include the consideration of elastic demands and captive users (e.g. Gonzales and Daganzo 2012). For the elastic demands, we only need to specify the travel demand $N_p(\text{TC}^c)$ as a monotonically decreasing function of the equilibrium travel cost (e.g. Arnott, de Palma, and Lindsey “A structural model of peak-period congestion” 1993; Zhou, Lam, and Heydecker 2005). Including captive users can be achieved by modifying $a_p(n)$ in Equation (15) as $a_p(n) = a_{pc} + a_{pf}(n)$, where $a_{pc}$ is the arrival rate of captive users, and $a_{pf}(n)$ is the arrival rate of flexible users for train $n$.

In this study, rail transit is assumed to be a homogeneous system in which both stations and passenger demand are evenly distributed. Thus, a train-FD model applicable to a heterogeneous railway system should be developed to deal with a more realistic situation. Considerations of heterogeneity in passenger preferences (i.e. the value of time) (Newell 1987; Akamatsu et al. 2021) and the costs/revenue of the transit agency in the optimization of timetable/fare settings are also important topics. The design of pricing schemes could be another fruitful future work. Since the analytical equilibrium solution is available of the proposed model, we could obtain insights into not only the first-best pricing scheme but also the second-best schemes that are generally formulated as MPEC (e.g. step tolls in Arnott, de Palma, and Lindsey 1990; Laih 1994; Lindsey, Van den Berg, and Verhoef 2012).

**Author contribution**

Conceptualization: Kentaro Wada; Methodology: Kentaro Wada and Jiahua Zhang; Formal analysis and investigation: Jiahua Zhang; Writing – original draft preparation: Jiahua Zhang; Writing – review and editing: Takashi Oguchi, Kentaro Wada and Jiahua Zhang; Funding acquisition: Kentaro Wada; Supervision: Takashi Oguchi and Kentaro Wada.

The preprint of this paper can be found in Zhang, Wada, and Oguchi (2021).

**Notes**

1. In these studies, the delay of trains was not considered. Even in such a situation, a considerable waiting (queueing) time of passengers (as well as in-vehicle crowding) can occur in an oversaturated railway system (Shi et al. 2018; Xu et al. 2019).
2. Throughout this paper, we do not consider the costs and revenue of the transit operator/agency. Thus, we treat the train operation exogenously except for Section 5 in which an optimal timetable setting is discussed from the passenger’s perspective.

3. The fundamental operating principle of this car-following model is consistent with practical control and existing studies (Carey and Kwieciński 1994; Higgins and Kozan 1998; Huisman et al. 2005), although it is more appropriate for moving block rather than fixed block railway signalling systems. Detailed validation of this modelling can be found in Seo, Wada, and Fukuda (2022).

4. Empirical investigations of the train-FD can be found in Fukuda, Imaoka, and Seo (2019) and Zhang and Wada (2019).

5. This paper does not consider dynamic pricing. Thus, a constant fare is excluded from the cost function.

6. Since we treat trains as a continuum or fluid, the number of trains can be a non-integer value.

7. Since the cumulative number of train departures from the system at time $t$, $D(t)$, is an increasing function of $t$, a one-to-one correspondence between the number of trains and their departure time exists. That is, $n = D(t) \Leftrightarrow t = D^{-1}(n)$.

8. It is equivalent to the condition in that the density $k(n)$ stays within the range of the train-FD for $q(n)$.

9. We have set the different $a_c$ for the three patterns just for avoiding overlapping.

**Acknowledgments**

The authors would like to express their sincere gratitude to the editor and four anonymous referees, whose helpful suggestions have improved the exposition of this paper.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

This study was financially supported by Japan Society for the Promotion of Science (JSPS KAKENHI) Grant No. JP17H03320.

**References**

Akamatsu, Takashi, Kentaro Wada, Takamasa Iryo, and Shunsuke Hayashi. 2021. “A New Look At Departure Time Choice Equilibrium Models with Heterogeneous Users.” *Transportation Research Part B: Methodological* 148: 152–182. doi:10.1016/j.trb.2021.04.003.

