ELSEVIER

19th International Symposium on Transportation and Traffic Theory

Design of a rail transit line for profit maximization in a linear transportation corridor

Zhi-Chun Li,[a,b] William H.K. Lam,[a,]* S.C. Wong,[c] A. Sumalee[a]

[a]*Department of Civil and Structural Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China*
[b]*School of Management, Huazhong University of Science and Technology, Wuhan 430074, China*
[c]*Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong, China*

Abstract

This paper addresses the design problem of a rail transit line located in a linear urban transportation corridor. The service variables designed are a combination of rail line length, number and locations of stations, headway and fare. Two profit maximization models, which account for the effects of different transit pricing structures (flat and distance-based fare regimes), are proposed. In the proposed models, the effects of passenger demand elasticity and population density along the urban corridor are explicitly considered. The solution properties of the proposed models are explored and compared analytically, and the indifference condition for the two fare regimes in terms of the operator's net profit is identified. A heuristic solution algorithm to solve the proposed models is presented. Numerical examples are provided to show the effects of the fare regimes, rail capital cost and urban configuration (in terms of urban population distribution and corridor length) on the design of the rail transit line and the profitability of the rail transit operations.
© 2011 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.

Keywords: Transportation corridor; rail line design; profit maximization; population density; transit pricing structure, urban form

1. Introduction

1.1. Background and motivation

Over the past few decades, rapid urban expansion due to urbanization and economic growth in some large Asian cities, including Shanghai and Hong Kong, has drastically increased the size of these cities. Traffic congestion has worsened due to the shortage of space for road expansion projects to accommodate the growing traffic demand in urban areas. To address this problem, the local authorities of these cities have launched rail transit development projects, which include extension of existing rail transit lines and construction of new rail transit lines. For instance, the Shanghai municipal government is currently extending Rail Line 11 about 5.76 km westwards and creating four new stations on the line, while the Hong Kong government recently approved a proposal to build a new metro line to

---

connect Shatin New Town to Central (i.e., the central business district (CBD) of Hong Kong), with a total length of 17 km and 10 stations. The construction of this new metro project is expected to start in 2011 and be completed in 2019.

In principle, the basic parameters to be determined in planning a rail transit line project include the rail line length, number and locations of stations, headway and fare (see, e.g. Vuchic, 2005; and the references shown in Table 1). The design of these parameters depends very much on the population density in the planning area. This is because the urban population density directly influences the level of passenger demand. For instance, in a sparsely populated city (e.g. many Western cities), operators prefer to short rail transit lines in order to minimize their costs (Spasovic and Schonfeld, 1993; Spasovic et al., 1994). However, in a densely populated city, such as Hong Kong in which most people use transit services for their daily travel, a benefit-driven operator has an incentive to extend the rail transit line from the city's CBD area into its outer areas so as to procure more profit. It is, therefore, important to address the relationship between the design parameters of the rail transit line and the urban population density.

Obviously, there are various tradeoffs between the extension of a rail transit line and its associated costs. For instance, the length of a rail line is closely related to its service coverage and its capital and operating costs. A longer rail line provides greater service coverage but incurs higher capital and operating costs, whereas a shorter rail line has lower capital and operating costs but offers less service coverage. The station spacing along a rail line directly affects the train operating speed and train dwell time at stations, and thus passenger demand on that line. In general, shorter station spacing can decrease the average passenger access time to stations. However, it also increases the average passenger in-vehicle travel time and train operating costs because of higher acceleration and deceleration delays caused by frequent stops. Conversely, longer station spacing can increase the train operating speed and decrease the average passenger in-vehicle travel time, but also increases the average passenger access time to stations. Since these tradeoffs are directly related to the revenues and thus profits of the rail transit operations, all of these parameters – the rail line length, number and spacing of stations, headway and fare – should be carefully designed. The present study addresses this design problem for strategic planning purposes.

1.2. Literature review

Significant progress has been made in transit service design models since the pioneering work of Vuchic and Newell (1968) in developing an analytical continuum model to optimize rail station spacing. For the convenience of readers, we have summarized in Table 1 some principal contributions of various analytical models to transit service design, which include: decision variables, such as the location or spacing of transit routes and/or stations, headway, and fare; the objective function, which is typically the minimization of the total system cost (i.e., the sum of operator and user costs); the transit mode involved, such as rail, bus or feeder bus; the geometry of transit lines, such as a linear structure or rectangular grid; and demand characteristics, including fixed or elastic demand, uniform or non-uniform demand distribution, and one-to-one, many-to-one, or many-to-many travel patterns.

Most existing transit models, as shown in Table 1, usually aim to minimize the sum of operator and user costs. This can be attributed to the relatively low population densities within large Western cities leading to government subsidy to the transit industry. However, this is not the case in some Eastern Asian cities, particularly those with high-density population development, such as Hong Kong, in which the transit industry can operate profitably without government subsidy. This is because a large proportion of the population in Hong Kong uses transit services as the main mode of transportation, and over 90% of the 11 million daily person-trips are served by privately operated public transit modes (Transport Department, 2003). Under Hong Kong's operating environment, the principal objective of private transit operators is neither welfare gain nor the efficient utilization of road space, but rather profit maximization (Lam and Zhou, 2000; Zhou et al., 2005; Li et al., 2009). It is thus important to address the interrelation between the population density and the transit operator's profit.

The literature review also reveals that the focuses of the previous related studies have mainly been on how to optimize the design variables of the transit services approximately and/or numerically. Little attention has been paid, however, to investigating the solution properties of transit design problems, such as the concavity of objective functions concerned with regard to their decision variables. In addition, there is scant research into the effects of different transit pricing structures. In reality, a flat fare regime and a distance-based one can have significantly different effects on the attractiveness of transit services and the level of passenger demand.

Table 1 Some major analytical models for transit service design

Citation	Decision variables	Objective function	Transit mode	Network geometry	Passenger demand
Vuchic and Newell (1968)	Station location and spacing	Min. total user travel time	Rail	Linear	Uniform, inelastic, many to one
Vuchic (1969)	Station location and spacing	Max. number of passengers	Rail	Linear	Uniform, inelastic, many to one
Hurdle (1973)	Route spacing and frequency (headway)	Min. operator and user cost	Feeder bus to rail	Rectangular grid	Piecewise uniform, inelastic, many to one
Hurdle and Wirasinghe (1980)	Station location or spacing	Min. operator and user cost	Rail	Rectangular grid	Uniform, inelastic, many to one
Wirasinghe and Ghoneim (1981)	Stop spacing	Min. operator and user cost	Bus	Linear	Piecewise uniform, inelastic, many to many
Kocur and Hendrickson (1982)	Route spacing, headway and fare	Max. operator profit, Max. user benefit	Bus	Rectangular grid	elastic, many to one
Wirasinghe and Seneviratne (1986)	Route length	Min. operator and user cost	Rail	Linear	Piecewise uniform, inelastic, many to one
Kuah and Perl (1988)	Route spacing, headway and stop spacing	Min. operator and user cost	Feeder bus to rail	Rectangular grid	Piecewise uniform, inelastic, many to one
Chang and Schonfeld (1993)	Zone size, route length, route spacing and headway	Min. operator and user cost	Bus	Rectangular grid	Uniform, inelastic, many to one
Spasovic et al. (1994)	Route spacing, length, headway and fare	Max. operator profit, Max. social welfare	Bus	Rectangular grid	Uniform, elastic, many to one
Liu et al. (1996)	Rail line location and length	Min. operator and user cost	Feeder bus to rail	Linear	Piecewise uniform, inelastic, many to many
Chien and Schonfeld (1997)	Route spacing and headway	Min. operator and user cost	Bus	Rectangular grid	Uniform, inelastic, many to many
Furth and Rahbee (2000)	Stop spacing	Min. operator and user cost	Bus	Linear	Uniform, inelastic, many to one
Saka (2001)	Stop spacing	Min. bus fleet size	Bus	Linear	inelastic, one to one
Wirasinghe et al. (2002)	Terminus location	Min. operator and user cost	Feeder bus to rail	Linear	Piecewise uniform, inelastic, many to many
Chien and Qin (2004)	Stop location	Min. operator and user cost	Bus	Linear	Discrete, inelastic, many to one

1.3. Problem statement and contributions

To address the foregoing issues, this paper develops analytical models for designing the service variables of a rail transit line located in a linear transportation corridor of length B, as shown in Fig. 1. In this figure, the rail line designed extends from the CBD of the city outward, and is represented by an ordered sequence of stations $\{1, 2, \cdots, N+1\}$. The symbol D_s represents the distance of station s from the CBD, and D_1 the length of the rail line. The service variables designed include the rail line length, D_1; number of stations, $N+1$; station locations, $D_s, s = 2, \cdots, N$; train headway, H; and fare for traveling from station s to the CBD, f_s. For presentation purpose, all variables and parameters used throughout this paper are defined in Table 2.
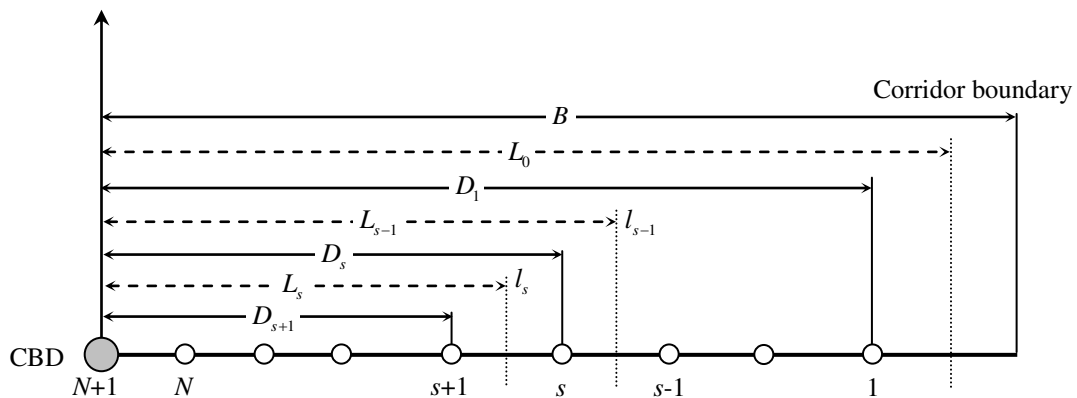

Fig. 1. The rail line configuration along a linear transportation corridor.

This paper makes three main contributions to the previous related studies. First, in order to examine the effects of different transit pricing regimes (flat and distance-based fare regimes), two profit maximization models are proposed. Second, the solution properties of the proposed models, particularly those with uniform population density and/or even (average) station spacing, are explored and compared analytically. The indifference condition of the operator's net profits for the flat and distance-based fare regimes is also identified. Third, a heuristic solution algorithm to solve the proposed models is developed. With the proposed models and solution algorithm, effects of some key model parameters, such as population density, rail capital cost and corridor length, are examined and evaluated in this paper.

The remainder of this paper is organized as follows. In the next section, some basic model assumptions are described and the passenger demand for each station is defined. Section 3 presents the profit maximization models for the flat and distance-based fare regimes, respectively. The solution properties of the two proposed models are then examined and discussed. In addition, the constraint conditions of the models are also presented. In Section 4, a heuristic solution algorithm is developed for jointly solving the design variables of the rail transit line. In Section 5, an example is used to illustrate the application of the proposed models and solution algorithm. Finally, conclusions are given in Section 6 together with recommendations for further studies.

Table 2 Notation

Symbol	Definition	Baseline value
B	length of transportation corridor (km)	-
C_O	train operating cost (\$/h)	-
C_L	rail line cost (\$/h)	-
C_S	rail station cost (\$/h)	-
D_s	distance of station s from the CBD; $\mathbf{D} = (D_s, s = 1, 2, \cdots, N)$ is corresponding vector (km)	-
e_a	sensitivity parameter for access time (1/h)	0.98
e_w	sensitivity parameter for wait time (1/h)	0.98
e_t	sensitivity parameter for in-vehicle time (1/h)	0.49
e_f	sensitivity parameter for fare (1/\$)	0.098
f_s	fare for traveling from station s to the CBD (\$ for flat fare, \$/km for distance-based fare)	-
F	fleet size (or number of trains) on the rail line (vehicles)	-
$g(x)$	population density at distance x from the CBD (persons/km^2)	-
g_0	population density in the CBD (persons/km^2)	-
G	total number of population in the planning area (persons)	-
H	train headway (h/vehicle)	-
K	capacity of vehicles, including seated and standing passengers (pass/veh)	1800
L_s	distance of the passenger watershed line l_s from the CBD (km)	-
$N+1$	total number of stations on the rail line	-
$P(x)$	potential passenger demand density at location x (pass/km-h)	-
P_0	potential passenger demand density in the CBD; $P_0 = \phi \eta g_0$ (pass/km-h)	-
$q(x,s)$	density of passenger demand for station s at location x (pass/km-h)	-
Q_s	passenger demand for station s (pass/h)	-
t_s	average passenger in-vehicle time from station s to the CBD (h)	-
T_{s1}	non-stop line-haul travel time from station s to the CBD (h)	-
T_{s2}	total train dwell time from station s to the CBD (h)	-
T_0	constant terminal time on the circular line (h)	0.08
$u_s(x)$	passenger access time to station s from location x (h)	-
V_t	average train cruise speed (km/h)	40
V_a	average walking speed of passengers (km/h)	4.0
w_s	average passenger wait time at station s (h)	-
π	net profit of operator ($\bar{\pi}$ for flat fare regime, and $\hat{\pi}$ for distance-based fare regime) (\$/h)	-
α	ratio of passenger waiting time to train headway	0.5
β_0	average train dwell time at a rail station (h)	0.01
γ_0	fixed component of rail line cost (\$/h)	750
γ_1	variable component of rail line cost (\$/km-h)	300
μ_0	fixed component of train operating cost (\$/h)	1350
μ_1	variable component of train operating cost (\$/veh-h)	540
Λ_0	fixed component of rail station cost (\$/h)	1250
Λ_1	variable component of rail station cost (\$/station-h)	500
θ	density gradient describing how rapidly the density falls as the distance increases (1/km)	-
η	average number of trips to the CBD per person per day	1.0
ϕ	peak-hour factor, i.e. the ratio of peak-hour flow to daily average flow	0.1
ζ	number of terminal times on the rail line	1.0

2. Basic considerations

2.1. Assumptions

To facilitate the presentation of the essential ideas without loss of generality, the following basic assumptions are made in this paper.

