Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 502708, 13 pages
http://dx.doi.org/10.1155/2014/502708

Research Article

Investigating the In-Vehicle Crowding Cost Functions for Public Transit Modes

Feifei Qin[1,2]

[1] *School of Urban Rail Transportation, Soochow University, No. 1 Shizi Street, Suzhou 215006, China*
[2] *Molde University College, P.O. Box 2110, 6402 Molde, Norway*

Correspondence should be addressed to Feifei Qin; merryqff@hotmail.com

Received 4 November 2013; Revised 12 December 2013; Accepted 13 December 2013; Published 23 February 2014

Academic Editor: X. Zhang

In the densely populated metropolitan area, empirical studies have found that overcrowding inside transit vehicles has become substantially worse and worse over recent years. Chronic in-vehicle crowding is not only caused by a lack of physical infrastructure, but also triggered by inadequate service provisions. Given the prevalence of overcrowded transit vehicles, this paper conducts both quantitative and qualitative studies, especially focusing on remodeling the in-vehicle crowding cost functions for different transit modes. Three numerical case studies show that applying distinct in-vehicle crowding cost functions to different transit modes has implications not only for the cost structure of transit systems and the magnitude of optimal service provisions but also for the presence of economies of scale in consumption.

1. Introduction

Residents in densely populated urban areas inevitably face travelling trauma. For example, during peak periods, the substantial growth of cars has quickly crammed road capacity, causing severe traffic congestion. Simultaneously, given the difficulties in providing additional rolling stocks, public transit facilities have not been able to cope with growing ridership. Thus, in rush hour, public transit passengers are often forced to tolerate the discomfort and stress associated with having to stand inside extremely overcrowded vehicles.

In the field of transport economics, a large number of theoretical researches and practical studies about road congestion have been conducted by many transport economists [1, 2]. However, the effect of in-vehicle crowding on efficient transit operations has not been carefully considered, even though some exploratory research attempted to address this problem as early as the 1970s [3, 4]. Specifically, the limited body of studies does not provide effective approaches to deal with the in-vehicle crowding effect, leaving a considerable gap between the theoretical research and the practical implementation. Reasons why the research on in-vehicle crowding is still in its infancy can be explored from three aspects. First, given the prevalence of overcrowding in transit, it is surprising that, until now, no definitive definition exists to describe the in-vehicle crowding issues. Second, owing to the diversity of transit modes, the existing research does not afford legible ways to measure the level of in-vehicle crowding, which in turn causes more confusion in terms of evaluating and modeling. Finally, there are many controversies about how to quantify and monetize the crowding costs and incorporate them into transit scheduling and pricing. Ambitiously aiming to solve the above problems in one shot, this paper comprehensively and systematically analyzes the in-vehicle crowding effect through qualitative and quantitative analysis.

The rest of paper is organized as follows. Section 2 reviews some relevant literature and seeks to develop a conceptual framework for precisely defining, measuring, and evaluating the in-vehicle crowding for different transit modes. In Section 3, after reexamining the extant crowding cost functions, three nonlinear function forms are set forth, which aim to capture the relationship between the average value of riding time and crowding levels. Additionally, to demonstrate the performances of the proposed nonlinear crowding cost functions, optimal frequencies are sought by minimizing the total system cost. In Section 4, we simulate the variation in optimal frequencies under a wide range of demand rate

for interurban bus, light rail, and heavy rail, respectively. Section 5 concludes with the main findings.

2. Conceptual Framework for the In-Vehicle Crowding Effect

Noting widespread dissatisfaction with overcrowding, a considerable number of studies have been conducted to specify the social effects of overcrowding from various disciplines, including sociology, psychology, and behavioral sciences. Unfortunately, in most transport economics studies, the in-vehicle crowding effect is incidental to the core research purpose. For example, Tian et al. [5] introduced an unspecified crowding cost function to study the equilibrium properties of many-to-one mass transit systems. de Palma et al. [6] optimized the location and pricing problems for metro stations through specifying one novel nonlinear crowding cost function. Only few scientific studies have been centered on addressing the in-vehicle crowding issues for public transit. To pave the way for the subsequent mathematical modeling and numerical experiments, different areas of research have been integrated into one conceptual framework, including passengers' crowding perceptions, quantitative and qualitative measurements, and the empirical estimates of crowding penalties.

As early as the 1970s, the complex nature of crowding problem attracted significant attention in the field of psychology [7, 8]. From a psychologist's point of view, the term "*crowding*" is made up of both objective and subjective elements [9]. Objective crowding is quantitatively measured by the number of people per unit of space, while the subjective crowding refers to a personal perceived state of mind that may occur when there is great disparity between expected interpersonal distance and actual one [10]. Thus, a central aspect of subjective crowding is the "*felt lack of behavioral freedom and privacy*" when the physical space becomes too limited. In the field of transport studies, "crowding" amounts to the unpleasant experiences of too many passengers fitting into a confined space, thus worsening passengers' well-being [11]. Accordingly, four key crowding effects have been specified in the literature, namely, in-vehicle crowding, platform crowding, excessive waiting time, and increased dwell time [12]. In this paper, only in-vehicle crowding effect is considered in the following analysis.