Arnott, Richard, André de Palma, and Robin Lindsey. 1990. “Economics of a Bottleneck.” *Journal of Urban Economics* 27 (1): 111–130. doi:10.1016/0094-1190(90)90028-L.

Arnott, Richard, André de Palma, and Robin Lindsey. 1993. “Properties of Dynamic Traffic Equilibrium Involving Bottlenecks, Including a Paradox and Metering.” *Transportation Science* 27 (2): 148–160. doi:10.1287/trsc.27.2.148.

Arnott, Richard, Andre de Palma, and Robin Lindsey. 1993. “A Structural Model of Peak-period Congestion: A Traffic Bottleneck with Elastic Demand.” *American Economic Review* 83 (1): 161–179.

Barrena, Eva, David Canca, Leandro C. Coelho, and Gilbert Laporte. 2014. “Single-line Rail Rapid Transit Timetabling Under Dynamic Passenger Demand.” *Transportation Research Part B: Methodological* 70: 134–150. doi:10.1016/j.trb.2014.08.013.

Braess, Dietrich. 1968. “Über Ein Paradoxon Aus Der Verkehrsplanung.” *Unternehmensforschung* 12 (1): 258–268.

Carey, Malachy, and Andrzej Kwieciński. 1994. “Stochastic Approximation to the Effects of Headways on Knock-on Delays of Trains.” *Transportation Research Part B: Methodological* 28 (4): 251–267. doi:10.1016/0191-2615(94)90001-9.

Daganzo, Carlos F. 1985. “The Uniqueness of a Time-dependent Equilibrium Distribution of Arrivals At a Single Bottleneck.” *Transportation Science* 19 (1): 29–37. doi:10.1287/trsc.19.1.29.

Daganzo, Carlos F. 2007. “Urban Gridlock: Macroscopic Modeling and Mitigation Approaches.” *Transportation Research Part B: Methodological* 41 (1): 49–62. doi:10.1016/j.trb.2006.03.001.

de Palma, André, Moez Kilani, and Stef Proost. 2015. “Discomfort in Mass Transit and Its Implication for Scheduling and Pricing.” *Transportation Research Part B: Methodological* 71: 1–18. doi:10.1016/j.trb.2014.10.001.

de Palma, Andre, Robin Lindsey, and Guillaume Monchambert. 2017. “The Economics of Crowding in Rail Transit.” *Journal of Urban Economics* 101: 106–122. doi:10.1016/j.jue.2017.06.003.

Edie, Leslie C. 1963. “Discussion of Traffic Stream Measurements and Definitions.” In *Proceedings of the Second International Symposium on the Theory of Traffic Flow*, edited by Joyce Almond, 139–154. Organisation for Economic Co-operation and Development.

Fosgerau, Mogens. 2015. “Congestion in the Bathtub.” *Economics of Transportation* 4 (4): 241–255. doi:10.1016/j.ecotra.2015.08.001.

Fosgerau, Mogens, and Kenneth A. Small. 2013. “Hypercongestion in Downtown Metropolis.” *Journal of Urban Economics* 76: 122–134. doi:10.1016/j.jue.2012.12.004.
Fukuda, Daisuke, Masahiro Imaoka, and Toru Seo. 2019. “Empirical Investigation of Fundamental Diagram for Urban Rail Transit Using Tokyo’s Commuter Rail Data.” In TRANSITDATA2019: 5th International Workshop and Symposium.

Geroliminis, N., and C. F. Daganzo. 2007. “Macroscopic Modeling of Traffic in Cities.” In Transportation Research Board 86th Annual Meeting.

Geroliminis, Nikolas, Jack Haddad, and Mohsen Ramezani. 2013. “Optimal Perimeter Control for Two Urban Regions with Macroscopic Fundamental Diagrams: A Model Predictive Approach.” IEEE Transactions on Intelligent Transportation Systems 14 (1): 348–359. doi: 10.1109/TITS.2012.2216877.