A1 The corridor connecting the city's CBD and suburb is assumed to be linear, following the previous related studies including those of Wang et al. (2004) and Liu et al. (2009) and some listed in Table 1.

A2 The study period is assumed to be a one-hour period, such as the morning peak hour, which is usually the most critical period in a day. Therefore, this paper mainly focuses on a many-to-one travel demand pattern.

A3 Passengers are assumed to board trains at the nearest rail station in terms of access time. Trains running along the rail line stop at each station on that line, and the average train dwell time at each station is assumed to be a constant. These assumptions have also been adopted in some previous related studies, such as those of Wirasinghe and Ghoneim (1981), Kuah and Perl (1988), Chien and Schonfeld (1997, 1998), and Chien and Qin (2004), but can be relaxed in further studies.

A4 The population density along the corridor is specified as a negative exponential function (see, e.g. Anas, 1982; O'Sullivan, 2000). We represent the population density at distance x from the CBD as $g(x) = g_0 \exp(-\theta x)$, $\forall x \in [0, B]$, where g_0 is the population density in the CBD and θ (≥ 0) is the density gradient describing how rapidly the density falls as the distance increases (see Fig. 5 later). The larger the value of θ, the smaller the population density at the city's edge and the more compact the city. That is, a smaller value of θ means a more decentralized city. In particular, when θ equals 0, the negative exponential population density function is reduced to a uniform one. With this assumption, the total number of population G in the study area is then given by $G = \int_0^B g_0 \exp(-\theta x) dx$.

A5 An elastic demand density function is defined to capture the responses of passengers to the quality of the rail transit line service, which is measured by a generalized travel cost that is a weighted combination of the access time to stations, wait time at stations, in-vehicle time, and fare. The responses include the decision of switching to an alternative mode (e.g., auto, bus, or walk) and the decision of not making the journey at all (Li et al., 2009).

2.2. Passenger demand for each station

As both stations on any segment of the rail line are competing for passengers between those two stations, there exists a *passenger watershed line* that divides the line segment between two adjacent stations into two sub-segments, as shown in Fig. 1. The passengers in the two sub-segments respectively use the upstream and downstream stations of the line segment. Let l_s be the passenger watershed line between stations s and $s + 1$, and L_s be the distance of the passenger watershed line l_s from the CBD. Based on A3, the watershed line l_s is located at the middle point of the line segment $(s, s+1)$, which implies

$$L_s = \frac{D_s + D_{s+1}}{2}, \ \forall s = 1, 2, \cdots, N , \tag{1}$$

where $D_{N+1} = 0$.

Let $q(x, s)$ denote the density of passenger demand (i.e. the number of passengers per unit of distance) for station s at location x. The total passenger demand for station s, Q_s, is then given by

$$Q_s = \int_{L_s}^{L_{s-1}} q(x, s) dx, \ \forall s = 1, 2, \cdots, N , \tag{2}$$

where $L_s, s = 1, 2, \cdots, N$ can be given by Eq. (1). L_0 represents the maximum location that the residents between station 1 and the corridor boundary will use the rail service, as shown in Fig. 1. Beyond that location (i.e. $x > L_0$), no one will patronize the rail system. Thus, L_0 holds

$$q(L_0, 1) = 0, \ L_0 \in [D_1, B] , \tag{3}$$

where $q(L_0, 1)$ is the density of passenger demand for station 1 at location L_0.

In order to determine $q(x, s)$, we first define the potential passenger demand density at location x, which is denoted by $P(x)$. Define η as the average number of trips to the CBD per person per day within the study area, then $\eta g(x)$ is the potential passenger demand density at location x per day in terms of A4. Let ϕ be the peak-hour factor, i.e. the ratio of peak-hour flow to the daily average flow. $P(x)$ can then be given by

$$P(x) = \phi \eta g_0 \exp(-\theta x) = P_0 \exp(-\theta x), \ \forall x \in [0, B] , \tag{4}$$

where P_0 is the (peak-hour) potential passenger demand density in the CBD and $P_0 = \phi \eta g_0$. The parameter ϕ is used to convert the traffic volume from a daily basis to an hourly basis.

Passenger demand for rail line service is usually sensitive to rail fare level and various time components (walk/access time to station, wait time and in-vehicle time) and thus is elastic. In order to model the effects of passenger demand elasticity, in this paper a linear elastic demand density function is used and specified as

$$q(x, s) = P(x)\big(1 - e_a u_s(x) - e_w w_s - e_t t_s - e_f f_s\big), \ \forall x \in [0, B], s = 1, 2, \cdots, N , \tag{5}$$

where $u_s(x)$ is the passenger access time to station s from location x, which is related to the distance of location x from station s; w_s is the average passenger wait time at station s; t_s is the passenger in-vehicle time from station s to the CBD; and f_s is the fare for traveling from station s to the CBD. e_a, e_w, e_t and e_f are the sensitivity parameters for the access time, wait time, in-vehicle time and fare, respectively.

To ensure the non-negativity of the passenger demand, the following condition should be satisfied

$$0 \leq 1 - e_a u_s(x) - e_w w_s - e_t t_s - e_f f_s \leq 1, \ \forall x \in [L_s, L_{s-1}], s = 1, 2, \cdots, N . \tag{6}$$

It should be pointed out that the parameters e_a, e_w, e_t and e_f in the linear demand function (5) are not the actual measures of demand elasticities. The ratios e_a/e_f, e_w/e_f and e_t/e_f determine the values of access time, wait time and in-vehicle time, respectively. The value of walking/access time is generally larger than the value of in-vehicle time (Chang and Schonfeld, 1991), and thus $e_a > e_t$ holds.

In addition, the linear demand function, as adopted here, has been extensively used in demand models due to its convenience for analytical tractability. Other alternative demand functions, such as exponential demand function, can also be adopted. However, it is usually difficult, if not impossible, to derive a closed-form solution. Alternatively, a non-linear demand function can be approximated as a linear demand function by using the first-order Taylor series expansion.

We now define the time components that are included in the linear demand function (5). The passenger access time $u_s(x)$ depends on the walking distance between location x and station s and the walking speed of passengers, V_a. It is expressed as

$$u_s(x) = \begin{cases} (D_s - x)/V_a , & \forall x \leq D_s \\ (x - D_s)/V_a , & \forall x > D_s \end{cases}, \ \forall x \in [0, B], s = 1, 2, \cdots, N . \tag{7}$$

The average passenger wait time at station s, w_s, can be calculated by

$$w_s = \alpha H, \quad \forall s = 1, 2, \cdots, N,$$ (8)

where H is the headway of the rail service, and α is a calibration parameter that depends on the distributions of train headway and passenger arrival. The value $\alpha = 0.5$ is commonly used to suggest a constant headway between trains and a uniform random passenger arrival distribution.

The passenger in-vehicle time from station s to the CBD, t_s, comprises the non-stop line-haul travel time, T_{s1}, and train dwell time, T_{s2}, at rail stations, i.e.,

$$t_s = T_{s1} + T_{s2}, \quad \forall s = 1, 2, \cdots, N,$$ (9)

where T_{s1} can be calculated by the in-vehicle length of the trip being made by the passengers divided by the average train cruise speed, V_t, i.e.,

$$T_{s1} = \frac{D_s}{V_t}, \quad \forall s = 1, 2, \cdots, N.$$ (10)

According to A3, the average train dwell time at each rail station is a constant. Consequently, the total train dwell time from station s to the CBD, T_{s2}, can be calculated by

$$T_{s2} = \beta_0 (N + 1 - s), \quad \forall s = 1, 2, \cdots, N,$$ (11)

where β_0 is the average train dwell time at a station, which can be calibrated with observed data (Lam et al., 1998).

Substituting Eqs. (5)-(11) into Eq. (2), Q_s can then be rewritten as

$$Q_s = \lambda_s \int_{L_s}^{L_{s-1}} P(x)dx - \frac{e_a}{V_a} \left(\int_{L_s}^{D_s} P(x)(D_s - x)dx + \int_{D_s}^{L_{s-1}} P(x)(x - D_s)dx \right), \quad \forall s = 1, 2, \cdots, N,$$ (12)

where

$$\lambda_s = 1 - e_w w_s - e_t t_s - e_f f_s$$

$$= 1 - e_w \alpha H - e_t \left(\frac{D_s}{V_t} + \beta_0 (N + 1 - s) \right) - e_f f_s, \quad \forall s = 1, 2, \cdots, N.$$ (13)

On the basis of Eqs. (3)-(13), the maximum service coverage L_0 of the rail transit line can be given by

$$L_0 = D_1 + \frac{V_a}{e_a} \lambda_1,$$ (14)

where λ_1 can be determined by Eq. (13).

3. Model formulation

3.1. Profit maximization models

As previously stated, the objective of the rail transit operator is to maximize its net profit. The net profit, denoted as π, is the total revenue, R, which is generated from the passenger fares, minus the total cost, C. It is expressed as

$$\pi = R - C.$$ (15)

In the following, we first define the total cost C. As described in Chien and Schonfeld (1997, 1998), the total cost C is incurred by train operations, rail line, and rail stations, and thus consists of the following three cost components: train operating cost, C_O; rail line cost, C_L; and rail station cost, C_S. It is represented as

$$C = C_O + C_L + C_S. \tag{16}$$

The train operating cost C_O comprises the fixed operating cost, μ_0, and variable operating cost, $\mu_1 F$, where F is the fleet size (or the number of trains) on that line and μ_1 is the hourly operating cost per train. It is formulated as

$$C_O = \mu_0 + \mu_1 F , \tag{17}$$

where F equals the vehicle round journey time, Θ, divided by headway H, i.e.,

$$F = \frac{\Theta}{H}, \tag{18}$$

where the round journey time Θ comprises the terminal time, line-haul travel time and train dwell time at stations, which is expressed as

$$\Theta = \zeta T_0 + 2(T_{11} + T_{12}), \tag{19}$$

where T_0 is the constant terminal time on the circular line and ζ is the number of terminal times on the line. T_{11} and T_{12} are, respectively, the total line-haul travel time and total dwell time for train operations from station 1 to the CBD. From Eqs. (10) and (11), we have $T_{11} = D_1/V_t$ and $T_{12} = \beta_0 N$.

The rail line cost, C_L, is the sum of the fixed costs, γ_0 (e.g., line overhead cost), and variable costs, $\gamma_1 D_1$ (e.g., land acquisition, line construction, maintenance, and labor costs), proportional to the rail line length D_1; i.e.,

$$C_L = \gamma_0 + \gamma_1 D_1 , \tag{20}$$

where γ_1 is the hourly rail line operating cost per kilometer.

The rail station cost, C_S, includes fixed costs (e.g., station overhead cost) and variable costs (e.g., station land acquisition, design and construction, operating, and maintenance costs). The total variable cost is determined by the number of stations multiplied by the average unit station cost. C_S can thus be expressed as

$$C_S = \Lambda_0 + \Lambda_1 (N+1) , \tag{21}$$

where Λ_0 is the fixed cost for station operations and Λ_1 is the hourly operating cost per station.

We now define the total operating revenue R that appears in Eq. (15). It is the sum of the number of passengers boarding at each station multiplied by the corresponding fare, i.e.,

$$R = \sum_{s=1}^{N} f_s Q_s , \tag{22}$$

where the passenger demand for station s, Q_s, is given by Eq. (12). The transit fare f_s depends on the fare regimes adopted.

Different transit fare regimes can lead to different levels of operating revenue and thus profit of the operator. In this paper, we consider two types of fare regimes: a flat fare regime in which all passengers are charged the same fare regardless of the length of their trips, and a distance-based regime in which the fares grow linearly with the distance travelled by passengers. Mathematically, the two fare regimes are, respectively, represented as follows.

For the flat fare regime,

$$f_s = \bar{f}, \ \forall s = 1, 2, \cdots, N , \tag{23}$$

where $\bar{f}$ is a constant.

For the distance-based fare regime,

$$f_s = f_0 + \hat{f}D_s, \ \forall s = 1, 2, \cdots, N, \tag{24}$$

where f_0 and $\hat{f}$ are the fixed and variable components of the distance-based fare, respectively.

In view of Eqs. (15)-(24), the profit maximization problems for the flat and distance-based fare regimes can, respectively, be formulated as

$$\max \bar{\pi}\left(\mathbf{D}, H, \bar{f}\right) = \bar{f}\sum_{s=1}^{N} Q_s - \left[\mu_0 + \frac{\mu_1}{H}\left(\zeta T_0 + \frac{2D_1}{V_t} + 2\beta_0 N \right) \right] - \left(\gamma_0 + \gamma_1 D_1 \right) - \left[\Lambda_0 + \Lambda_1 (N+1) \right], \tag{25}$$

and

$$\max \hat{\pi}\left(\mathbf{D}, H, \hat{f}\right) = \sum_{s=1}^{N} \left(f_0 + \hat{f}D_s \right) Q_s - \left[\mu_0 + \frac{\mu_1}{H}\left(\zeta T_0 + \frac{2D_1}{V_t} + 2\beta_0 N \right) \right] - \left(\gamma_0 + \gamma_1 D_1 \right) - \left[\Lambda_0 + \Lambda_1 (N+1) \right], \tag{26}$$

where the bolded symbol " $\mathbf{D}$ " is the vector of station locations, i.e. $\mathbf{D} = \left(D_s, \forall s = 1, 2, \cdots, N \right)$. Q_s can be determined by Eq. (12). The decision variables include the rail line length D_1, station locations $D_2, \cdots, D_N$, train headway H, and fare $\bar{f}$ for the flat fare regime or $\hat{f}$ for the distance-based fare regime.

The optimal solutions for the rail line length, station location (or spacing), headway and fare can be obtained by setting the partial derivatives of each objective function with respect to its decision variables equal to zero and solving them simultaneously. We then have the following proposition (the proof is given in Appendix A).

Proposition 1. *The optimal rail line length, station location (or spacing), headway and fare solutions for the flat and distance-based fare regimes satisfy the systems of equations as shown in Table 3, respectively.*

We now look at a special case with an even station spacing, which can be regarded as the average station spacing. The even (or average) station spacing solution can serve as a benchmark indicator for the planning and design of rail transit line service, particularly at the early stage of the design of the transit line.

Let δ represent the even (or average) station spacing of the rail line, one then obtains $D_s = (N+1-s)\delta$ and thus $L_s = (N-s+\frac{1}{2})\delta$ in terms of Eq. (1). Substituting them into Eqs. (12)-(14) and (25)-(26), we can then derive the first-order optimality conditions of the optimization models (25) and (26) as follows (this proof is similar to Proposition 1 and omitted here).