Since different transit modes are characterized by different interior layout with varying amounts of seats and standing space, passengers have different expectations of getting seats or travelling standing up. For example, because intercity buses and trains are merely designed for seating, passengers believe that purchasing one ticket entitles them to obtain one seat. Once they are forced to stand, they feel great stress and discomfort. In this case, crowding occurs when passengers are unable to obtain seats as they initially expected. On the contrary, the urban rail-based modes and some buses are designed to carry large number of standees rather than to provide as many seats as possible. Consequently, passengers do not view standing as crowding. As more passengers board, passengers start to perceive overcrowding when standing density exceeds one threshold, leading to an "*invasion of privacy*" [7]. In Table 1, the disparate perceptions of in-vehicle crowding are shown for different transit modes. Since transit modes are different in physical capacity design, we cannot apply one common definition for all modes to specify under what conditions passengers can perceive the in-vehicle crowding.

Table 1: The perception of in-vehicle crowding in different transit modes.

	Modes	Passengers' crowding feeling
Intercity (longer journeys)	Intercity rail	No space for seating
	Interurban bus	No space for seating
Commuter (shorter journeys)	Bus	Standees having less than 0.35 m^2 of space
	Light rail	Standees having less than 0.25 m^2 of space
	Heavy rail	Standees having less than 0.17 m^2 of space

Source: edited by author according to different vehicle design guidelines.

Since the crowding perceptions may differ among transit modes, the in-vehicle crowding is less easily measurable. The empirical research chooses either qualitative descriptions or quantitative measures to gauge its level. For example, to model the competition occurring on one rail-based route, Accent Marketing Research and Hague Consulting Group [13] merely described crowding levels from two ways: all seats occupied and easy to find a seat. Noting the limitations of this two-level measurement, Lam et al. [14], Accent [15], and Maunsell and MacDonald [16] all provided the multi-level qualitative descriptions to crudely specify whether the coach condition is crushed or not. Although these qualitative descriptions to some extent assist in specifying the crowding levels, the lack of accuracy and quantification prevents them from being widely used in transportation research. Consequently, to specify the crowding levels as precisely as possible, two quantitative metrics, namely, Load Factor and passenger standing density (pass/m^2), have been extensively adopted. As mentioned above, for interurban buses and trains where users feel overcrowding when they cannot find seats, it is appropriate to measure the crowding degree in terms of Load Factor (LF). However, for massive capacity transit modes, the Load Factor cannot be sensibly utilized, because a high standing capacity can easily make Load Factor exceed 300%. Thus, instead of consistently choosing Load Factor, current empirical studies [11, 17] have recommended adopting standing density (i.e., the number of standing passengers per square meter) as the proper measurement. Table 2 lists seat capacity, crush capacity, and proper crowding measurements for different transit modes.

As early as the 1970s, Goodwin [18] pointed out that overcrowding substantially affects the monetary values that passengers place on travelling time savings. As the crowding level increases, the disutility of travelling increases, so that passengers have to attach significantly high values of riding time to be equally well-off [19]. Similar to transfer penalty, the in-vehicle crowding can be evaluated by crowding penalties.

Table 2: Vehicle capacity and crowding measurements for transit modes.

Urban transport mode	Vehicle type	Seat capacity	Crush capacity[a]	Crowding perception	Crowding measurement
Intercity train	All types	65–95	100–170	Impossibility in finding seats	Load Factor
Interurban bus	All types	38–50	55–65	Impossibility in finding seats	Load Factor
Tram	All types	40–60	90–150	Standing in crush-loaded condition	Standing density
Bus	Single-decker	35–50	55–85	Impossibility in finding seats	Load Factor
	Double-decker	60–75	90–100	Impossibility in finding seats	Load Factor
	Articulated	60–80	100–160	Standing in crush-loaded condition	Load Factor
Light rail	All types	60–80	160–180	Standing in crush-loaded condition	Standing density
Subway	All types	60–80	200–300	Standing in crush-loaded condition	Standing density

[a]The crush capacity is the maximum number of people that it is physically possible to squeeze onto a transit vehicle.

In the published research on crowding valuation, three alternative ways are used to express crowding penalties, namely the time multiplier, the monetary value per time unit, and the monetary value per trip [20]. Since the time multiplier is inherently more transferable across different contexts compared with the monetary valuations, it has become the predominant approach.

Historically, the empirical estimates of crowding penalties have been reported in many unpublished consultancy repots. Only very recently has a small amount of research been published in academic journals. Of all these studies, the one conducted by Wardman and Whelan [21] is the most comprehensive in academic scope, as it estimated crowding penalties not only in terms of load factor but also in terms of standing density (see Table 3). The process can be described as follows: first, the levels of crowding inside transit vehicles are measured by the Load Factor and/or standing density. Then, a Stated Preference (SP) survey is conducted to investigate the passengers' trade-off between waiting time, fares, and the level of crowding. Finally, through Logit models (mainly multinomial Logit (MNL)), the crowding penalties are typically estimated by calculating the ratio of crowding coefficients to the noncrowding coefficient.

This paper largely builds on those values recommend by Whelan and Crockett [11] for the London area. The potential applications of the proposed crowding cost function in other regions are also possible if the city-specified crowding penalties can be obtained by SP/RP survey (see Li and Hensher [20] for Sydney region).

A close look at the magnitudes of crowding penalties given in Table 3 yields some interesting insights. First, the in-vehicle crowding degree, in terms of load factor or standing density, influences the value of riding time spent standing and seating. As the crowding degree increases to crush capacity (load factor >150% or standing density (pass/m^2) >6), the values of riding time get substantially higher. Second, the valuations of crowding penalties are 1.5 to 1.85 when few others are standing. When the vehicle is at crush capacity, crowding penalties exceed 2, which is very much in line with the recommended premium typically attached to walking time and waiting time. Finally, it is obvious that the variation of crowding penalties between successive Load Factors (from 1.5 at 100% to 2.37 at 200%) is much larger than the variation in terms of standing density (from 1.62 for one pass/m^2 to 2.04 for six pass/m^2).