Geroliminis, Nikolas, and David M. Levinson. 2009. “Cordon Pricing Consistent with the Physics of Overcrowding.” In Transportation and Traffic Theory 2009: Golden Jubilee, edited by William H. K. Lam, S. C. Wong, and Hong K. Lo, 219–240. Boston, MA: Springer US.

Gonzales, Eric J., and Carlos F. Daganzo. 2012. “Morning Commute with Competing Modes and Distributed Demand: User Equilibrium, System Optimum, and Pricing.” Transportation Research Part B: Methodological 46 (10): 1519–1534. doi:10.1016/j.trb.2012.07.009.

Hao, Sijia, Rui Song, and Shiwei He. 2022. “Timetabling for a Congested Urban Rail Transit Network Based on Mixed Logic Dynamic Model.” Transportmetrica B: Transport Dynamics 10 (1): 139–158. doi:10.1080/21680566.2021.1965925.

Hendrickson, Chris, and George Kocur. 1981. “Schedule Delay and Departure Time Decisions in a Deterministic Model.” Transportation Science 15 (1): 62–77. doi:10.1287/trsc.15.1.62.

Higgins, Andrew, and Erhan Kozan. 1998. “Modeling Train Delays in Urban Networks.” Transportation Science 32 (4): 346–357. doi:10.1287/trsc.32.4.346.

Huisman, Dennis, Leo G. Kroon, Ramon M. Lentink, and Michiel JCM Vromans. 2005. “Operations Research in Passenger Railway Transportation.” Statistica Neerlandica 59 (4): 467–497. doi:10.1111/stan.2005.59.issue-4.

Jiang, Y., and W. Y. Szeto. 2016. “Reliability-based Stochastic Transit Assignment: Formulations and Capacity Paradox.” Transportation Research Part B: Methodological 93: 181–206. doi:10.1016/j.trb.2016.06.008.

Kariyazaki, Keiji, Naohiko Hibino, and Shigeru Morichi. 2015. “Simulation Analysis of Train Operation to Recover Knock-on Delay Under High-frequency Intervals.” Case Studies on Transport Policy 3 (1): 92–98. doi:10.1016/j.cstsp.2014.07.007.

Kato, Hironori, Yuichiro Kaneko, and Yoshihiko Soyama. 2012. “Departure-Time Choices of Urban Rail Passengers Facing Unreliable Service: Evidence from Tokyo.” In Proceedings of the International Conference on Advanced Systems for Public Transport 2012.

Kraus, Marvin, and Yuichiro Yoshida. 2002. “The Commuter’s Time-of-use Decision and Optimal Pricing and Service in Urban Mass Transit.” Journal of Urban Economics 51 (1): 170–195. doi:10.1006/juec.2001.2242.

Laith, Chen-Hsiu. 1994. “Queueing At a Bottleneck with Single-and Multi-step Tolls.” Transportation Research Part A: Policy and Practice 28 (3): 197–208. doi:10.1016/0191-2615(94)90007-8.

Li, Shukai, Maged M. Dessouky, Lixing Yang, and Ziyou Gao. 2017. “Joint Optimal Train Regulation and Passenger Flow Control Strategy for High-frequency Metro Lines.” Transportation Research Part B: Methodological 99: 113–137. doi:10.1016/j.trb.2017.01.010.

Li, Zhi-Chun, Hai-Jun Huang, and Hai Yang. 2020. “Fifty Years of the Bottleneck Model: A Bibliometric Review and Future Research Directions.” Transportation Research Part B: Methodological 139: 311–342. doi:10.1016/j.trb.2020.06.009.

Li, Zhi-Chun, William H. K. Lam, S. C. Wong, and Agachai Sumalee. 2010. “An Activity-based Approach for Scheduling Multimodal Transit Services.” Transportation 37 (5): 751–774. doi:10.1017/s11116-010-9291-z.

Lindsey, C. Robin, Vincent A. C. Van den Berg, and Erik T. Verhoef. 2012. “Step Tolling with Bottleneck Queuing Congestion.” Journal of Urban Economics 72 (1): 46–59. doi:10.1016/j.juec.2012.02.001.