Proposition 2. *The optimal even (average) station spacing, headway and fare solutions for the flat and distance-based fare regimes satisfy the systems of equations as shown in Table 4, respectively.*

Table 3 Optimal rail line length, station location, headway and fare solutions for different fare regimes

Flat fare regime	Distance-based fare regime

$$\begin{cases} \dfrac{\partial \pi}{\partial D_s} = \bar{f}\sum_{i=s-1}^{s+1}\dfrac{\partial Q_i}{\partial D_s} - \Delta_s\left(\dfrac{2\mu_1}{HV_t}+\gamma_1\right) = 0, \; s=1,\cdots,N, \\[4mm] H = \sqrt{\dfrac{\mu_1\left(\zeta T_0 + \dfrac{2D_1}{V_t}+2\beta_0 N\right)}{\alpha e_w \bar{f}\sum_{s=1}^{N}\int_{L_s}^{L_{s-1}}P(x)dx}}, \\[6mm] \bar{f} = \dfrac{\sum_{s=1}^{N}Q_s}{e_f \sum_{s=1}^{N}\int_{L_s}^{L_{s-1}}P(x)dx}, \end{cases}$$

$$\begin{cases} \dfrac{\partial \hat{\pi}}{\partial D_s} = \hat{f}Q_s + \sum_{i=s-1}^{s+1}\left(f_0+\hat{f}D_i\right)\dfrac{\partial Q_i}{\partial D_s} - \Delta_s\left(\dfrac{2\mu_1}{HV_t}+\gamma_1\right) = 0, \; s=1,\cdots,N, \\[4mm] H = \sqrt{\dfrac{\mu_1\left(\zeta T_0 + \dfrac{2D_1}{V_t}+2\beta_0 N\right)}{\alpha e_w \sum_{s=1}^{N}\left(f_0+\hat{f}D_s\right)\int_{L_s}^{L_{s-1}}P(x)dx}}, \\[6mm] \hat{f} = \dfrac{\sum_{s=1}^{N}D_s Q_s - f_0 e_f \sum_{s=1}^{N}D_s\int_{L_s}^{L_{s-1}}P(x)dx}{e_f \sum_{s=1}^{N}D_s^{\,2}\int_{L_s}^{L_{s-1}}P(x)dx}, \end{cases}$$

where $\Delta_s = 1$ if $s=1$, and 0 otherwise. $\partial Q_i/\partial D_s$ can be given by

$$\begin{cases} \dfrac{\partial Q_{s-1}}{\partial D_s} = \dfrac{1}{2}P(L_{s-1})\left(-\lambda_{s-1}+\dfrac{e_a}{V_a}(D_{s-1}-L_{s-1})\right), \; \forall s=2,\cdots,N, \\[4mm] \dfrac{\partial Q_1}{\partial D_1} = \dfrac{\partial\lambda_1}{\partial D_1}\int_{L_1}^{L_0}P(x)dx + \lambda_1\left(P(L_0)\left(1+\dfrac{V_a}{e_a}\dfrac{\partial\lambda_1}{\partial D_1}\right)-\dfrac{1}{2}P(L_1)\right)-\dfrac{e_a}{V_a}\left(\int_{L_1}^{D_1}P(x)dx - \int_{D_1}^{L_0}P(x)dx\right) \\[4mm] \qquad +\dfrac{e_a}{V_a}\left(\dfrac{1}{2}P(L_1)(D_1-L_1)-\lambda_1 P(L_0)\dfrac{V_a}{e_a}\left(1+\dfrac{V_a}{e_a}\dfrac{\partial\lambda_1}{\partial D_1}\right)\right), \\[4mm] \dfrac{\partial Q_s}{\partial D_s} = \dfrac{\partial\lambda_s}{\partial D_s}\int_{L_s}^{L_{s-1}}P(x)dx + \dfrac{1}{2}\lambda_s\left(P(L_{s-1})-P(L_s)\right)-\dfrac{e_a}{V_a}\left(\int_{L_s}^{D_s}P(x)dx - \int_{D_s}^{L_{s-1}}P(x)dx\right) \\[4mm] \qquad +\dfrac{e_a}{2V_a}\left(P(L_s)(D_s-L_s)-P(L_{s-1})(L_{s-1}-D_s)\right), \; \forall s=2,\cdots,N, \\[4mm] \dfrac{\partial Q_{s+1}}{\partial D_s} = \dfrac{1}{2}P(L_s)\left(\lambda_{s+1}-\dfrac{e_a}{V_a}(L_s-D_{s+1})\right), \; \forall s=1,2,\cdots,N-1, \end{cases}$$

where $\dfrac{\partial\lambda_s}{\partial D_s} = \begin{cases} -e_t/V_t, & \text{for the flat fare regime} \\ -\left(e_t/V_t + e_f\hat{f}\right), & \text{for the distance-based fare regime} \end{cases}, \; \forall s=1,2,\cdots,N$.

Table 4 Optimal even (average) station spacing, headway and fare solutions for different fare regimes

Flat fare regime	Distance-based fare regime

$$\begin{cases} \dfrac{\partial \overline{\pi}}{\partial \delta} = \overline{f}\sum_{s=1}^{N}\dfrac{\partial Q_s}{\partial \delta} - \left(\dfrac{2\mu_1}{HV_t}+\gamma_1\right)N = 0, \\[2mm] H = \sqrt{\dfrac{\mu_1\left(\zeta T_0 + \dfrac{2N\delta}{V_t}+2\beta_0 N\right)}{\alpha e_w \overline{f}\sum_{s=1}^{N}\int_{L_s}^{L_{s-1}}P(x)dx}}, \\[2mm] \overline{f} = \dfrac{\sum_{s=1}^{N}Q_s}{e_f\sum_{s=1}^{N}\int_{L_s}^{L_{s-1}}P(x)dx}, \end{cases}$$

$$\begin{cases} \dfrac{\partial \hat{\pi}}{\partial \delta} = \sum_{s=1}^{N}\left(\hat{f}(N+1-s)Q_s + \left(f_0 + \hat{f}\delta(N+1-s)\right)\dfrac{\partial Q_s}{\partial \delta}\right) - \left(\dfrac{2\mu_1}{HV_t}+\gamma_1\right)N = 0, \\[2mm] H = \sqrt{\dfrac{\mu_1\left(\zeta T_0 + \dfrac{2N\delta}{V_t}+2\beta_0 N\right)}{\alpha e_w \sum_{s=1}^{N}\left(f_0 + \hat{f}\delta(N+1-s)\right)\int_{L_s}^{L_{s-1}}P(x)dx}}, \\[2mm] \hat{f} = \dfrac{\sum_{s=1}^{N}(N+1-s)Q_s - f_0 e_f\sum_{s=1}^{N}(N+1-s)\int_{L_s}^{L_{s-1}}P(x)dx}{\delta e_f\sum_{s=1}^{N}(N+1-s)^2\int_{L_s}^{L_{s-1}}P(x)dx}, \end{cases}$$

where $\partial Q_s / \partial \delta$ can be given by

$$\begin{cases} \dfrac{\partial Q_1}{\partial \delta} = \dfrac{\partial \lambda_1}{\partial \delta}\int_{L_1}^{L_0}P(x)dx + \lambda_1\left(\left(N+\dfrac{V_a}{e_a}\dfrac{\partial \lambda_1}{\partial \delta}\right)P(L_0) - \left(N-\dfrac{1}{2}\right)P(L_1)\right) - \dfrac{e_a}{V_a}N\left(\int_{L_1}^{D_1}P(x)dx - \int_{D_1}^{L_0}P(x)dx\right) \\[2mm] \qquad + \dfrac{e_a}{V_a}\left(\dfrac{\delta}{2}\left(N-\dfrac{1}{2}\right)P(L_1) - \lambda_1 P(L_0)\dfrac{V_a}{e_a}\left(N+\dfrac{V_a}{e_a}\dfrac{\partial \lambda_1}{\partial \delta}\right)\right), \\[3mm] \dfrac{\partial Q_s}{\partial \delta} = \dfrac{\partial \lambda_s}{\partial \delta}\int_{L_s}^{L_{s-1}}P(x)dx + \lambda_s\left(\left(N-s+\dfrac{3}{2}\right)P(L_{s-1}) - \left(N-s+\dfrac{1}{2}\right)P(L_s)\right) - \dfrac{e_a}{V_a}(N+1-s)\left(\int_{L_s}^{D_s}P(x)dx - \int_{D_s}^{L_{s-1}}P(x)dx\right) \\[2mm] \qquad + \dfrac{\delta}{2}\dfrac{e_a}{V_a}\left(\left(N-s+\dfrac{1}{2}\right)P(L_s) - \left(N-s+\dfrac{3}{2}\right)P(L_{s-1})\right), \quad \forall s = 2,\cdots,N, \end{cases}$$

where $\dfrac{\partial \lambda_s}{\partial \delta} = \begin{cases} -(N+1-s)e_t/V_t, & \text{for the flat fare regime} \\[2mm] -(N+1-s)\left(e_t/V_t + e_f\hat{f}\right), & \text{for the distance-based fare regime} \end{cases}, \quad \forall s = 1,2,\cdots,N.$

3.2. Properties of models

In this section, we examine the properties of the models (25) and (26) that are proposed in the previous section. On the basis of the first-order optimality conditions presented in Table 3, the second-order partial derivatives of $\overline{\pi}(\cdot)$ with respect to headway H and fare $\overline{f}$ are, respectively, given by

$$\begin{cases} \dfrac{\partial^2 \overline{\pi}}{\partial H^2} = -\alpha e_w \overline{f}P(L_0)\dfrac{\partial L_0}{\partial H} - \dfrac{2\mu_1}{H^3}\left(\zeta T_0 + \dfrac{2D_1}{V_t}+2\beta_0 N\right) = (\alpha e_w)^2\,\overline{f}P(L_0)\dfrac{V_a}{e_a} - \dfrac{2\mu_1}{H^3}\left(\zeta T_0 + \dfrac{2D_1}{V_t}+2\beta_0 N\right), \\[3mm] \dfrac{\partial^2 \overline{\pi}}{\partial \overline{f}^2} = -e_f \overline{f}P(L_0)\dfrac{\partial L_0}{\partial \overline{f}} - 2e_f\sum_{s=1}^{N}\int_{L_s}^{L_{s-1}}P(x)dx = e_f^{\,2}\overline{f}P(L_0)\dfrac{V_a}{e_a} - 2e_f\sum_{s=1}^{N}\int_{L_s}^{L_{s-1}}P(x)dx, \\[3mm] \dfrac{\partial^2 \overline{\pi}}{\partial H \partial \overline{f}} = -\alpha e_w\sum_{s=1}^{N}\int_{L_s}^{L_{s-1}}P(x)dx + \alpha e_w e_f \overline{f}P(L_0)\dfrac{V_a}{e_a}. \end{cases} \qquad (27)$$

For the distance-based fare regime, the second-order partial derivatives of $\hat{\pi}(\cdot)$ with respect to H and $\hat{f}$ can be derived as follows:

$$
\begin{cases}
\dfrac{\partial^2 \hat{\pi}}{\partial H^2} = \left(\alpha e_{\mathrm{w}}\right)^2 \left(f_0 + \hat{f} D_1\right) P(L_0) \dfrac{V_{\mathrm{a}}}{e_{\mathrm{a}}} - \dfrac{2\mu_1}{H^3}\left(\zeta T_0 + \dfrac{2D_1}{V_{\mathrm{t}}} + 2\beta_0 N\right), \\[3mm]
\dfrac{\partial^2 \hat{\pi}}{\partial \hat{f}^2} = \left(e_{\mathrm{f}} D_1\right)^2 \left(f_0 + \hat{f} D_1\right) P(L_0) - 2 e_{\mathrm{f}} \displaystyle\sum_{s=1}^{N} D_s^2 \int_{L_s}^{L_{s-1}} P(x)\,dx, \\[3mm]
\dfrac{\partial^2 \hat{\pi}}{\partial H \partial \hat{f}} = -\alpha e_{\mathrm{w}} \displaystyle\sum_{s=1}^{N} D_s \int_{L_s}^{L_{s-1}} P(x)\,dx + \alpha e_{\mathrm{w}} e_{\mathrm{f}} D_1 \left(f_0 + \hat{f} D_1\right) P(L_0) \dfrac{V_{\mathrm{a}}}{e_{\mathrm{a}}}.
\end{cases} \tag{28}
$$

Eqs. (27) and (28) show that the signs of the second-order partial derivatives of $\bar{\pi}(\cdot)$ and $\hat{\pi}(\cdot)$ with respect to H and $\bar{f}$ (or $\hat{f}$) are related to the maximum service coverage L_0 of the rail transit line and the city's population density. All the second-order partial derivatives may be negative, positive, or zero. Therefore, given all other variables, the concavity of $\bar{\pi}(\cdot)$ (or $\hat{\pi}(\cdot)$) with respect to H and/or $\bar{f}$ (or H and/or $\hat{f}$) and thus the uniqueness of the optimal headway and/or fare solutions cannot be guaranteed. In addition, from Eqs. (27) and (28), we have

Proposition 3. *Given the number N and location vector $\mathbf{D}$ of stations, the profit function $\bar{\pi}(\cdot)$ (or $\hat{\pi}(\cdot)$) is concave with respect to H and $\bar{f}$ (or H and $\hat{f}$) if and only if $\dfrac{\partial^2 \bar{\pi}}{\partial H^2} < 0$, $\dfrac{\partial^2 \bar{\pi}}{\partial \bar{f}^2} < 0$ and $\dfrac{\partial^2 \bar{\pi}}{\partial H^2}\dfrac{\partial^2 \bar{\pi}}{\partial \bar{f}^2} - \left(\dfrac{\partial^2 \bar{\pi}}{\partial H \partial \bar{f}}\right)^2 > 0$ (or $\dfrac{\partial^2 \hat{\pi}}{\partial H^2} < 0$, $\dfrac{\partial^2 \hat{\pi}}{\partial \hat{f}^2} < 0$ and $\dfrac{\partial^2 \hat{\pi}}{\partial H^2}\dfrac{\partial^2 \hat{\pi}}{\partial \hat{f}^2} - \left(\dfrac{\partial^2 \hat{\pi}}{\partial H \partial \hat{f}}\right)^2 > 0$) hold simultaneously.*

However, when the population density follows a uniform distribution, i.e. $g(x) = g_0$ (or, equivalently, the potential passenger demand density $P(x) = P_0$. According to Eq. (4), both can be used alternately but not causing confusion), we have the following result.