Table 3: Crowding penalties for rail-based modes for London area.

Load Factor (LF)	Crowding penalties for standing	Stand density (pass./m^2)	Crowding penalties for standing
		0	1.53
100	1.5	1	1.62
120	1.67	2	1.70
140	1.85	3	1.79
160	2.02	4	1.87
180	2.2	5	1.96
200	2.37	6	2.04

Source: Whelan and Crockett (2009) [11].

3. Mathematical Modeling for In-Vehicle Crowding Cost

The current empirical studies suggest no matter what type of indicator is employed, crowding penalties monotonically increase with crowding levels. However, the appropriate function form employed to describe the mathematical relationship between the crowding penalties and crowding levels is still underresearched. Furthermore, since the perceptions of the in-vehicle crowding vary among transit modes, we cannot use one common crowding cost function for all transit modes. In what follows, after reexamining the classical function form proposed by Kraus (1991) [22] and other subsequent versions, three nonlinear functions are set out with some explanations.

3.1. Reinvestigating Existing In-Vehicle Crowding Cost Functions. In contrast to the conventional research that generally assigns the same time value regardless of in-vehicle conditions, Kraus (1991) [22] first investigated the effect of the in-vehicle crowding on optimal fare settings. Viewed as a seminal study, Kraus's work divided the value of riding time into two parts: the value of time for passengers who secure seats (P_{v0}) and the value of time in crowded situations where

passengers have to stand ($P_{v0} + \pi$). The generic form of the model is given as

$$P_v = \begin{cases} P_{v0} & \text{No people stand} \\ P_{v0} + \pi & \text{some people stand.} \end{cases} \quad (1)$$

In (1), π is the extra part perceived by standees. Since this simple function form provides an understandable way to capture the crowding externality, a large number of subsequent studies either directly adopted or slightly revised it based on their specific research purposes.

Jansson [23] and Tian et al. [5], respectively, developed their abstract crowding cost functions in which the value of riding time is a function of occupancy rate. But the exact relationship was not explicitly specified. Jara-Diaz and Gschwender [24] updated Kraus's piecewise-constant function by assuming that the values of riding time are no longer constant but rather a continuous on-decreasing function

$$P_v = P_{v0}\left(1 + \rho\phi\right), \quad (2)$$

where, as a markup on the value of riding time (P_{v0}), the crowding penalty (ρ) linearly varies with occupancy rate (ϕ). In contrast to this widely used linear function, a few studies have developed nonlinear ones as incidental to their core research [6, 25].

Although the traditional linear function reflects the additional discomfort and inconvenience associated with in-vehicle crowding, some shortcomings can be easily detected. First, the linear relationship between crowding penalties and crowding levels has been questioned by many empirical observations. A large amount of evidence suggests that the nonlinearity may be empirically supported, especially when psychological elements are factored into crowding perceptions. Second, the existing crowding cost functions only offer a deterministic form and do not specify the probability of having to stand or get a seat. Actually, introducing the probability of having to stand (or sit) as a weight may provide a more convincing way to present the average value of time. Third, a crucial assumption of the linear form is that the in-vehicle crowding effect is sufficiently smooth. Actually, the crowding perception would be sharply deteriorated at a particular crowding level. Finally, since transit modes are different in many dimensions, applying a common crowding cost function for all transit modes and disregarding the specific characteristics of each would lead to undesirable biases.

3.2. Developing Nonlinear Crowding Cost Functions for Different Transit Modes. Whelan and Crockett [11] not only provided a detailed estimation of crowding penalties but also examined possible function forms, such as linear, exponential, and Gompertz. The goodness of fit test for alternative forms indicated that the nonlinearity of crowding penalties with respect to crowding degrees is presented [21]. In addition, to deal with in-vehicle crowding problem for massive capacity modes, such as heavy rail, step function forms could be proper in that it allows discrete jumps among different standing densities [26].

Formula 1: One-Step Nonlinear Function for Crowding Costs. For some modes, such as interurban buses and rails, when the number of passengers inside the vehicle is lower than the seating capacity, the crowding penalty is not active. Thus, the value of riding time remains constant. However, as long as one passenger has to stand, the discomfort and stress resulting from standing would cause the value he attaches to riding time savings to be much higher than those seated passengers. To capture this, a one-step nonlinear function can be derived as follows:

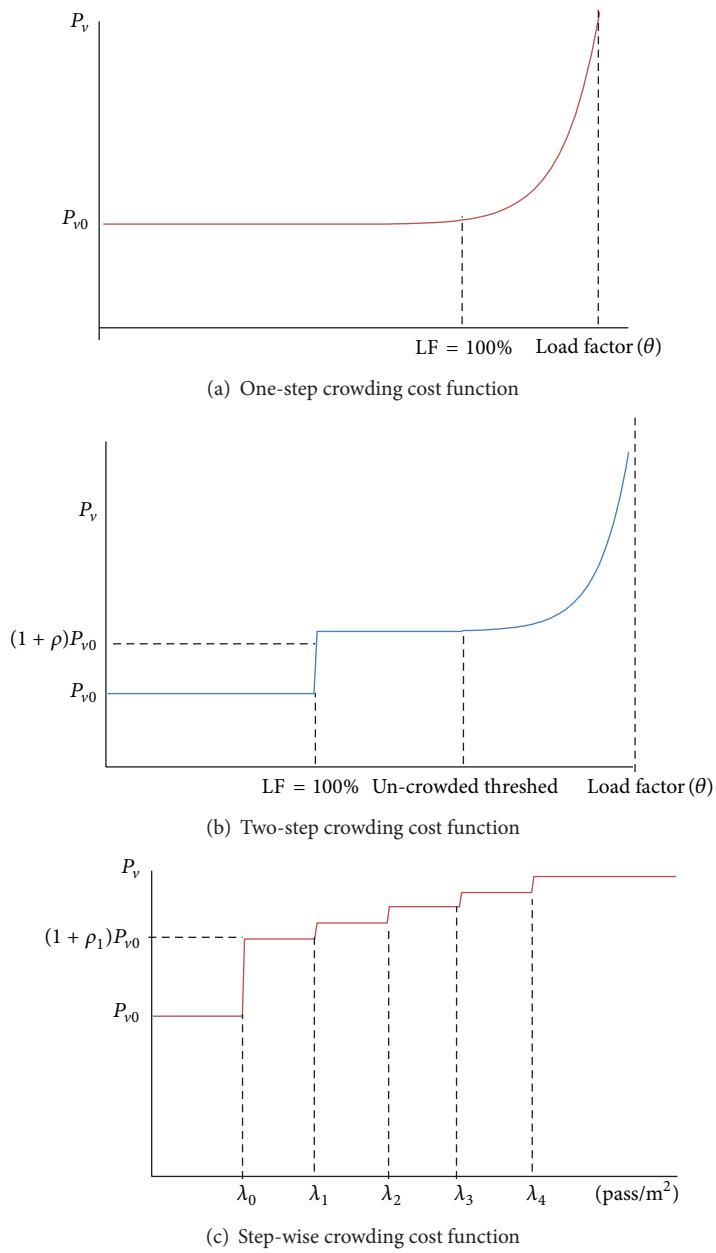
$$P_v = P_{v0}\frac{k}{N} + \left(1 - \frac{k}{N}\right)P_{v0}\left(1 + \beta e^{\alpha(\theta-1)}\right). \quad (3)$$

Equation (3) expresses the weighted value of riding time with a probabilistic form. If we denote k by seating capacity and N by number of passengers on board, the ratio of k/N shows the probability of getting a seat. To clarify the distinction between seating and standing, we assign a constant value of riding time (P_{v0}) to seated passengers and assume that crowding penalties for standees exponentially increase with load factor θ (i.e., N/k). In the exponential part, a certain value of α (more than five) can be arbitrarily given, which serves as an adjustment device to enlarge or shrink the exponential part. By contrast, another parameter β needs to be carefully calibrated from the empirical estimation of crowding penalties. For easy exploring, (3) can be re-arranged as

$$P_v = P_{v0} + \left(1 - \frac{k}{N}\right)P_{v0}\beta e^{\alpha(\theta-1)}. \quad (4)$$

It is envisaged that if the number of passengers is fewer than the number of seats ($N < k$), the item ($\theta - 1$) takes a negative sign. Through the scale parameter α, the exponential part $\beta e^{\alpha(\theta-1)}$ will get close to zero, causing the average value of riding time to be P_{v0}. If the number of passengers equals seating capacity ($N = k$), the item ($1-k/N$) is zero, indicating that the resulting average value of riding time is P_{v0}. As more passengers have to stand in the vehicle, a mark-up ($\beta e^{\alpha(\theta-1)}$) appears as the variable crowding penalty. In the most acute crowding condition, the value of riding time could approach infinity. To explain this function more clearly, we plot the weighted value of riding time against the load factor (θ) in Figure 1(a).

Formula 2: Two-Step Nonlinear Function for Crowding Costs. In contrast to interurban modes, most rail-based urban transit modes (such as tram and light rail) are designed to carry large numbers of standees. For those modes, in-vehicle crowding only takes place when the crowding level exceeds a certain threshold. Although until now no specific research has thus far focused on exploring this threshold, a general accepted observation is that in-vehicle crowding is active when load factor reaches 140% or standing density is above four pass/m^2. Observing this feature, only discriminating standing or sitting cannot describe the impact of crowding on



(a) One-step crowding cost function

(b) Two-step crowding cost function

(c) Step-wise crowding cost function

Figure 1: The graphical illustration of three non-linear crowding cost functions.

passengers' riding time savings. A new function form is thus called for to help differentiate the value of riding time among seating, standing in uncrowded conditions, and standing in crowded vehicles. The following two-step crowding cost function is actually borrowed from de Palma et al.'s work [6]. By slightly adjusting some parameters, the average value of time can be expressed as

$$P_v = P_{v0} + \frac{\rho P_{v0}}{1 + e^{\alpha(1-\theta)}} + \beta P_{v0} e^{\gamma(\theta-\delta)}. \quad (5)$$

Similarly, the first term on the right-hand side of (5) (P_{v0}) corresponds to the value of riding time in the base case where all passengers can find seats. The parameters α and γ can be given arbitrary values as they do not fundamentally affect the value of riding time. However, the parameters β and ρ should be carefully calibrated based on crowding penalties under different LFs. To illustrate the working mechanism, some explanations are made for three exclusive cases. If the load factor is less than one ($\theta < 1$), the denominator of the second term ($1 + e^{\alpha(1-\theta)}$) is scaled up to infinite value through any arbitrarily higher value of α, which in turn makes the second-term approaches zero. Meanwhile, since load factor is less than the threshold ($\theta < \delta$), the third term ($\beta P_{v0} e^{\gamma(\theta-\delta)}$) vanishes as it gets close to zero. Accordingly, the average value of riding time keeps the constant value—P_{v0}. As more passengers board, the limited seating capacity means that only a small proportion of them can find seats. The uncrowned threshold (δ) allows some passengers

Table 4: The standing density design guide for heavy rail.