MLIT. 2020. “The Visualization of Railway Lines’ Delay in Tokyo Metropolitan Area.” [online] https://www.mlit.go.jp/report/press/content/001328948.pdf.

Newell, Gordon F. 1987. “The Morning Commute for Nonidentical Travelers.” Transportation Science 21 (2): 74–88. doi:10.1287/trsc.21.2.74.

Newell, Gordon Frank. 2002. “A Simplified Car-following Theory: a Lower Order Model.” Transportation Research Part B: Methodological 36 (3): 195–205. doi:10.1016/S0191-2615(00)00044-8.

Pas, Eric L., and Shari L. Principio. 1997. “Braess’ Paradox: Some New Insights.” Transportation Research Part B: Methodological 31 (3): 265–276. doi:10.1016/s0191-2615(96)00024-0.

Robenek, Tomáš, Yousef Maknoon, Shadi Sharif Azadeh, Jianghang Chen, and Michel Bierlaire. 2016. “Passenger Centric Train Timetabling Problem.” Transportation Research Part B: Methodological 89: 107–126. doi:10.1016/j.trb.2016.04.003.

Seo, Toru, Kentaro Wada, and Daisuke Fukuda. 2017. “A Macroscopic and Dynamic Model of Urban Rail Transit with Delay and Congestion.” In TRB 96th Annual Meeting Compendium of Papers.

Seo, Toru, Kentaro Wada, and Daisuke Fukuda. 2022. “Fundamental Diagram of Urban Rail Transit Considering Train–Passenger Interaction.” Transportation, doi:10.1007/s11116-022–10281–0.

Shi, Jungang, Lixing Yang, Jing Yang, and Ziyou Gao. 2018. “Service-oriented Train Timetabling with Collaborative Passenger Flow Control on An Oversaturated Metro Line: An Integer Linear Optimization Approach.” Transportation Research Part B: Methodological 110: 26–59. doi:10.1016/j.trb.2018.02.003.

Smith, Michael J. 1984. “The Existence of a Time-dependent Equilibrium Distribution of Arrivals At a Single Bottleneck.” Transportation Science 18 (4): 385–394. doi:10.1287/trsc.18.4.385.
Tang, Yili, Hai Yang, Bo Wang, Jianling Huang, and Yunyun Bai. 2020. “A Pareto-improving and Revenue-neutral Scheme to Manage Mass Transit Congestion with Heterogeneous Commuters.” Transportation Research Part C: Emerging Technologies 113: 245–259. doi:10.1016/j.trc.2019.05.016.

Tian, Qiong, Hai-Jun Huang, and Hai Yang. 2007. “Equilibrium Properties of the Morning Peak-period Commuting in a Many-to-one Mass Transit System.” Transportation Research Part B: Methodological 41 (6): 616–631. doi:10.1016/j.trb.2006.10.003.

Tirachini, Alejandro, David A. Hensher, and John M. Rose. 2013. “Crowding in Public Transport Systems: Effects on Users, Operation and Implications for the Estimation of Demand.” Transportation Research Part A: Policy and Practice 53: 36–52. doi:10.1016/j.tra.2013.06.005.

Vickrey, William S. 1969. “Congestion Theory and Transport Investment.” American Economic Review 59 (2): 251–260. https://www.jstor.org/stable/1823678.

Vuchic, Vukan R. 2005. Urban Transit: Operations, Planning, and Economics. John Wiley & Sons.

Wada, Kentaro, Sangjoon Kil, Takashi Akamatsu, and Minoru Osawa. 2012. “A Control Strategy to Prevent Delay Propagation in High-frequency Railway Systems.” Journal of Japan Society of Civil Engineers, Ser. D3 (Infrastructure Planning and Management) 68 (5): I-1025–I-1034. doi:10.2208/jscejipm.68.I_1025.