Proposition 4. *For the flat fare regime, given the number N of stations, headway H and fare $\bar{f}$, when the population density along the rail line is uniformly distributed, the profit function $\bar{\pi}(\cdot)$ is concave with regard to the rail line length D_1 and station locations $D_2, D_3, \ldots, D_N$.*

The proof of Proposition 4 is provided in Appendix B. Proposition 4 indicates that for the flat fare regime and a uniform population distribution, when other variables are given, the optimal solutions for the rail line length and the station location are unique. However, this property is not satisfied for the distance-based fare regime. For illustration purposes, here is an example.

Example 1. Consider a rail line that extends from the city's CBD outwards. There are three stations: one is located at the CBD area (i.e. $D_3 = 0$), and the locations of other two stations (i.e. D_1 and D_2) are unknown and need to be determined. Assume that the rail fare is a distance-based one and is given by $f_s = \hat{f} D_s$, $s = 1, 2$. In the following, we show that for a given train headway H, a unit-distance fare $\hat{f}$, and a uniform population density (i.e. $P(x) = P_0$), the profit function $\hat{\pi}(\cdot)$, which is given by Eq. (26), may be non-concave with respect to D_1 and D_2. To do so, we need to check the negative definiteness of the following Hessian matrix:

$$H_2(\hat{\pi}) = \left(\frac{\partial^2 \hat{\pi}}{\partial D_i \partial D_j}\right) = \begin{pmatrix} \dfrac{\partial^2 \hat{\pi}}{\partial D_1^{\;2}} & \dfrac{\partial^2 \hat{\pi}}{\partial D_1 \partial D_2} \\[2mm] \dfrac{\partial^2 \hat{\pi}}{\partial D_2 \partial D_1} & \dfrac{\partial^2 \hat{\pi}}{\partial D_2^{\;2}} \end{pmatrix}. \tag{29}$$

According to Eqs. (12)-(14) and $P(x) = P_0$, we have

$$Q_s = \lambda_s P_0 \left(L_{s-1} - L_s\right) - \frac{e_a P_0}{2V_a}\left(\left(D_s - L_s\right)^2 + \left(L_{s-1} - D_s\right)^2\right), \quad \forall s = 1,2, \tag{30}$$

where $\lambda_s = 1 - e_w \alpha H - e_t \beta_0 (3-s) - \left(\dfrac{e_t}{V_t} + e_f \hat{f}\right) D_s$, $s = 1,2$, and $L_s, s = 0,1,2$ are given by Eqs. (1) and (14).

The second-order partial derivatives of $\hat{\pi}(\cdot)$ with respect to D_s, $s = 1,2$ can then be derived as follows:

$$\begin{cases} \dfrac{\partial^2 \hat{\pi}}{\partial D_1^{\;2}} = -\hat{f}P_0\left(\left(\dfrac{e_t}{V_t} + e_f \hat{f}\right)\left(2\lambda_1 \dfrac{V_a}{e_a} + 2D_1 - D_2\right) + \dfrac{1}{4}\dfrac{e_a}{V_a}(3D_1 - D_2) - D_1 \dfrac{V_a}{e_a}\left(\dfrac{e_t}{V_t} + e_f \hat{f}\right)^2 - \lambda_1\right), \\[3mm] \dfrac{\partial^2 \hat{\pi}}{\partial D_1 \partial D_2} = \hat{f}P_0\left(\dfrac{1}{2}\left(\dfrac{e_t}{V_t} + e_f \hat{f}\right)(D_1 - D_2) + \dfrac{1}{4}\dfrac{e_a}{V_a}(D_1 + D_2) + \dfrac{1}{2}(\lambda_2 - \lambda_1)\right), \\[3mm] \dfrac{\partial^2 \hat{\pi}}{\partial D_2^{\;2}} = -\hat{f}P_0\left(\left(\dfrac{e_t}{V_t} + e_f \hat{f}\right)D_1 - \dfrac{e_a}{V_a}\left(\dfrac{1}{4}D_1 - \dfrac{3}{2}D_2\right)\right). \end{cases} \tag{31}$$

Let $\dfrac{\partial^2 \hat{\pi}}{\partial D_2^{\;2}} = 0$, one then obtains

$$D_2 = \left(\frac{1}{6} - \frac{2}{3}\frac{V_a}{e_a}\left(\frac{e_t}{V_t} + e_f \hat{f}\right)\right)D_1. \tag{32}$$

As a result, the second-order leading principle minor of $H_2(\hat{\pi})$ in Eq. (29) is always less than zero, i.e.

$$\det\left(H_2(\hat{\pi})\right) = -\left(\frac{\partial^2 \hat{\pi}}{\partial D_1 \partial D_2}\right)^2 < 0. \tag{33}$$

This implies that there is at least one feasible station location solution such that $(-1)^s \det\left(H_s(\hat{\pi})\right) > 0$ (a sufficient and necessary condition that a symmetric matrix is negative definite, see Strang, 2006) is not satisfied. Hence, $\hat{\pi}(\cdot)$ may not be concave with respect to the station location vector **D** even for a uniform population density. Thus, given other variables, the uniqueness of the optimal rail line length and station location solutions cannot be guaranteed.

However, the following proposition shows that the optimal even (or average) station spacing solutions for the both fare regimes are unique. The proof of this proposition is provided in Appendix C.

Proposition 5. *Given the number N of stations, headway H and fare $\bar{f}$ (or $\hat{f}$), when the population density along the rail line is uniformly distributed, the optimal even (or average) station spacing solutions for both the flat and distance-based fare regimes are unique and given by Eqs. (C.5) and (C.13), respectively.*

It should be pointed out that the concavity of the objective function $\bar{\pi}(\cdot)$ or $\hat{\pi}(\cdot)$ and thus the uniqueness of the model solutions cannot be guaranteed when the rail line length, station location (or spacing), headway and fare are jointly optimized because the negative definiteness of the resultant Hessian matrix cannot be ensured.

3.3. Comparison of fare regimes

In the previous section, we have discussed the solution properties for two profit maximization models with different fare regimes. It has been shown that the flat and distance-based fare regimes can lead to significant differences in the solution properties of the models. In this section, we further compare the net profits resulting from any two fare levels (not necessarily optimal solutions) for the two fare regimes, and identify the condition under which the two fare regimes are indifferent in terms of the net profit of the operator.

For fair comparison, the rail line configuration (the rail line length and the number and locations of stations) and train headway are assumed to be identical for the two regimes. Let $\hat{Q}_s$ and $\bar{Q}_s$ represent the resultant passenger demand for station s by the flat and distance-based fare regimes, respectively. From Eqs. (25) and (26), the difference in the net profits for the two regimes is given by

$$\hat{\pi} - \bar{\pi} = \sum_{s=1}^{N} \left(f_0 + \hat{f} D_s \right) \hat{Q}_s - \bar{f} \sum_{s=1}^{N} \bar{Q}_s \,, \tag{34}$$

where $\hat{Q}_s$ and $\bar{Q}_s$ can be determined by Eqs. (12)-(14), respectively.

In order to gain some preliminary and valuable insights, we consider a special case with a uniform population density (i.e., $P(x) = P_0$) and an even (or average) station spacing δ, which leads to the passenger demand pattern as shown in Eqs. (C.1) and (C.2). Substituting Eqs. (C.1) and (C.2) into Eq. (34) yields

$$\begin{aligned}
\hat{\pi} - \bar{\pi} &= \sum_{s=1}^{N} \left(f_0 + \hat{f} D_s \right) \hat{Q}_s - \bar{f} \sum_{s=1}^{N} \bar{Q}_s \\
&= \left(f_0 + \hat{f} \delta N \right) \left(\frac{P_0}{2} \frac{V_a}{e_a} \hat{\lambda}_1^{\,2} + \frac{P_0}{2} \hat{\lambda}_1 \delta - \frac{P_0}{8} \frac{e_a}{V_a} \delta^2 \right) + \sum_{s=2}^{N} \left(f_0 + \hat{f} \delta (N+1-s) \right) \left(P_0 \hat{\lambda}_s \delta - \frac{P_0}{4} \frac{e_a}{V_a} \delta^2 \right) \\
&\quad - \bar{f} \left(\frac{P_0}{2} \frac{V_a}{e_a} \bar{\lambda}_1^{\,2} + \frac{P_0}{2} \bar{\lambda}_1 \delta - \frac{P_0}{8} \frac{e_a}{V_a} \delta^2 \right) - \bar{f} \sum_{s=2}^{N} \left(P_0 \bar{\lambda}_s \delta - \frac{P_0}{4} \frac{e_a}{V_a} \delta^2 \right),
\end{aligned} \tag{35}$$

where

$$\hat{\lambda}_s = 1 - e_w \alpha H - e_f f_0 - (N+1-s) \left(\left(\frac{e_t}{V_t} + e_f \hat{f} \right) \delta + e_t \beta_0 \right), \ \forall s = 1, 2, \cdots, N \,, \text{ and} \tag{36}$$

$$\bar{\lambda}_s = 1 - e_w \alpha H - e_f \bar{f} - (N+1-s) \left(\frac{e_t}{V_t} \delta + e_t \beta_0 \right), \ \forall s = 1, 2, \cdots, N \,. \tag{37}$$

To identify the sign of ($\hat{\pi} - \bar{\pi}$), we set Eq. (35) equal to zero and solve it. We then obtain

Proposition 6. *Given the number N of stations, train headway H, and fares $\bar{f}$ and $\hat{f}$, the net profits for the flat and distance-based fare regimes are indifferent if and only if the (even or average) station spacing δ satisfies the following cubic equation*

$$\delta^3 + a_1 \delta^2 + a_2 \delta + a_3 = 0 \,, \tag{38}$$

where the coefficients $a_i, i = 1, 2, 3$ are given by Eqs. (D.1)-(D.5) in Appendix D, respectively.

For presentation purposes, the station spacing that satisfies Eq. (38) is referred to as "*indifference station spacing*" and denoted as δ^*. The corresponding rail line length and net profit are referred to as "*indifference rail line length*" and "*indifference net profit*", respectively. According to Eqs. (D.1)-(D.5), the indifference station spacing

δ^* is independent of the population density P_0. The solution expressions for Eq. (38) are shown in entry (ii) of Appendix D. For more details, the reader can refer to Spiegel et al. (2009, Page 13).

For illustrating the application of Proposition 6, an example is provided as below.

Example 2. In this example, it is assumed that the uniform population density is 34,000 persons/km^2, which is the average population density of Hong Kong. The number of stations is fixed as $N = 15$ and the train headway is 3.0 minutes. The fixed and variable components of the distance-based fare are assumed to be \$0.25 and \$0.15/km, respectively. The values of other input parameters are identical with those shown in Table 2. In the following, the indifference solutions between the flat fares of \$1.30, \$1.48 and \$1.60 and the above specified distance-based fare are presented, respectively.

By Eqs. (38) and (D.1)-(D.5), one can obtain the indifference station spacing solutions between the three flat fares and the specified distance-based fare, as shown in Table 5. It can be seen that there are two indifference station spacing solutions between the flat fare of \$1.30 and the given distance-based fare, i.e. $\delta^* = 1.95$ or 1.03 km (another negative root -0.12 is meaningless and thus discarded). They are associated with the indifference rail line lengths of 29.25 and 15.45 km, respectively, which result in the indifference net profits of \$27,358, and \$19,246 per hour, respectively. When the flat fare is increased to \$1.48, only one indifference station spacing solution exists, i.e. $\delta^* = 1.52$ km. The resultant indifference rail line length and indifference net profit are 22.80 km and \$31,281 per hour, respectively. When the flat fare is further increased to \$1.60, no indifference station spacing solution exists.

Table 5 The indifference solutions for three flat fares and a specified distance-based fare

Flat fare (\$)	Indifference station spacing δ^* (km)	Indifference rail line length D_1^* (km)	Indifference net profit π^* (\$/h)
1.30	(1.95; 1.03; -0.12)	(29.25; 15.45; ×)	(27,358; 19,246; ×)
1.48	1.52	22.80	31,281
1.60	×	×	×

Note: "×" means that no appropriate (real) solution exists at the corresponding item.
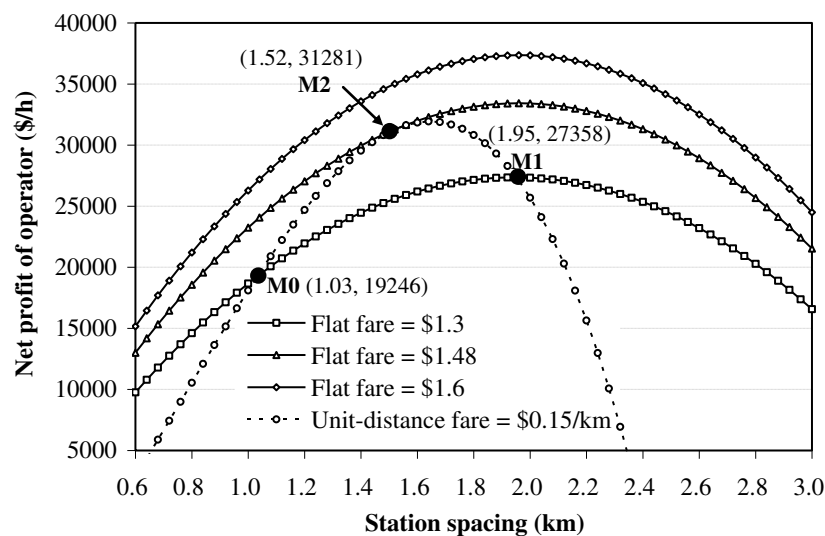

Fig. 2. Indifference station spacings between three flat fares and a distance-based fare.

In order to verify the correctness of the indifference solutions that are generated by Eq. (38), a graphical analysis approach is used to plot the profit curves for the above three flat fares and the above specified distance-based fare when the station spacing is changed from 0.6 to 3.0 km, as shown in Fig. 2. It can be seen in this figure that the intersections M0, M1 and M2 between these profit curves are really consistent with the indifference solutions that are obtained from Eq. (38).

In addition, it can also be seen that, in contrast to the given distance-based fare, a flat fare of higher than \$1.48 can always lead to higher net profit and thus is a better option. However, when the flat fare is lower than \$1.48, there is some station spacing such that the distance-based fare is better than the flat fare in terms of the net profit of the operator. For example, when the station spacing is in between 1.03 and 1.95 km, the specified distance-based fare can yield higher profit than the flat fare of \$1.30.