Standing capacity	The crowding description	Japan	UK	Russia
0-1 pax/m^2	Fewer standing inside vehicle	Comfort	Comfort	Good
1-2 pax/m^2	Some standees	Comfort	Comfort	Good
2-3 pax/m^2	Standees can be free circulation	Comfort	Busy	Good
3-4 pax/m^2	0.25 m^2 of space among passengers Free circulation in the aisles	Little crowded (threshold)	Crowded	Normal
4-5 pax/m^2	0.20 m^2 of space among passengers Some restrictions in movement High probability of physical contact	Some crowded	Very crowded	Normal
5-6 pax/m^2	0.16 m^2 of space among passengers Very restricted circulation Frequent physical contact	Very crowded but tolerable	Maximal	Crowded
6-7 pax/m^2	Less 0.14 m^2 among passengers Impossible movement	Very crowded, crush loaded	Beyond capacity	Crowded
7-8 pax/m^2	Less 0.1 m^2 among passengers Physically squeeze into a car	Intolerable	Beyond capacity	unbearable

to travel by standing in uncrowded travelling conditions. Analogously, the scale parameter α causes the denominator of second-term $(1 + e^{\alpha(1-\theta)})$ not to appreciably differentiate from 1. Simultaneously, parameter γ scales down the third term $(\beta P_{v0} e^{\gamma(\theta-\delta)})$ to zero. Thus, in uncrowded condition, the impact of standing on the average value of time can be reflected by multiplying one mark-up to the value of time in the base case, which is $(1 + \rho)P_{v0}$. If load factor exceeds the threshold $(\theta > \delta)$, the vehicle gets more crowded and passengers have to stand in cramped conditions. In this case, the third term $(\beta P_{v0} e^{\gamma(\theta-\delta)})$ can no longer be ignored and it grows exponentially as each extra passenger boards the vehicle. The graphical representation of (5) is shown in Figure 1(b).

Formula 3: Stepwise Function for Crowding Costs. In contrast to bus-based modes, the physical capacity design guidelines for heavy rail usually adopt standing density to specify crowding levels instead of using load factor. In Table 4, eight levels of standing density are outlined for three representative countries. Furthermore, for each standing density, the travel conditions are described in terms of the available standing space and ability to move.

To this end, the average value of riding time can be modeled as a stepwise function. In this paper, the distinction of value of time is only made for six levels of standing density (i.e., from 0 pax/m^2 to 6 pax/m^2). The specific stepwise function for heavy rail is thus

$$P_v = P_{v0}\left(1 + \frac{\rho_1}{1 + e^{\gamma(\lambda_0 - N)}} + \frac{\rho_2}{1 + e^{\gamma(\lambda_1 - N)}} + \frac{\rho_3}{1 + e^{\gamma(\lambda_2 - N)}} + \frac{\rho_4}{1 + e^{\gamma(\lambda_3 - N)}} + \frac{\rho_5}{1 + e^{\gamma(\lambda_4 - N)}} + \frac{\rho_6}{1 + e^{\gamma(\lambda_5 - N)}}\right), \tag{6}$$

where λ_0 specifies the seating capacity. A new series of parameters λ_i $(i = 1, 2, \ldots, 5)$ denotes the number of passengers for the ith standing density level. The value of crowding penalty ρ_i $(i = 1, 2, \ldots, 6)$ is the incremental change between two consecutive levels. Through the scale-up parameter γ, the average value of time remains constant in each level but jumps discretely from one to another. The diagram of Figure 1(c) illustrates the shape of this function.

3.3. Total System Cost Minimization. In transit operation optimization, one of the approaches is the minimization of total system cost with respect to operating elements, such as fare, frequency, vehicle size, and routes. Since vehicle capacity and route density are assumed to be exogenously given, the only control variable here is frequency. As far as the modeling method is concerned, we adopt a very similar approach to Tirachini et al. [25].

As one of inputs supplied by users, waiting time cost makes up an appreciable part of total system costs. Since travelers can change their behaviors according to service type, we divide transit services into two types: frequency-based service and schedule-based service. For frequency-based transit services, passengers arrive at stations randomly so that the rule-of-thumb "*wait equals half headway*" is an approximation for waiting time. By contrast, for schedule-based services, passengers refer to the timetables before their departure. In this application, we assume that when frequency is greater than twelve vehicles/hour, the service can be viewed as a frequency-based type. When frequency is less than 12 vehicles/hour, the waiting behavior of passengers comprises two parts: passive waiting at other places and active waiting time at stations. It is interesting to note, since passive waiting time can be spent in a productive way, the value of time attached to passive waiting is less than the value of active waiting time at stations [24]. Thus, the general formula of average waiting time cost (w) for two service types is

$$w = P_w\left(t_w + \frac{2\mu}{f}\right), \tag{7}$$

where t_w is the fixed safety time that passengers spend waiting at stations. μ is the ratio of value of time for passive waiting to active waiting. P_w denotes the value of waiting time and f is frequency. Based on the above notations, total waiting time costs, denoted by C_{u-w}, are

$$C_{u-w} = P_w \left(t_0 + \frac{2\psi}{f} \right) Q, \tag{8}$$

where

$$t_0 = \begin{cases} 0, & \text{if } f \geq 12\,\text{veh/h}, \\ t_w, & \text{if } f < 12\,\text{veh/h}, \end{cases}$$
$$\psi = \begin{cases} 1, & \text{if } f \geq 12\,\text{veh/h}, \\ \mu, & \text{if } f < 12\,\text{veh/h}. \end{cases} \tag{9}$$