Xu, Xinyue, Haiying Li, Jun Liu, Bin Ran, and Lingqiao Qin. 2019. “Passenger Flow Control with Multi-station Coordination in Subway Networks: Algorithm Development and Real-world Case Study.” Transportmetrica B: Transport Dynamics 7 (1): 446–472. doi:10.1080/21680566.2018.1434020.

Yang, Ruixia, Baoming Han, Qi Zhang, Zhenyu Han, and Yuxuan Long. 2022. “Integrated Optimization of Train Route Plan and Timetable with Dynamic Demand for the Urban Rail Transit Line.” Transportmetrica B: Transport Dynamics 11 (1): 93–126. doi:10.1080/21680566.2022.2040064.

Yang, Hai, and Yili Tang. 2018. “Managing Rail Transit Peak-hour Congestion with a Fare-reward Scheme.” Transportation Research Part B: Methodological 110: 122–136. doi:10.1016/j.trb.2018.02.005.

Yuan, Fuya, Huijun Sun, Liujie Kang, and Si Zhang. 2022. “Joint Optimization of Train Scheduling and Dynamic Passenger Flow Control Strategy with Headway-dependent Demand.” Transportation Research Part B: Transport Dynamics 10 (1): 627–651. doi:10.1080/21680566.2022.2025951.

Zhang, Junlin, Robin Lindsey, and Hai Yang. 2018. “Public Transit Service Frequency and Fares with Heterogeneous Users Under Monopoly and Alternative Regulatory Policies.” Transportation Research Part B: Methodological 117: 190–208. doi:10.1016/j.trb.2018.08.020.

Zhang, Jiahua, and Kentaro Wada. 2019. “Fundamental Diagram of Urban Rail Transit: An Empirical Investigation by Boston’s Subway Data.” In hEART 2019: 8th Symposium of the European Association for Research in Transportation.

Zhang, Jiahua, Kentaro Wada, and Takashi Oguchi. 2021. “Morning Commute in Congested Urban Rail Transit System: A Macroscopic Model for Equilibrium Distribution of Passenger Arrivals.” arXiv:2102.13454.

Zhang, Junlin, Hai Yand, Robin Lindsey, and Xinwei Li. 2020. “Modeling and Managing Congested Transit Service with Heterogeneous Users Under Monopoly.” Transportation Research Part B: Methodological 132: 249–266. doi:10.1016/j.trb.2019.04.012.

Zhang, Fangni, Hai Yang, and Wei Liu. 2014. “The Downs–Thomson Paradox with Responsive Transit Service.” Transportation Research Part A: Policy and Practice 70: 244–263. doi:10.1016/j.tra.2014.10.022.

Zhou, Jing, William H. K. Lam, and Benjamin G. Heydecker. 2005. “The Generalized Nash Equilibrium Model for Oligopolistic Transit Market with Elastic Demand.” Transportation Research Part B: Methodological 39 (6): 519–544. doi:10.1016/j.trb.2004.07.003.

Zhu, Ling, Shukai Li, Yuting Hu, and Bin Jia. 2022. “Robust Collaborative Optimization for Train Timetabling and Short-turning Strategy in Urban Rail Transit Systems.” Transportmetrica B: Transport Dynamics 11 (1): 147–173. doi:10.1080/21680566.2022.2048120.

Zorn, Lisa, Elizabeth Sall, and Daniel Wu. 2012. “Incorporating Crowding Into the San Francisco Activity-based Travel Model.” Transportation 39 (4): 755–771. doi:10.1007/s11116-012-9405-x.
Appendices

Appendix 1

The main notations used in the formulation are summarized in Table A1.