3.4. Constraints

Thus far, the models proposed in the previous section have not taken into account the effects of constraint conditions. To make the models more realistic, in the following the capacity constraint and rail line length constraint are presented, respectively. The capacity constraint ensures that the supply of the rail transit service satisfies the passenger demand, i.e.

$$\sum_{s=1}^{N} Q_s \leq \frac{K}{H}, \tag{39}$$

where K is the capacity of vehicles (i.e. the maximum number of passengers allowed in a vehicle, both seated and standing).

The capacity constraint (39) can further be represented as a bound constraint as below

$$H \leq H_{\max}, \tag{40}$$

where $H_{\max} = K \big/ \sum_{s=1}^{N} Q_s$.

On the other hand, the rail line length designed should not exceed the corridor length, i.e.

$$D_1 \leq B. \tag{41}$$

Particularly, when addressing the even or average station spacing δ, constraint (41) can further be written as

$$\delta \leq \frac{B}{N}. \tag{42}$$

Eq. (42) is actually a bound constraint on the station spacing.

In order to ensure that the capacity and rail line length constraints of the rail transit service are satisfied, the optimized values of the decision variables, such as the train headway, rail line length and station spacing, should be verified with Eqs. (40)-(42).

4. Solution algorithm

In this section, a heuristic solution algorithm is developed to solve the proposed models (25) and (26) with bound constraints (40)-(42). The solution algorithm developed below is directly based on the first-order optimality conditions of the proposed models (see Table 3 or Table 4). The step-by-step procedure is given as follows.

Step 1. *Initialization.* Choose an initial value for each of the design variables of the rail line, including rail line length $D_1^{(0)}$, station locations $D_s^{(0)}$ ($s = 2, \cdots, N$), headway $H^{(0)}$ and fare $\bar{f}^{(0)}$ or $\hat{f}^{(0)}$. Determine the corresponding passenger demand for each station on the rail line, $Q_s^{(0)}$ ($s = 1, 2, \cdots, N$) by Eqs. (12)-(14) and the corresponding value of the objective function (25) or (26). Set iteration counter $j = 1$.

Step 2. *Updating of the design variables.* Sequentially update the values of the headway, fare, rail line length and station location (or spacing) according to the first-order optimality conditions given in Table 3 or Table 4.

Step 2.1. Update $H^{(j)}$ with fixed $D_s^{(j-1)}$ ($s = 1, 2, \cdots, N$) and $\bar{f}^{(j-1)}$ or $\hat{f}^{(j-1)}$. Check whether the resultant headway $H^{(j)}$ satisfies the capacity constraint (40) and the non-negative passenger demand constraint (6). If it exceeds some constraint bound, then it is set at the corresponding bound.

Step 2.2. Update $\bar{f}^{(j)}$ or $\hat{f}^{(j)}$ with fixed $D_s^{(j-1)}$ ($s = 1, 2, \cdots, N$) and $H^{(j)}$. Check the non-negative passenger demand constraint (6) for the resultant fare $\bar{f}^{(j)}$ or $\hat{f}^{(j)}$. If it exceeds the constraint bound, then it is set at the corresponding bound.

Step 2.3. Update $D_s^{(j)}$ ($s = 1, 2, \cdots, N$) with fixed $H^{(j)}$ and $\bar{f}^{(j)}$ or $\hat{f}^{(j)}$. Check the rail line length constraint (41) or station spacing constraint (42) and the non-negative passenger demand constraint (6) for the resultant $D_s^{(j)}$ ($s = 1, 2, \cdots, N$). If it exceeds some constraint bound, then it is set at the corresponding bound.

Step 3. *Updating of the passenger demand and objective function.* Update the passenger demand for each station on the rail line, $Q_s^{(j)}$ ($s = 1, 2, \cdots, N$) by Eqs. (12)-(14), and the resultant value of the objective function (25) or (26).

Step 4. *Termination check.* If the resultant objective function values for successive iterations are sufficiently close, then terminate the algorithm and output the optimal solution $\left\{ \mathbf{D}^*, H^*, \bar{f}^* \text{ or } \hat{f}^* \right\}$ and the corresponding objective function value $\bar{\pi}^*$ or $\hat{\pi}^*$. Meanwhile, the optimal fleet size F^* can also be obtained with Eq. (18). Otherwise, set $j = j + 1$, and go to Step 2.

It should be mentioned that the above solution procedure is based on a fixed number of stations. Note that the number of stations is an integer variable, which makes it difficult to solve the resultant mixed integer programming problem. Fortunately, the number of stations on a rail transit line is a finite number. Therefore, a simple approach for finding the optimal number of stations is to compare the resultant objective function values with different numbers of stations.

In Step 2, the values of the design variables – the train headway, fare, and the station location/spacing – are sequentially updated one at a time while holding the values of other variables fixed. The associated constraints should be immediately checked after the updating of each decision variable such that the resultant solutions at each iteration of the solution process always satisfy all the constraints.

The equations with regard to the station location/spacing $D_s^{(j)}$ and the fare $\bar{f}^{(j)}$ (or $\hat{f}^{(j)}$) contain the passenger demand variable $Q_s^{(j)}$ ($s = 1, 2, \cdots, N$), which are a function of $D_s^{(j)}$ and $\bar{f}^{(j)}$ (or $\hat{f}^{(j)}$), respectively. Thereby, the solving of $D_s^{(j)}$ and the fare $\bar{f}^{(j)}$ (or $\hat{f}^{(j)}$) in Step 2 is equivalent to solving a fixed-point problem regarding that decision variable itself. This can be easily implemented by using the bisection method or Newton's method (Epperson, 2007, Chapter 3). In this paper, the bisection method is adopted. It should be pointed out that when the objective functions are not concave with regard to some decision variable, both the bisection method and the Newton's method may terminate at some local optimum.

5. Numerical studies

In this section, two test scenarios are used to illustrate the application of the proposed models and solution algorithm and the contributions of this paper. The first scenario aims to show the effects of the population density

and the rail capital cost on the rail line design. The second scenario is used to evaluate the effects of urban forms in terms of the population distribution and the corridor length. The alignment of the rail line concerned is shown in Fig. 1. In the following analyses, unless specifically stated otherwise, the length of the corridor is fixed as 30 km, the fixed component of the distance-based fare is $1.5, and the baseline values for other input parameters are the same with those as shown in Table 2.

5.1. Scenario 1

We first investigate the effects of the population density and the rail capital cost on the net profit of the operator. In reality, the rail capital (or fixed) cost usually changes over time and space dimensions and is thus uncertain. It is, thus, necessary for a revenue-driven investor to ascertain the minimum (average) population density requirement for different rail capital costs such that the investment project is profitable. To do so, we conduct numerical experiments by changing the population density from 4,000 to 36,000 persons per square kilometer and scaling the baseline values of the rail capital cost (i.e. the model parameters μ_0, γ_0 and Λ_0) by multiplied from 0.5 to 4.0.
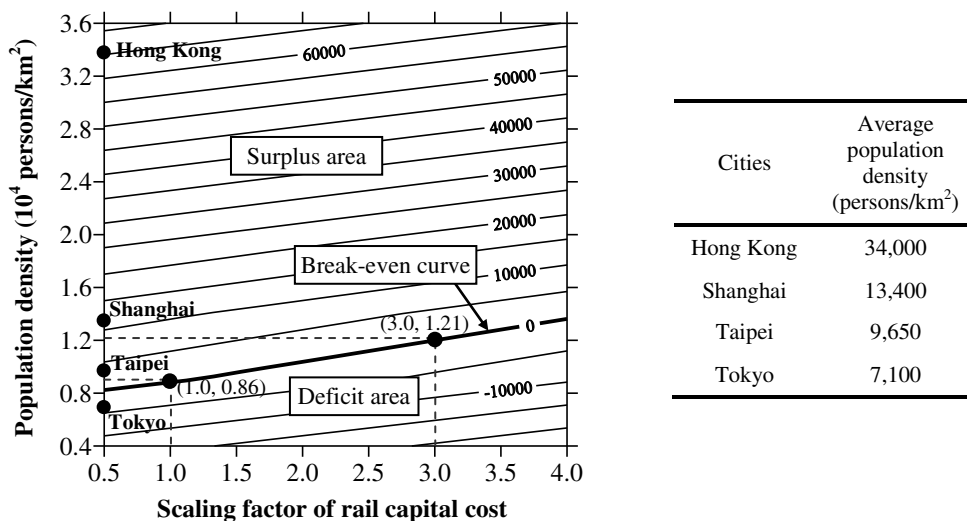

Fig. 3. Change in net profit with population density and rail capital cost.

Fig. 3 shows the change in the optimized net profit for various combinations of the population density and the rail capital cost for the flat fare regime. It can be observed that different combinations of the population density and the rail capital cost can lead to three possible outcomes: surplus, break-even, or deficit. As the rail capital cost increases, the minimum population density required for making the rail project financially viable increases. For instance, at the level of the baseline value of the rail capital cost, the average population density to ensure the profitability of the rail line operations must exceed 8,600 persons per square kilometer. When the rail capital cost is 3.0 times of the baseline value, the minimum viable population density reaches 12,100 persons per square kilometer. Fig. 3 also shows the profitability of the rail line operations for different rail capital costs for four cities with different average population densities. It can be seen that, as the rail capital cost changes from 0.5 time to 4.0 times of the baseline value, the rail transit services in Hong Kong can always operate profitably and those in Tokyo would require direct government subsidies. When the rail capital cost reaches 1.5 times of the baseline value, the rail operations in Taipei become break-even. For the Shanghai's rail transit services, a deficit occurs at 4.0 times of the baseline value of the rail capital cost.

We now look at the effects of population density on the optimal design of the rail transit line in terms of the net profit of the operator. We take Hong Kong and Taipei as examples. Their average population densities are 34,000 and 9,650 persons per square kilometer respectively, as shown in Fig. 3. Fig. 4 shows the changes in the optimized

net profit with different numbers of stations and different fare regimes for the two cities, respectively. It can be seen that, for all cases, as the number of stations increases, the resultant net profit of the operator first increases and then decreases. The optimal numbers of stations for the flat and distance-based fare regimes are, respectively, 19 and 15 for the Hong Kong case (shown in Fig. 4a), and 8 and 6 for the Taipei case (shown in Fig. 4b). This means that the higher the urban population density, the bigger the number of stations.
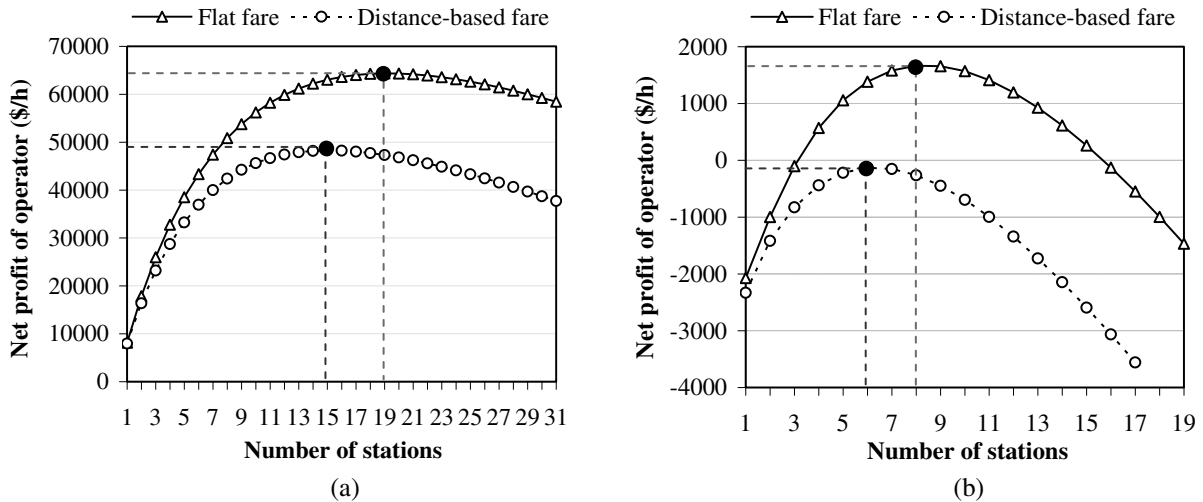

Fig. 4. Net profit of operator against number of stations: (a) Hong Kong case; (b) Taipei case.

Table 6 Optimal solutions with different average population densities

Optimal solution	Hong Kong case		Taipei case	
	Flat fare	Distance-based fare	Flat fare	Distance-based fare
Number of stations	**19**	**15**	**8**	**6**
Rail line length (km)	27.05	24.00	13.67	11.05
Average station spacing (km)	1.42	1.60	1.71	1.84
Fare*	3.46	0.26	3.61	0.36
Headway (h)	0.06	0.05	0.14	0.14
Fleet size (no. of vehicles)	31	30	7	6
Total passenger demand (pass/h)	30,478	33,180	4,791	4,557
Net profit ($/h)	**64,346**	**48,273**	**1,665**	**-132**

* The flat fare and the distance-based fare are measured in $ and $/km, respectively.

Table 6 displays the optimal solutions for the design variables of the rail line with two different average population densities. It can be seen that, for a given fare regime, a city with a higher population density requires a longer rail line, a higher station density (i.e. a shorter average station spacing), a lower fare, a smaller headway and a larger fleet size, and vice versa. In addition, for a given population density, in contrast to the distance-based fare regime, the flat fare regime can lead to a higher net profit, which needs an investment of a longer rail line, a larger fleet size, and a higher station density. In particular, for a low-density city the distance-based fare regime can induce a negative profit (-$132/h).

5.2. Scenario 2

To explore the effects of the density gradient in the population distribution function (see A4 or Eq. (4)), Fig. 5 shows different population distributions with the same number of population ($G = 1,020,000$) and the same corridor length ($B = 30$ km) for three different density gradients: $\theta = 0$, 0.05 and 0.1. It can be observed in this figure that a smaller θ-value implies a higher level of dispersion in population distribution along the corridor, whereas a larger θ-value indicates a more compact city. Particularly, when θ equals 0, the inhabitants are uniformly distributed along the corridor.
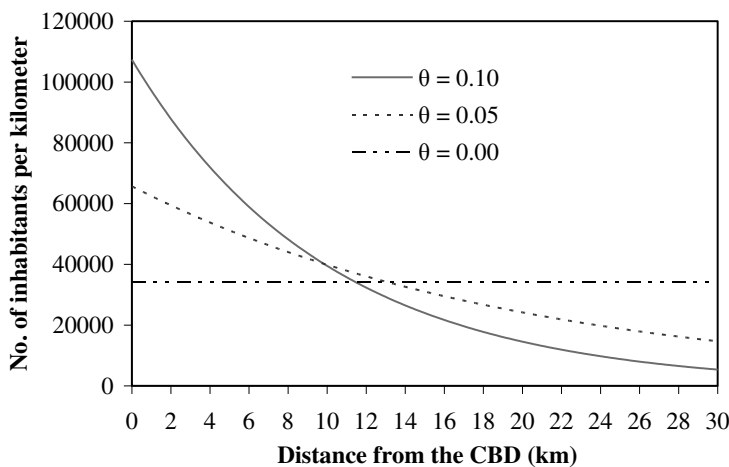

Fig. 5. Urban forms with different values of density gradient θ.