Another key component of costs—in-vehicle riding time costs—can be specified as the product of the average value of riding time (P_v) and the average riding time (T_v). If l is average trip length and L presents the route length, average riding time can be modeled as a fraction of the cycle time (t_c):

$$T_v = t_c \left(\frac{l}{2L} \right). \tag{10}$$

As passengers' value of riding time grows with crowding degree, the impact of crowding on riding time costs, can be dealt with by replacing P_v with the developed nonlinear crowding cost functions:

$$C_{u-v} = P_v t_c \left(\frac{l}{2L} \right) Q. \tag{11}$$

The so-called vehicle loaded (N) is the average number of passengers aboard, which can be calculated as the ratio of the total number of hourly passengers to the frequency:

$$N = \left(\frac{Q}{f} \right) * \left(\frac{l}{2L} \right). \tag{12}$$

Conventionally, the operating costs are divided into fixed and variable costs incurred in running services. Represented by C_{op}, the operating cost is

$$C_{\text{op}} = c_0 + c_1 B\eta + c_2 VB. \tag{13}$$

The first element is fixed costs per hour, c_o, which includes cost items that do not change with outputs. The second part (c_1) is unit cost per hour, which is determined by the fleet required in the peak period. The third one (c_2) is unit cost per vehicle kilometer. In practice, to avoid unexpected breakdowns, some vehicles remain unused at depots as back-ups. Consequently, a reserve rate of fleet, η, is introduced to reflect this. Denoting V by average running speed, the required fleet (B) can be formulated as

$$B = f t_c = f \left(\frac{2L}{V} \right). \tag{14}$$

Using (14) to eliminate B from (13), the operating cost function is

$$C_{\text{op}} = c_0 + c_1 \left(\frac{2L}{V} \right) \eta f + c_2 2Lf. \tag{15}$$

Ultimately, the objective is to minimize total system cost with respect to frequency (f):

$$\underset{f}{\text{Min}}\ \text{TSC} = C_{u-w} + C_{u-v} + C_{\text{op}}. \tag{16}$$

In practice, service frequency should be neither too low nor too high for operation and safety reasons. In inequity (17), $f_{\min}$ is the public desired minimum level of service and $f_{\max}$ is the maximum feasible frequency decided by the station capacity and safety considerations:

$$f_{\min} \leq f \leq f_{\max}. \tag{17}$$

Besides frequency constraints, the line capacity should be sufficient to accommodate demand:

$$\omega \left(\frac{Q}{2f} \right) \leq K, \tag{18}$$

where ω is the fraction of passengers traveling across the most loaded section and K denotes the physical capacity limit. Optimal frequency is sought by setting the first derivative of the total system cost function equal to zero and solving it subject to the above two constrains.

4. Numerical Experiments for Three Transit Modes

To illustrate the feasibility of incorporating the proposed nonlinear crowding cost functions into optimization, three cases are conducted for interurban bus, light rail, and metro. Two data sets are used. The first one is taken from Tirachini et al. [25]. The second one that has been used to calibrate crowding functions is taken from Whelan and Crockett's study. The values of parameters are summarized in Table 5.

To assess the relative merits of alternative crowding cost functions and the impact of crowding functions on system optimization, the following three scenarios are developed for each numerical case:

Scenario 1. Minimizing total system cost without considering the crowding effect.

Scenario 2. Minimizing total system cost with a classic linear crowding cost function.

Scenario 3. Minimizing total system cost with a nonlinear crowding cost function.

These simple numerical experiments could probably tell us at least three factors of general interests, which are also valid for more complex research. First, we want to know whether the proposed nonlinear functions are applicable for microeconomic modeling research. Second, we are interested