| Notation | Description |
|----------|-------------|
| $A(t), D(t)$ | Cumulative train arrival and departure at the rail transit system by time $t$ |
| $a(t), d(t)$ | Train arrival and departure flow rate at the rail transit system at time $t$ |
| $A_p(t), D_p(t)$ | Cumulative passenger arrival and departure at rail transit system by time $t$ |
| $t_0$ | Buffer time of a train at a station |
| $\nu_f$ | Desired cruising speed of a train |
| $L$ | Common trip length of passengers |
| $l$ | Average distance between adjacent stations |
| $\mu$ | Maximum flow rate of passenger boarding |
| $\tau$ | Reaction time of a train |
| $\delta$ | Minimum spacing of a train |
| $\alpha$ | Time value for a travel delay |
| $\beta, \gamma$ | Time values for earliness or lateness |
| $T_0$ | Minimum travel time outside equilibrium period |
| $t^*$ | Desired departure time of passengers |
| $s(t, t^*)$ | Schedule delay cost |
| $T(t), T(n)$ | Travel time of train $n$ departing from rail transit system at $t$ |
| $h_0(n), h_0(n)$ | Arriving and departing headway of train $n$ |
| $q(n)$ | Train average flow with respect to train $n$ during its trip |
| $a_p(n)$ | Passenger arrival rate for train $n$ at the stations along its trip |
| $t_0$ | Start time of the equilibrium period |
| $t_{ed}$ | End time of the equilibrium period |
| $t_m$ | Time when the travel time is maximized |
| $N_p$ | Total travel demand of passengers |
| $W_p(t)$ | Cumulative number of passengers who want to depart from rail transit by time $t$ |

Appendix 2

To simplify the discussion, the available train inflow is assumed to be sufficient, which means that constraint Equation (38) is inactive. According to Equation (39), $a_1$ and $a_2$ should be as large as possible to minimize $TC^e$ under a given $N_p$. Meanwhile, to ensure the free-flow operation, the following conditions similar to Equation (28) should be satisfied:

\[
\begin{align*}
\begin{cases}
\xi_1 a_1 \frac{T_0 + \frac{\beta}{\alpha} (t_m - t_0)}{L} \leq \frac{1}{1} \left[ 1 + \xi_1 a_1 \left( \frac{l - \delta}{\nu_f} - \tau \right) \right] \quad (A1) \\
\xi_2 a_2 \frac{T_0 + \frac{\beta}{\alpha} (t_m - t_0)}{L} \leq \frac{1}{1} \left[ 1 + \xi_2 a_2 \left( \frac{l - \delta}{\nu_f} - \tau \right) \right]
\end{cases}
\end{align*}
\]

Substituting $TC^e(a_1, a_2 | N_p) = \beta (t_m - t_0)$ into these two conditions, we have

\[
\begin{align*}
\begin{cases}
G_1 (a_1, a_2 | N_p) = \xi_1 a_1 \frac{l}{L} \left( T_0 + \frac{TC^e(a_1, a_2 | N_p)}{\alpha} \right) - \frac{l - \delta}{\nu_f} + \tau \leq 1 \\
G_2 (a_1, a_2 | N_p) = \xi_2 a_2 \frac{l}{L} \left( T_0 + \frac{TC^e(a_1, a_2 | N_p)}{\alpha} \right) - \frac{l - \delta}{\nu_f} + \tau \leq 1
\end{cases} \quad (A2)
\end{align*}
\]

We then prove that $a_1$ and $a_2$ are maximized when $G_1 = 1$ and $G_2 = 1$ hold simultaneously, which implies Equation (40). First, if $G_1 = 1$ and $G_2 = 1$, we cannot increase any of $a_1$ and $a_2$ because $\partial G_1 / \partial a_1 > 0$ and $\partial G_2 / \partial a_2 > 0$ (i.e. the conditions (A2) are violated). Thus, the remaining part of the proof is to show that if any of two inequalities is smaller than 1, both $a_1$ and $a_2$ can be increased while not violating the conditions (A2). Suppose $G_1 = 1$ but $G_2 < 1$. Then, we can increase $a_2$ until $G_2 = 1$ while maintaining $a_1$ unchanged because $\partial G_2 / \partial a_2 > 0$. Meanwhile, increasing $a_2$ leads to $G_1 < 1$ because $\partial G_1 / \partial a_2 < 0$, i.e. the conditions (A2) hold. The same discussion holds true for $G_1 < 1$ but $G_2 = 1$. Therefore, Equation (40) is proved.