Table 7 Optimal solutions with different density gradients

Optimal solution	$\theta = 0.05$		$\theta = 0.1$	
	Flat fare	Distance-based fare	Flat fare	Distance-based fare
Number of stations	19	14	19	14
Rail line length (km)	20.59	17.17	16.31	13.40
Average station spacing (km)	1.08	1.23	0.86	0.96
Fare[*]	3.86	0.21	4.10	0.31
Headway (h)	0.06	0.05	0.05	0.05
Fleet size (no. of vehicles)	26	24	24	22
Total passenger demand (pass/h)	31,145	34,511	33,636	38,284
Net profit ($/h)	85,487	65,008	106,136	81,464

* The flat fare and the distance-based fare are measured in $ and $/km, respectively.

Table 7 indicates the optimal solutions with different values of density gradient θ for the flat and distance-based fare regimes (the results with $\theta = 0$ are associated with the Hong Kong case as shown in Table 6). It is noted that, for a given fare regime, as the density gradient θ increases (i.e. the city becomes more compact, see Fig. 5), the optimal rail line length, average station spacing and fleet size decrease, the optimal headway almost remains unchanged, and the optimal fare increases. As a result, the total passenger demand and the associated net profit rise, respectively. In addition, for a given value of the density gradient θ, compared to the distance-based fare regime,

although the flat fare regime has a lower attractiveness for passengers, it can still create a higher net profit at the cost of a longer rail line, a larger fleet size, and a higher station density (i.e. a shorter average station spacing).

We now examine the effects of the urban forms on the rail line design by changing the value of the density gradient from 0 to 0.2 and the value of the corridor length from 9 to 14 km. For consistent comparison, the total number of population is fixed as 1,020,000. In the following, only the flat fare regime is taken as an example because the distance-based fare regime can yield the similar conclusion. Fig. 6 plots the change in the optimized net profit for different corridor lengths and different density gradients. It can be seen that for a given density gradient, as the corridor length decreases, the net profit of the operator increases, and vice versa. This means that a high-density and small-scale city is more profitable than a low-density and large-scale city from the perspective of a revenue-driven operator.
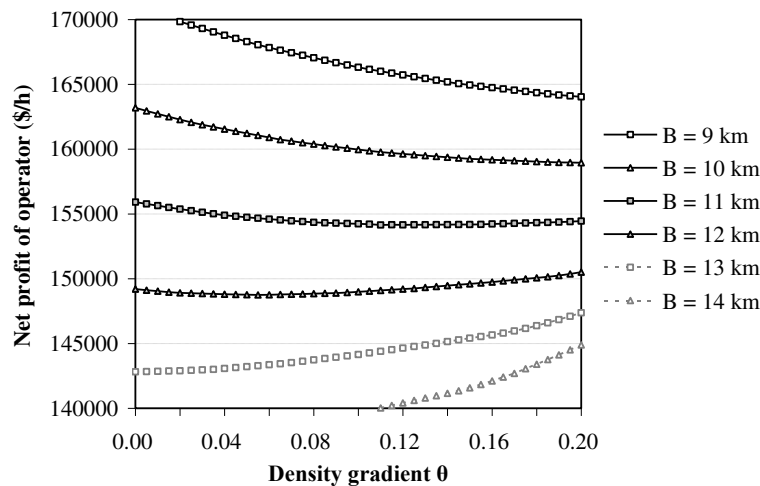

Fig. 6. Change in net profit of operator with density gradient and corridor length.

However, for a given corridor length, the change in the net profit of the operator with the density gradient exhibits diverse tendencies. Specifically, for an urban corridor with a length of less than 10 km, as the density gradient θ increases from 0 to 0.2, the net profit of the operator always descends. The maximum net profit occurs at the case of $\theta = 0$, which is associated with a uniform population distribution. This means that for a high-density and small-scale city, a decentralized urban form is more profitable for a benefit-driven operator than a compact urban form. However, for an urban corridor with a length of larger than 13 km, the net profit of the operator always ascends as the density gradient θ increases. This implies that for a low-density and large-scale city, a compact urban form is more profitable than a decentralized urban form. When the corridor length is around 11-12 km, as the density gradient θ changes from 0 to 0.2, the net profit of the operator first decreases and then increases. Therefore, there are two different density gradients, which are respectively associated with a compact urban form and a diffuse urban form, such that they can achieve the same profitability for profit-maximizing transit services.

6. Conclusions and further studies

In this paper, analytical models were proposed for optimizing the design variables of a rail transit line in a linear urban transportation corridor. The rail line length, number and locations of stations, headway and fare were optimized simultaneously. The effects of passenger demand elasticity, fare regimes and urban population distribution were explicitly considered in the proposed models. Two profit maximization models, based on the flat and distance-based fare regimes, were presented. The first-order optimality conditions for the two proposed models were derived (see Propositions 1 and 2) and their solution properties were investigated and compared (see

Propositions 3-6). A heuristic solution algorithm for jointly determining the optimal design variables was developed. The applications of the proposed models to the comparison of fare regimes and to the evaluation of rail capital cost and urban structure have contributed to some new insights and important findings. It has been shown that the fare regimes, rail capital cost, urban population density, density gradient and city's length have significant effects on the design of the rail line and/or the profitability of the rail transit operations. The proposed models can serve as a useful tool for long-term strategic planning of rail transit services and urban development and for evaluation of various rail transit and land use policies.

Although it has been shown that the models developed in this paper have well-defined properties, some important features of transit services were omitted and should be considered in future studies. Firstly, our models did not consider the effects of passenger crowding within train carriages and at railway stations. Previous studies have shown that crowding discomfort has an important effect on passenger's choice of transit service (Huang, 2000; Li et al., 2009). Therefore, it will be useful to relax this assumption in future studies particularly for congested transit networks in Asia. Secondly, our models mainly focused on a many-to-one travel demand pattern during commuting period. In reality, individual trips take place at various origins and destinations. Thereby, there is a need to extend the proposed models to consider a many-to-many travel demand pattern (Wirasinghe and Ghoneim, 1981; Liu et al., 1996; Chien and Schonfeld, 1997; Wirasinghe et al., 2002). Thirdly, our models also assumed that passengers would get on the train at the nearest station. However, in reality, some passengers may have a preference for the upstream station particularly during the peak periods. This is because, at the upstream station, there is a greater possibility of obtaining a seat and/or a higher chance of getting on the train, and thus the risk of not being able to board the train is reduced (Sumalee et al., 2009). An investigation of the choice of upstream and downstream stations is interesting and important but outside the scope of this paper and thus left to future research. Fourthly, while the models developed in this paper focused on the transit operator's interest, namely, profit maximization, an extension of the proposed models to consider the user's perspective (i.e., minimization of total user cost) or society's perspective (i.e., maximization of social welfare) could lead toward more comprehensive policy analysis.

Acknowledgements

References

Anas, A. (1982). *Residential location markets and urban transportation*. Academic Press, New York, USA.

Chang, S. K., & Schonfeld P. M. (1991). Multiple period optimization of bus transit systems. *Transportation Research B*, 25 (6), 453-478.

Chang, S. K., & Schonfeld, P. M. (1993). Optimal dimensions of bus service zones. *Journal of Transportation Engineering-ASCE*, 119 (4), 567-585.

Chien, S., & Schonfeld, P. M. (1997). Optimization of grid transit system in heterogeneous urban environment. *Journal of Transportation Engineering-ASCE*, 123 (1), 28-35.

Chien, S., & Schonfeld, P. M. (1998). Joint optimization of a rail transit line and its feeder bus system. *Journal of Advanced Transportation*, 32 (3), 253-284.

Chien, S., & Qin, Z. (2004). Optimization of bus stop locations for improving transit accessibility. *Transportation Planning and Technology*, 27 (3), 211-227.

Epperson, J. F. (2007). *An introduction to numerical methods and analysis*. John Wiley & Sons, NJ, USA.

Furth, P., & Rahbee, A. (2000). Optimal bus stop spacing through dynamic programming and geographic modeling. *Transportation Research Record*, 1731, 15-22.

Huang, H. J. (2000). Fares and tolls in a competitive system with transit and highway: The case with two groups of commuters. *Transportation Research E*, 36 (4), 267-284.

Hurdle, V. F. (1973). Minimum cost locations for parallel public transit lines. *Transportation Science*, 7 (4), 340-350.

Hurdle, V. F., & Wirasinghe, S. C. (1980). Location of rail stations for many to one travel demand and several feeder modes. *Journal of Advanced Transportation*, 14 (1), 24-45.

Kocur, G., & Hendrickson, C. (1982). Design of local bus service with demand equilibrium. *Transportation Science*, 16 (2), 149-170.

Kuah, G. K., & Perl, J. (1988). Optimization of feeder bus routes and bus stop spacing. *Journal of Transportation Engineering-ASCE*, 114 (3), 341-354.

Lam, W. H. K., Cheung, C. Y., & Poon, Y. F. (1998). A study of train dwelling time at the Hong Kong mass transit railway system. *Journal of Advanced Transportation*, 32 (3), 285-296.

Lam, W. H. K., & Zhou, J. (2000). Optimal fare structure for transit networks with elastic demand. *Transportation Research Record*, 1733, 8-14.

Li, Z. C., Lam, W. H. K., & Wong, S. C. (2009). The optimal transit fare structure under different market regimes with uncertainty in the network. *Networks and Spatial Economics*, 9 (2), 191-216.

Liu, G., Quain, G. J., & Wirasinghe, S. C. (1996). Rail line length in a cross-town corridor with many to many demand. *Journal of Advanced Transportation*, 30 (1), 95-114.

Liu, T. L., Huang, H. J., Yang, H., & Zhang, X. (2009). Continuum modeling of park-and-ride services in a linear monocentric city with deterministic mode choice. *Transportation Research B*, 43 (6), 692-707.

O'Sullivan, A. (2000). *Urban economics*. Irwin/McGraw-Hill, Boston, USA.

Saka, A. A. (2001). Model for determining optimum bus-stop spacing in urban areas. *Journal of Transportation Engineering-ASCE*, 127 (3), 195-199.

Spasovic, L. N., & Schonfeld, P. M. (1993). A method for optimizing transit service coverage. *Transportation Research Record*, 1402, 28-39.

Spasovic, L., Boile, M. P., & Bladikas, A. (1994). Bus transit service coverage for maximum profit and social welfare. *Transportation Research Record*, 1451, 12-22.

Spiegel, M. R., Lipschutz, S., & Liu, J. (2009). *Mathematical handbook of formulas and tables*. McGraw-Hill, New York, USA.

Strang, G. (2006). *Linear algebra and its applications*. Thomson, Brooks/Cole, Belmont, CA, USA.

Sumalee, A., Tan, Z. J., & Lam, W. H. K. (2009). Dynamic stochastic transit assignment with explicit seat allocation model. *Transportation Research B*, 43 (8-9), 895-912.

Transport Department (2003). *Travel Characteristics Survey 2002 – Final Report, Hong Kong*. Transport Department of Hong Kong.

Vuchic, V. R., & Newell, G. F. (1968). Rapid transit interstation spacing for minimum travel time. *Transportation Science*, 2 (4), 303-339.

Vuchic, V. R. (1969). Rapid transit interstation spacing for maximum number of passengers. *Transportation Science*, 3 (3), 214-232.

Vuchic, V. R. (2005). *Urban transit: operations, planning and economics*. John Wiley & Sons, Hoboken, NJ, USA.

Wang, J. Y. T., Yang, H., & Lindsey, R. (2004). Locating and pricing park-and-ride facilities in a linear monocentric city with deterministic mode choice. *Transportation Research B*, 38 (8), 709-731.

Wirasinghe, S. C., & Ghoneim, N. S. A. (1981). Spacing of bus-stops for many to many travel demand. *Transportation Science*, 15 (3), 210-221.

Wirasinghe, S. C., & Seneviratne, P. N. (1986). Rail line length in an urban transportation corridor. *Transportation Science*, 20 (4), 237-245.

Wirasinghe, S. C., Quain, G. J., & Bandara, J. M. S. J. (2002). Optimal terminus location for a rail line with many to many travel demand. *Proceedings of the 15th International Symposium on Transportation and Traffic Theory* (Taylor, M. A. P., ed.), pp. 75-97. Elsevier, Oxford.

Zhou, J., Lam, W. H. K., & Heydecker, B. G. (2005). The generalized Nash equilibrium model for oligopolistic transit market with elastic demand. *Transportation Research B*, 39 (6), 519-544.

Appendix A. Proof of Proposition 1

We first look at the flat fare regime. To obtain the optimal solutions for the rail line length and station location, we set the partial derivative of the objective function $\bar{\pi}(\cdot)$ with respect to D_s to zero. Then, we have

$$\frac{\partial \bar{\pi}}{\partial D_s} = \bar{f} \sum_{i=1}^{N} \frac{\partial Q_i}{\partial D_s} - \Delta_s \left(\frac{2\mu_1}{HV_t} + \gamma_1 \right) = 0, \ \forall s = 1, 2, \cdots, N, \tag{A.1}$$

where $\Delta_s = 1$ if $s = 1$, and 0 otherwise.