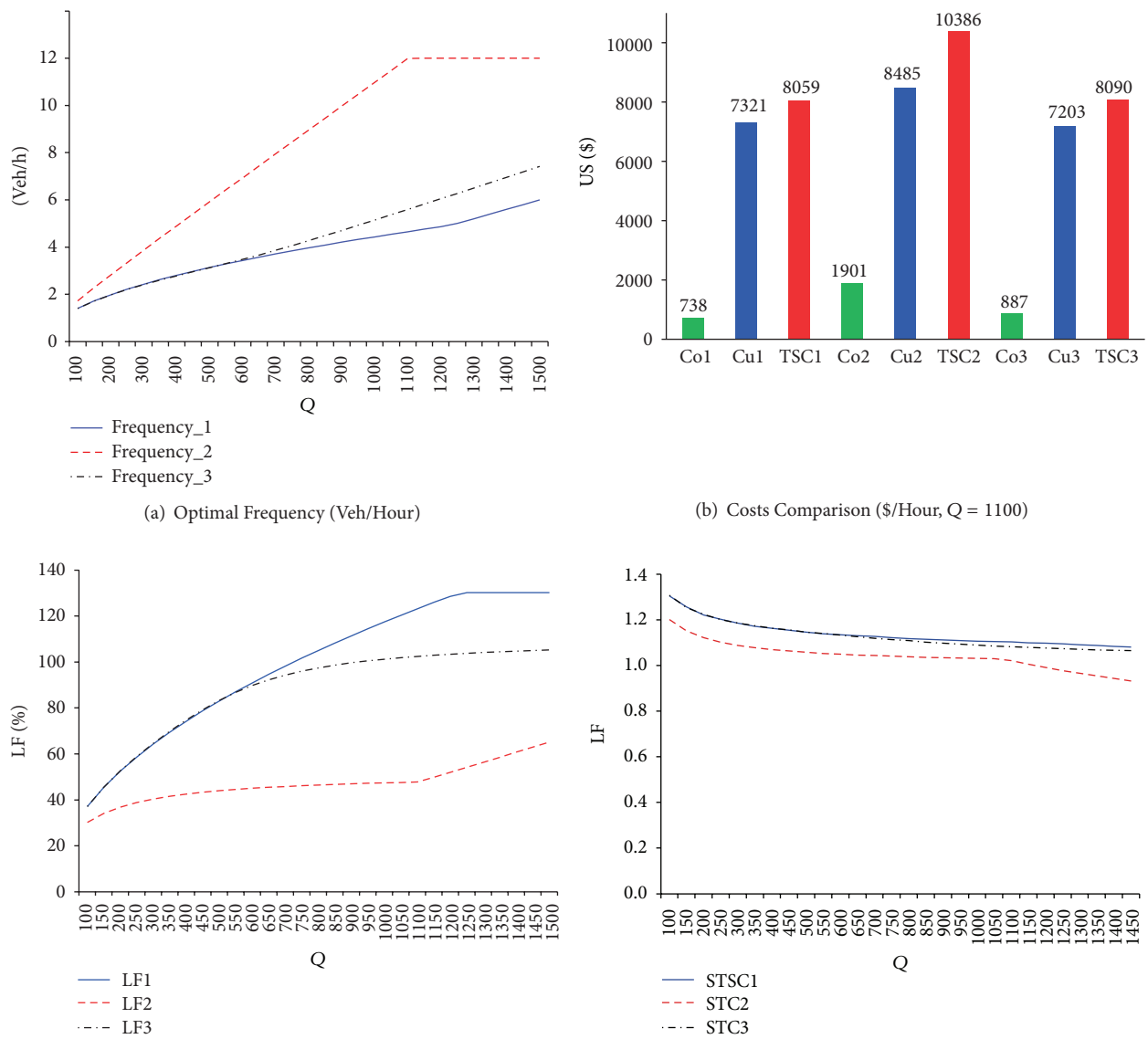
Table 5: Summaries of notations and parameters.

	Parameter	Description	Bus	Light rail	Heavy rail
Operating cost parameters	c_0	Fixed cost (\$/hour)	0	14,866	24,910
	c_1	Unit cost per vehicle-hour (\$/veh/h)	54	164	354
	c_2	Unit cost per vehicle-kilometer (\$/veh/km)	1.13	1.83	3.32
Mode operating parameters	V	Commercial speed (km/hour)	20	35	40
	t_c	Route travel trip time (hour)	2	1.14	1
	k	Seating capacity (seats/vehicle)	48	64	12
	K	Crush capacity (seats + standing/veh)	65	166 (5/m^2)	750 (6/m^2)
	$f_{\max}$	Maximum frequency allowed (veh/h)	200	80	40
Common parameters	P_v	Value of riding time (\$/hour)		9.45	
	P_w	Value of waiting time (\$/hour)		18.9	
	μ	Ratio of value of waiting time at home to at station		0. 33	
	t_w	Safety threshold time (min.)		4	
	L	Route length (km)		20	
	l	Average trip length (km)		10	
	η	Reserve rate for fleet		1.05	
	ω	Fraction of passengers across the loaded sections		30%	
	α	Scale parameter		5	
	γ	Scale parameter		20	

in finding out to what extent the optimal frequencies deviate among the three scenarios for different transit modes. In other words, compared with the linear crowding cost function, do the nonlinear crowding cost functions tend to overstate or understate optimal service frequency? Third, we want to know in which conditions the diseconomies of scale resulting from the "*Kraus effect*," offset the economies of scale arising from the "*Moring effect*," which finally leads to the diseconomies of scale on the consumption side.

4.1. Numerical Case One for Low Capacity Mode. In terms of low capacity transit modes, such as interurban buses and intercity rail, the design guidelines for internal standing areas are strict. For instance, comfortable loading for interurban buses, which allow standees on relatively short trips, should provide at least 0.45 m^2 for each standing passenger. Thus for these low capacity transit modes, the one-step function form for crowding costs is regarded as appropriate. Over a certain range of demand, we can examine the optimal outcomes for three scenarios.

First, embedding the linear crowding cost function into riding time costs generates the highest optimal frequencies across all demand levels. However, as depicted in Figure 2(a), there is no noticeable difference between Scenario 1 and Scenario 3 for low demand. However, the divergence is gradually evident when demand exceeds 700 pax/h. This divergent result can be explained by the fact that the optimal frequencies in Scenario 1 proportionally increase with the square root of demand. However, owning to the inclusion of the in-vehicle crowding externality, the optimal frequencies in Scenario 3 proportionally increase with demand rather than with the square root of demand for high patronage. Additionally, an in-depth comparison finds the linear function in Scenario 2 places more weight on riding time costs than the nonlinear function in Scenario 3 does. Thus, at any demand rate, the values of optimal frequencies in Scenario 2 are well above the optimal frequencies in Scenario 3. Second, Figure 2(b) depicts that when the demand rate is 1100 pax/h, Scenario 2 yields the highest operation costs (1901 \$/h) of the three scenarios, followed by Scenario 3 (887 \$/h) in second place. When the crowding cost is interpreted as a linear form, increased riding time cost dominate reduced waiting time cost, which finally yields the highest users costs in Scenario 2. By contrast, with a relatively smooth slope, the nonlinear function places less weight on riding time cost. Thus, in spite of its relatively higher frequency compared with Scenario 1, Scenario 3 provides lower consumer expenditures than Scenario 1 does. Third, in Figure 2(c), regarding optimal load factors, Scenario 2 intends to accommodate fewer passengers by providing more frequent services, resulting in the inefficient utilization of the vehicle facility. By contrast, by disregarding the crowding externality, Scenario 1 attempts to accommodate passengers as much as possible, which easily causes load factor to exceed 100%. As an intermediate case, Scenario 3 is likely to reduce the incidence of crowding with more efficient usage of transit vehicle. Finally, to assess the degree of scale economies, we introduce a new variable ζ, which is the ratio of the average costs to marginal costs. If $\zeta > 1$, we can confirm that there exist scale economies. The opposite case ($\zeta < 1$) denotes diseconomies of scale, while $\zeta = 1$ means neutral case [27]. From Figure 2(d), we can observe across all tested demand that Scenario 1 and Scenario 3 present the economies of scale. However, since the linear crowding cost curve is steeper than the nonlinear one, Scenario 2 puts more weight on the riding time cost. Thus, a considerable diseconomy of scale is present, particularly when ridership is high.