According to Eqs. (12)-(14), Q_s is a function of D_s, λ_s, L_{s-1}, and L_s, which are functions of D_{s-1}, D_s, and D_{s+1} in terms of Eqs. (1)-(11); i.e.,

$$Q_s = Q_s \left(D_{s-1}, D_s, D_{s+1} \right), \ \ \forall s = 1, 2, \cdots, N. \tag{A.2}$$

Hence, the following equation holds

$$\frac{\partial Q_i}{\partial D_s} = 0, \ \ \forall i \neq s-1, s, s+1. \tag{A.3}$$

Substituting Eq. (A.3) into Eq. (A.1), one immediately obtains

$$\bar{f} \sum_{i=s-1}^{s+1} \frac{\partial Q_i}{\partial D_s} - \Delta_s \left(\frac{2\mu_1}{HV_t} + \gamma_1 \right) = 0, \ \forall s = 1, 2, \cdots, N. \tag{A.4}$$

The partial derivative of the objective function $\bar{\pi}(\cdot)$ with respect to headway H is

$$\frac{\partial \bar{\pi}}{\partial H} = \bar{f} \sum_{s=1}^{N} \frac{\partial Q_s}{\partial H} + \frac{\mu_1}{H^2} \left(\zeta T_0 + \frac{2D_1}{V_t} + 2\beta_0 N \right) = 0. \tag{A.5}$$

From Eq. (12), Q_1 is a function of L_0, which is a function of λ_1, and thus a function of headway H in terms of Eq. (13). However, L_s $(s = 2, \cdots, N)$ is independent of headway H according to Eq. (1). Therefore, we have

$$\begin{cases} \dfrac{\partial Q_1}{\partial H} = \dfrac{\partial \lambda_1}{\partial H} \displaystyle\int_{L_1}^{L_0} P(x)dx + \lambda_1 P(L_0) \dfrac{\partial L_0}{\partial H} - \dfrac{e_a}{V_a} P(L_0)(L_0 - D_1) \dfrac{\partial L_0}{\partial H} = \dfrac{\partial \lambda_1}{\partial H} \displaystyle\int_{L_1}^{L_0} P(x)dx = -\alpha e_w \displaystyle\int_{L_1}^{L_0} P(x)dx, \\[4mm] \dfrac{\partial Q_s}{\partial H} = \dfrac{\partial \lambda_s}{\partial H} \displaystyle\int_{L_s}^{L_{s-1}} P(x)dx = -\alpha e_w \displaystyle\int_{L_s}^{L_{s-1}} P(x)dx, \ \forall s = 2, \cdots, N. \end{cases} \tag{A.6}$$

Combining Eqs. (A.5) and (A.6) yields

$$H = \sqrt{\frac{\mu_1 \left(\zeta T_0 + \dfrac{2D_1}{V_t} + 2\beta_0 N \right)}{\alpha e_w \bar{f} \sum_{s=1}^{N} \int_{L_s}^{L_{s-1}} P(x)dx}}. \tag{A.7}$$

The partial derivative of $\bar{\pi}(\cdot)$ with respect to fare $\bar{f}$ is

$$\frac{\partial \bar{\pi}}{\partial \bar{f}} = \sum_{s=1}^{N} Q_s + \bar{f} \sum_{s=1}^{N} \frac{\partial Q_s}{\partial \bar{f}} = 0, \tag{A.8}$$

where

$$\begin{cases} \dfrac{\partial Q_1}{\partial \bar{f}} = \dfrac{\partial \lambda_1}{\partial \bar{f}} \int_{L_1}^{L_0} P(x)dx + \lambda_1 P(L_0)\dfrac{\partial L_0}{\partial \bar{f}} - \dfrac{e_a}{V_a} P(L_0)(L_0 - D_1)\dfrac{\partial L_0}{\partial \bar{f}} = \dfrac{\partial \lambda_1}{\partial \bar{f}} \int_{L_1}^{L_0} P(x)dx = -e_f \int_{L_1}^{L_0} P(x)dx, \\ \dfrac{\partial Q_s}{\partial \bar{f}} = \dfrac{\partial \lambda_s}{\partial \bar{f}} \int_{L_s}^{L_{s-1}} P(x)dx = -e_f \int_{L_s}^{L_{s-1}} P(x)dx, \quad \forall s = 2,\cdots,N. \end{cases}$$ (A.9)

Substituting Eq. (A.9) into Eq. (A.8), one obtains

$$\bar{f} = \dfrac{\sum_{s=1}^{N} Q_s}{e_f \sum_{s=1}^{N} \int_{L_s}^{L_{s-1}} P(x)dx} .$$ (A.10)

In view of the above, the system of equations, which consist of Eqs. (A.4), (A.7) and (A.10), defines the optimal rail line length, station location, headway and fare for the flat fare regime. Similarly, one can derive the first-order optimality conditions for the distance-based fare regime. Its proof is omitted here due to the paper length constraint.

Appendix B. Proof of Proposition 4

We need to check the $N \times N$ Hessian matrix, $H_N(\bar{\pi}) = \left(\frac{\partial^2 \bar{\pi}}{\partial D_i \partial D_j} \right)$, of $\bar{\pi}(\cdot)$ with respect to $D_1, D_2,\ldots,D_N$. From Eq. (25) and Eqs. (A.1)-(A.3), the second-order partial derivatives of $\bar{\pi}(\cdot)$ with respect to $D_1, D_2,\ldots,D_N$ are given by

$$\begin{cases} \dfrac{\partial^2 \bar{\pi}}{\partial D_1^2} = \bar{f}\left(\dfrac{\partial^2 Q_1}{\partial D_1^2} + \dfrac{\partial^2 Q_2}{\partial D_1^2} \right), \text{ and } \dfrac{\partial^2 \bar{\pi}}{\partial D_N^2} = \bar{f}\left(\dfrac{\partial^2 Q_{N-1}}{\partial D_N^2} + \dfrac{\partial^2 Q_N}{\partial D_N^2} \right), \\ \dfrac{\partial^2 \bar{\pi}}{\partial D_s^2} = \bar{f}\left(\dfrac{\partial^2 Q_{s-1}}{\partial D_s^2} + \dfrac{\partial^2 Q_s}{\partial D_s^2} + \dfrac{\partial^2 Q_{s+1}}{\partial D_s^2} \right), \quad \forall s = 2,\cdots,N-1, \\ \dfrac{\partial^2 \bar{\pi}}{\partial D_{s-1}\partial D_s} = \bar{f}\left(\dfrac{\partial^2 Q_{s-1}}{\partial D_{s-1}\partial D_s} + \dfrac{\partial^2 Q_s}{\partial D_{s-1}\partial D_s} \right), \quad \forall s = 2,\cdots,N, \\ \dfrac{\partial^2 \bar{\pi}}{\partial D_s \partial D_i} = 0, \quad \forall i \neq s-1, s, s+1. \end{cases}$$ (B.1)

According to $\dfrac{\partial Q_i}{\partial D_s}$ ($i = 1,2,\cdots,N$) shown in Table 3 and (B.1), the Hessian matrix $H_N(\bar{\pi})$ becomes

$$H_N(\bar{\pi}) = \left(\dfrac{\partial^2 \bar{\pi}}{\partial D_i \partial D_j} \right) = \begin{pmatrix} \dfrac{\partial^2 \bar{\pi}}{\partial D_1^2} & \dfrac{\partial^2 \bar{\pi}}{\partial D_1 \partial D_2} & 0 & \cdots & 0 \\ \dfrac{\partial^2 \bar{\pi}}{\partial D_2 \partial D_1} & \dfrac{\partial^2 \bar{\pi}}{\partial D_2^2} & \dfrac{\partial^2 \bar{\pi}}{\partial D_2 \partial D_3} & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & \dfrac{\partial^2 \bar{\pi}}{\partial D_{N-1}\partial D_{N-2}} & \dfrac{\partial^2 \bar{\pi}}{\partial D_{N-1}^2} & \dfrac{\partial^2 \bar{\pi}}{\partial D_{N-1}\partial D_N} \\ 0 & \cdots & 0 & \dfrac{\partial^2 \bar{\pi}}{\partial D_N \partial D_{N-1}} & \dfrac{\partial^2 \bar{\pi}}{\partial D_N^2} \end{pmatrix}.$$ (B.2)

When the population density is uniformly distributed (i.e., $P(x) = P_0$), the total passenger demand for station s, Q_s (see Eqs. (12)-(13)), can then be expressed as

$$\begin{cases} Q_1 = \lambda_1 P_0 \left(L_0 - \dfrac{D_1 + D_2}{2} \right) - \dfrac{e_a P_0}{V_a} \left(\dfrac{\left(D_1 - D_2 \right)^2}{8} + \dfrac{\left(L_0 - D_1 \right)^2}{2} \right), \\[3mm] Q_s = \dfrac{\lambda_s P_0}{2} \left(D_{s-1} - D_{s+1} \right) - \dfrac{e_a P_0}{8 V_a} \left(\left(D_s - D_{s+1} \right)^2 + \left(D_{s-1} - D_s \right)^2 \right), \quad \forall s = 2, \cdots, N, \end{cases} \tag{B.3}$$

where L_0 is determined by Eq. (14) and

$$\lambda_s = 1 - e_w \alpha H - e_t \left(\dfrac{D_s}{V_t} + \beta_0 \left(N + 1 - s \right) \right) - e_f \bar{f}, \quad \forall s = 1, 2, \cdots, N. \tag{B.4}$$

Combining Eqs. (14), (B.1), (B.3) and (B.4), we obtain

$$\begin{cases} \dfrac{\partial^2 \bar{\pi}}{\partial D_1^2} = -\bar{f} P_0 \left(\dfrac{2 e_t}{V_t} \left(\dfrac{1}{2} - \dfrac{e_t V_a}{e_a V_t} \right) + \dfrac{e_a}{V_a} \left(\dfrac{1}{2} + \left(\dfrac{e_t V_a}{e_a V_t} \right)^2 \right) \right), \text{ and } \dfrac{\partial^2 \bar{\pi}}{\partial D_N^2} = -\dfrac{3 e_a}{4 V_a} \bar{f} P_0, \\[3mm] \dfrac{\partial^2 \bar{\pi}}{\partial D_s^2} = -\bar{f} P_0 \dfrac{e_a}{V_a}, \quad \forall s = 2, \cdots, N-1, \\[3mm] \dfrac{\partial^2 \bar{\pi}}{\partial D_s \partial D_{s+1}} = \bar{f} P_0 \dfrac{e_a}{2 V_a}, \quad \forall s = 1, 2, \cdots, N-1. \end{cases} \tag{B.5}$$

Substituting Eq. (B.5) into (B.2), the Hessian matrix $H_N(\bar{\pi})$ can then be written as

$$H_N(\bar{\pi}) = \bar{f} P_0 \begin{pmatrix} -a_{11} & \dfrac{e_a}{2 V_a} & 0 & \cdots & 0 \\[2mm] \dfrac{e_a}{2 V_a} & -\dfrac{e_a}{V_a} & \dfrac{e_a}{2 V_a} & \cdots & 0 \\[2mm] \cdots & \cdots & \cdots & \cdots & \cdots \\[2mm] 0 & \cdots & \dfrac{e_a}{2 V_a} & -\dfrac{e_a}{V_a} & \dfrac{e_a}{2 V_a} \\[2mm] 0 & \cdots & 0 & \dfrac{e_a}{2 V_a} & -a_{NN} \end{pmatrix}, \tag{B.6}$$

where

$$a_{11} = \dfrac{2 e_t}{V_t} \left(\dfrac{1}{2} - \dfrac{e_t V_a}{e_a V_t} \right) + \dfrac{e_a}{V_a} \left(\dfrac{1}{2} + \left(\dfrac{e_t V_a}{e_a V_t} \right)^2 \right), \text{ and } a_{NN} = \dfrac{3 e_a}{4 V_a}. \tag{B.7}$$

In order to show the negative definiteness of the matrix $H_N(\bar{\pi})$, one only needs to check the negative definiteness of the quadric form $Y^T H_N(\bar{\pi}) Y$, where the superscript "T" denotes the transpose of a vector and Y is an N-dimensional column vector, i.e. $Y = \left(y_1, y_2, \cdots, y_N \right)^T$. The quadric form $Y^T H_N(\bar{\pi}) Y$ can be expressed as

$$Y^T H_N(\bar{\pi}) Y = \bar{f} P_0 \left(-a_{11} y_1^2 + \dfrac{e_a}{V_a} y_1 y_2 - \dfrac{e_a}{V_a} y_2^2 + \dfrac{e_a}{V_a} y_2 y_3 - \dfrac{e_a}{V_a} y_3^2 + \cdots - \dfrac{e_a}{V_a} y_{N-1}^2 + \dfrac{e_a}{V_a} y_{N-1} y_N - a_{NN} y_N^2 \right)$$

$$= -\bar{f}P_0\left(\left(a_{11} - \frac{e_a}{2V_a}\right)y_1^2 + \frac{e_a}{2V_a}(y_1 - y_2)^2 + \frac{e_a}{2V_a}(y_2 - y_3)^2 + \cdots + \frac{e_a}{2V_a}(y_{N-1} - y_N)^2 + \left(a_{NN} - \frac{e_a}{2V_a}\right)y_N^2\right). \quad (B.8)$$

According to Eq. (B.7), one obtains

$$a_{NN} - \frac{e_a}{2V_a} = \frac{3e_a}{4V_a} - \frac{e_a}{2V_a} = \frac{e_a}{4V_a} > 0 \text{ , and} \quad (B.9)$$

$$a_{11} - \frac{e_a}{2V_a} = \frac{2e_t}{V_t}\left(\frac{1}{2} - \frac{e_t V_a}{e_a V_t}\right) + \frac{e_a}{V_a}\left(\frac{1}{2} + \left(\frac{e_t V_a}{e_a V_t}\right)^2\right) - \frac{e_a}{2V_a} = \frac{e_t}{V_t}\left(1 - \frac{e_t V_a}{e_a V_t}\right). \quad (B.10)$$

As previously stated, $e_a > e_t$ and the average walking speed of passengers V_a is less than the average train operating speed V_t (i.e. $V_a < V_t$). Accordingly, $\frac{e_t V_a}{e_a V_t} < 1$, and thus $a_{11} - \frac{e_a}{2V_a} > 0$.

Consequently, for any non-zero vector Y, the following inequality always holds

$$Y^T H_N(\bar{\pi})Y < 0. \quad (B.11)$$

This means that the $N \times N$ Hessian matrix $H_N(\bar{\pi})$ is negative definite (Strang, 2006), and thus $\bar{\pi}(\cdot)$ is concave with respect to $D_1, D_2, \ldots, D_N$. This completes the proof of Proposition 4.

Appendix C. Proof of Proposition 5

Let δ represent the even (or average) station spacing of the rail line, then $D_s = (N+1-s)\delta$ and $L_s = (N-s+\frac{1}{2})\delta$. Substituting them into Eq. (B.3), we then obtain the passenger demand for station s as below

$$\begin{cases} Q_1 = \frac{P_0}{2}\frac{V_a}{e_a}\lambda_1^2 + \frac{P_0}{2}\lambda_1\delta - \frac{P_0}{8}\frac{e_a}{V_a}\delta^2, \\ Q_s = P_0\lambda_s\delta - \frac{P_0}{4}\frac{e_a}{V_a}\delta^2, \ \forall s = 2,\cdots,N, \end{cases} \quad (C.1)$$

where

$$\lambda_s = \begin{cases} 1 - e_w\alpha H - e_f\bar{f} - (N+1-s)\left(\frac{e_t}{V_t}\delta + e_t\beta_0\right), \ \forall s = 1,2,\cdots,N, \text{ for flat fare regime,} \\ 1 - e_w\alpha H - e_f f_0 - (N+1-s)\left(\left(\frac{e_t}{V_t} + e_f\hat{f}\right)\delta + e_t\beta_0\right), \ \forall s = 1,2,\cdots,N, \text{ for distance-based fare regime.} \end{cases}$$

$$(C.2)$$

Given the number N of stations, headway H and fare $\bar{f}$ (or $\hat{f}$), and a uniform population density (i.e., $P(x) = P_0$), in the following we derive the optimal even (average) station spacing solution for the flat and distance-based fare regimes, respectively.

(i) For the flat fare regime, from Eq. (25) and $D_s = (N+1-s)\delta$, we have

$$\frac{\partial \bar{\pi}}{\partial \delta} = \bar{f}\sum_{s=1}^{N}\frac{\partial Q_s}{\partial \delta} - N\left(\frac{2\mu_1}{HV_\mathrm{t}} + \gamma_1\right) = 0 , \tag{C.3}$$

where

$$\begin{cases} \dfrac{\partial Q_1}{\partial \delta} = -\delta P_0\left(\dfrac{e_\mathrm{t}}{V_\mathrm{t}}N + \dfrac{1}{4}\dfrac{e_\mathrm{a}}{V_\mathrm{a}} - \left(\dfrac{e_\mathrm{t}}{V_\mathrm{t}}N\right)^2\dfrac{V_\mathrm{a}}{e_\mathrm{a}}\right) + P_0\left(1 - e_\mathrm{w}\alpha H - e_\mathrm{f}\bar{f} - e_\mathrm{t}\beta_0 N\right)\left(\dfrac{1}{2} - \dfrac{e_\mathrm{t}V_\mathrm{a}}{e_\mathrm{a}V_\mathrm{t}}N\right), \\[2ex] \dfrac{\partial Q_s}{\partial \delta} = -(N+1-s)P_0 e_\mathrm{t}\left(\dfrac{2\delta}{V_\mathrm{t}} + \beta_0\right) + P_0\left(1 - e_\mathrm{w}\alpha H - e_\mathrm{f}\bar{f}\right) - \dfrac{P_0}{2}\dfrac{e_\mathrm{a}}{V_\mathrm{a}}\delta, \ \forall s = 2,\cdots,N. \end{cases} \tag{C.4}$$

Substituting Eq. (C.4) into Eq. (C.3) and carrying out some algebraic operations, we obtain the even (average) station spacing solution as below

$$\delta = \frac{A_1}{A_2} - \frac{N}{\bar{f}P_0 A_2}\left(\frac{2\mu_1}{HV_\mathrm{t}} + \gamma_1\right), \tag{C.5}$$

where

$$\begin{cases} A_1 = \left(1 - e_\mathrm{w}\alpha H - e_\mathrm{f}\bar{f}\right)\left(N - \dfrac{1}{2}\right) - \dfrac{e_\mathrm{t}V_\mathrm{a}}{e_\mathrm{a}V_\mathrm{t}}N\left(1 - e_\mathrm{w}\alpha H - e_\mathrm{f}\bar{f} - e_\mathrm{t}\beta_0 N\right) - \dfrac{1}{2}e_\mathrm{t}\beta_0 N^2, \\[2ex] A_2 = \dfrac{e_\mathrm{t}}{V_\mathrm{t}}N^2\left(1 - \dfrac{e_\mathrm{t}V_\mathrm{a}}{e_\mathrm{a}V_\mathrm{t}}\right) + \dfrac{1}{2}\dfrac{e_\mathrm{a}}{V_\mathrm{a}}\left(N - \dfrac{1}{2}\right). \end{cases} \tag{C.6}$$

(ii) For the distance-based fare regime, we have

$$\frac{\partial \hat{\pi}}{\partial \delta} = f_0\sum_{s=1}^{N}\frac{\partial Q_s}{\partial \delta} + \hat{f}\sum_{s=1}^{N}(N+1-s)\left(Q_s + \delta\frac{\partial Q_s}{\partial \delta}\right) - N\left(\frac{2\mu_1}{HV_\mathrm{t}} + \gamma_1\right) = 0 , \tag{C.7}$$

where $Q_s, s = 1,2,\cdots,N$ are determined by Eq. (C.1), and $\dfrac{\partial Q_s}{\partial \delta}, s = 1,2,\cdots,N$ are given by

$$\begin{cases} \dfrac{\partial Q_1}{\partial \delta} = -\delta P_0\left(\dfrac{e_\mathrm{a}}{4V_\mathrm{a}} + N\left(\dfrac{e_\mathrm{t}}{V_\mathrm{t}} + e_\mathrm{f}\hat{f}\right) - N^2\dfrac{V_\mathrm{a}}{e_\mathrm{a}}\left(\dfrac{e_\mathrm{t}}{V_\mathrm{t}} + e_\mathrm{f}\hat{f}\right)^2\right) + P_0\left(1 - e_\mathrm{w}\alpha H - e_\mathrm{f}f_0 - e_\mathrm{t}\beta_0 N\right)\left(\dfrac{1}{2} - N\dfrac{V_\mathrm{a}}{e_\mathrm{a}}\left(\dfrac{e_\mathrm{t}}{V_\mathrm{t}} + e_\mathrm{f}\hat{f}\right)\right), \\[2ex] \dfrac{\partial Q_s}{\partial \delta} = -(N+1-s)P_0\left(2\delta\left(\dfrac{e_\mathrm{t}}{V_\mathrm{t}} + e_\mathrm{f}\hat{f}\right) + e_\mathrm{t}\beta_0\right) + P_0\left(1 - e_\mathrm{w}\alpha H - e_\mathrm{f}f_0\right) - \dfrac{P_0}{2}\dfrac{e_\mathrm{a}}{V_\mathrm{a}}\delta, \ \forall s = 2,\cdots,N. \end{cases} \tag{C.8}$$

On the basis of Eqs. (C.7) and (C.8), one obtains

$$b_1\delta^2 + b_2\delta + b_3 = 0 , \tag{C.9}$$

where

$$b_1 = \hat{f} P_0 N \left(\frac{3}{8} \frac{e_a}{V_a} N + \left(N^2 + \frac{1}{2} \right) \left(\frac{e_t}{V_t} + e_f \hat{f} \right) - \frac{3}{2} \frac{V_a}{e_a} N^2 \left(\frac{e_t}{V_t} + e_f \hat{f} \right)^2 \right),$$
(C.10)

$$b_2 = \hat{f} P_0 N \left(2N \frac{V_a}{e_a} \left(\frac{e_t}{V_t} + e_f \hat{f} \right) \left(1 - e_w \alpha H - e_f f_0 - e_t \beta_0 N \right) + \frac{1}{3} e_t \beta_0 \left(2N^2 + 1 \right) - N \left(1 - e_w \alpha H - e_f f_0 \right) \right)$$

$$+ f_0 P_0 \left(N^2 \left(\frac{e_t}{V_t} + e_f \hat{f} \right) - \frac{V_a}{e_a} N^2 \left(\frac{e_t}{V_t} + e_f \hat{f} \right)^2 + \frac{1}{4} \frac{e_a}{V_a} (2N - 1) \right),$$
(C.11)

$$b_3 = N \left(\frac{2\mu_1}{HV_t} + \gamma_1 \right) - \frac{1}{2} \hat{f} P_0 N \frac{V_a}{e_a} \left(1 - e_w \alpha H - e_f f_0 - e_t \beta_0 N \right)^2$$

$$- f_0 P_0 \left(\left(N - \frac{1}{2} \right) \left(1 - e_w \alpha H - e_f f_0 \right) - \frac{1}{2} N^2 e_t \beta_0 - N \frac{V_a}{e_a} \left(\frac{e_t}{V_t} + e_f \hat{f} \right) \left(1 - e_w \alpha H - e_f f_0 - e_t \beta_0 N \right) \right).$$
(C.12)

It can be shown that the extremal point $\left(-b_2 - \sqrt{b_2^2 - 4b_1 b_3} \right) \big/ 2b_1$ of Eq. (C.9) leads to the minimum profit. Therefore, the optimal even (average) station spacing solution for the distance-based fare regime is

$$\delta = \frac{-b_2 + \sqrt{b_2^2 - 4b_1 b_3}}{2b_1} .$$
(C.13)

This completes the proof of Proposition 5.

Appendix D. Coefficients and solution of Eq. (38) in Proposition 6

(i) The coefficients of Eq. (38)

$$a_1 = \frac{k_1}{k_0} , \quad a_2 = \frac{k_2}{k_0} , \quad \text{and} \quad a_3 = \frac{k_3}{k_0} ,$$
(D.1)

where

$$k_0 = \hat{f} \left(N^3 \left(\frac{e_t}{V_t} + e_f \hat{f} \right) \left(\frac{1}{2} \frac{V_a}{e_a} \left(\frac{e_t}{V_t} + e_f \hat{f} \right) - \frac{1}{3} \right) - \frac{1}{8} \frac{e_a}{V_a} N^2 - \frac{1}{6} \left(\frac{e_t}{V_t} + e_f \hat{f} \right) N \right),$$
(D.2)

$$k_1 = \hat{f} \left(N \left(1 - e_w \alpha H - e_f f_0 - N e_t \beta_0 \right) \left(\frac{1}{2} - N \frac{V_a}{e_a} \left(\frac{e_t}{V_t} + e_f \hat{f} \right) \right) + \left(\frac{1}{2} \left(1 - e_w \alpha H - e_f f_0 \right) - \frac{1}{6} e_t \beta_0 \left(2N - 1 \right) \right) \left(N - 1 \right) N \right)$$

$$+ \frac{1}{2} N^2 \frac{V_a}{e_a} \left(f_0 \left(\frac{e_t}{V_t} + e_f \hat{f} \right)^2 - \bar{f} \left(\frac{e_t}{V_t} \right)^2 \right) + \frac{1}{4} \frac{e_a}{V_a} \left(\frac{1}{2} - N \right) \left(f_0 - \bar{f} \right) - \frac{1}{2} N^2 e_f \hat{f} ,$$
(D.3)

$$k_2 = \frac{1}{2} \hat{f} N \frac{V_a}{e_a} \left(1 - e_w \alpha H - e_f f_0 - N e_t \beta_0 \right)^2 + \frac{1}{2} \left(f_0 \left(1 - e_w \alpha H - e_f f_0 - N e_t \beta_0 \right) - \bar{f} \left(1 - e_w \alpha H - e_f \bar{f} - N e_t \beta_0 \right) \right)$$

$$-N\frac{V_\mathrm{a}}{e_\mathrm{a}}\left(\overline{f}_0\left(1-e_\mathrm{w}\alpha H-e_\mathrm{f}f_0-Ne_\mathrm{t}\beta_0\right)\left(\frac{e_\mathrm{t}}{V_\mathrm{t}}+e_\mathrm{f}\hat{f}\right)-\overline{f}\left(1-e_\mathrm{w}\alpha H-e_\mathrm{f}\overline{f}-Ne_\mathrm{t}\beta_0\right)\frac{e_\mathrm{t}}{V_\mathrm{t}}\right)$$

$$+\left(N-1\right)\left(f_0\left(1-e_\mathrm{w}\alpha H-e_\mathrm{f}f_0\right)-\overline{f}\left(1-e_\mathrm{w}\alpha H-e_\mathrm{f}\overline{f}\right)\right), \tag{D.4}$$

$$k_3=\frac{1}{2}\frac{V_\mathrm{a}}{e_\mathrm{a}}\left(f_0\left(1-e_\mathrm{w}\alpha H-e_\mathrm{f}f_0-Ne_\mathrm{t}\beta_0\right)^2-\overline{f}\left(1-e_\mathrm{w}\alpha H-e_\mathrm{f}\overline{f}-Ne_\mathrm{t}\beta_0\right)^2\right). \tag{D.5}$$

(ii) The solution of Eq. (38)

Let $\Phi=\left(3a_2-a_1^2\right)\big/9$, $\Psi=\left(9a_1a_2-27a_3-2a_1^3\right)\big/54$, $\xi_1=\sqrt[3]{\Psi+\sqrt{\Phi^3+\Psi^2}}$, and $\xi_2=\sqrt[3]{\Psi-\sqrt{\Phi^3+\Psi^2}}$. Then, the roots of Eq. (38) are as follows

$$\begin{cases}\delta_1^*=\xi_1+\xi_2-\tfrac{1}{3}a_1,\\[2mm]\delta_2^*=-\dfrac{1}{2}\left(\xi_1+\xi_2\right)-\tfrac{1}{3}a_1+\dfrac{i\sqrt{3}}{2}\left(\xi_1-\xi_2\right),\\[2mm]\delta_3^*=-\dfrac{1}{2}\left(\xi_1+\xi_2\right)-\tfrac{1}{3}a_1-\dfrac{i\sqrt{3}}{2}\left(\xi_1-\xi_2\right).\end{cases} \tag{D.6}$$

The above roots can be real or complex, which is dependent on the sign of $\Phi^3+\Psi^2$.

(i) If $\Phi^3+\Psi^2>0$, then one root is real and two are complex conjugate.

(ii) If $\Phi^3+\Psi^2=0$, then all roots are real and at least two are equal.

(ii) If $\Phi^3+\Psi^2<0$, then all roots are real and unequal, and given as follows

$$\begin{cases}\delta_1^*=2\sqrt{-\Phi}\cos(\tfrac{1}{3}\omega)-\tfrac{1}{3}a_1,\\[2mm]\delta_2^*=2\sqrt{-\Phi}\cos(\tfrac{1}{3}\omega+120^\circ)-\tfrac{1}{3}a_1,\\[2mm]\delta_3^*=2\sqrt{-\Phi}\cos(\tfrac{1}{3}\omega+240^\circ)-\tfrac{1}{3}a_1,\end{cases} \tag{D.7}$$

where $\cos\omega=\Psi\big/\sqrt{-\Phi^3}$.