(a) Optimal Frequency (Veh/Hour)

(b) Costs Comparison ($/Hour, $Q = 1100$)

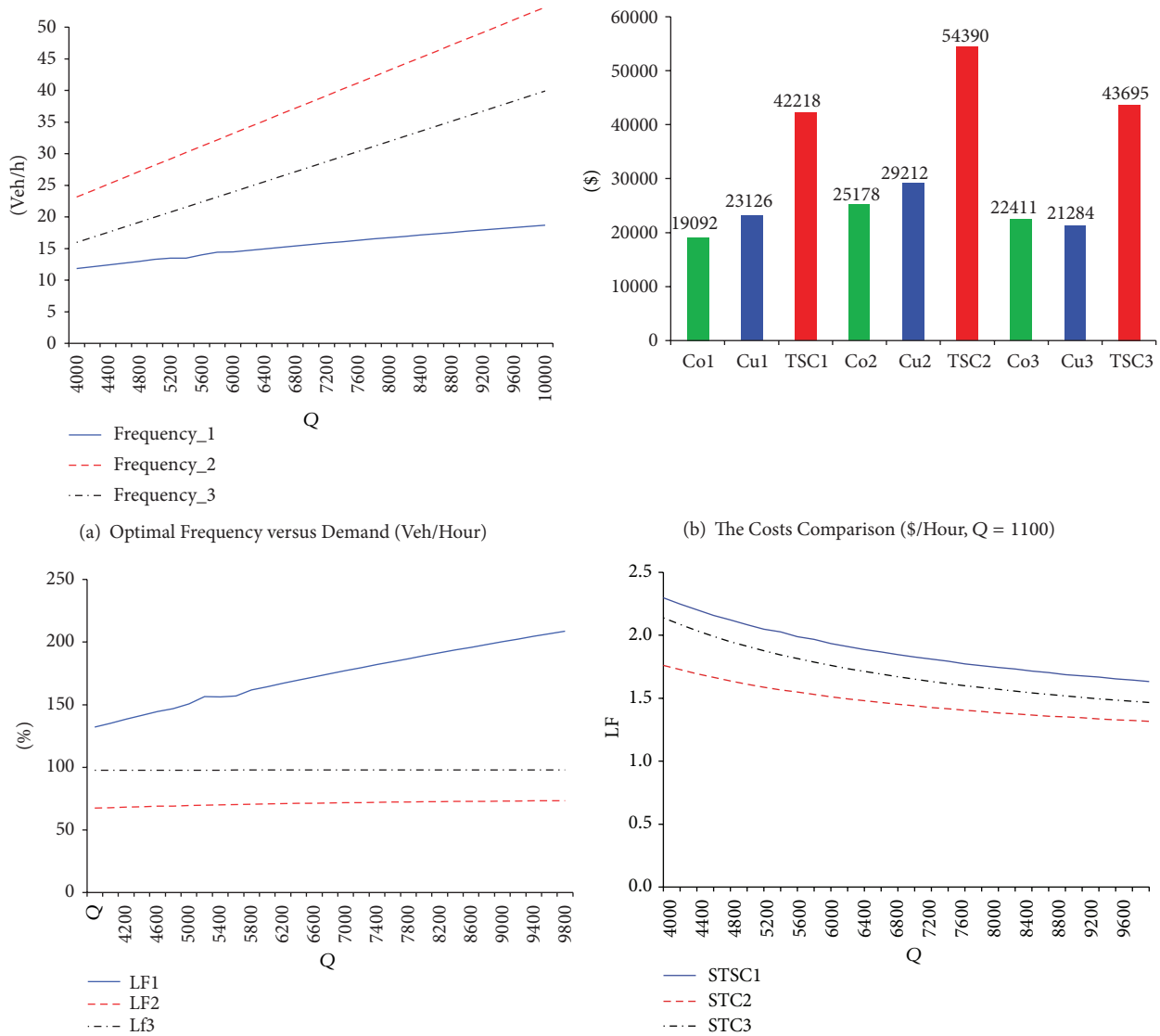
(c) Load Factor versus demand (%)

(d) The Existence of Economies of Scale

Figure 2: The Optimal Outputs for Numerical Case One (Here, C_o denotes the operating cost of the transit operator. C_u is the user cost and TSC means total system cost).

4.2. Numerical Case Two for High Capacity Modes. In the context of transit systems, most rail-based transit modes are designed to carry more standees than seated passengers so that passengers have high tolerance of in-vehicle crowding. Consequently, besides keeping the value of riding time for seating constant, it is reasonable to assume, when crowding level is below a certain threshold, the value of riding time also keeps constant but takes a higher value. Once crowding level is greater than threshold, the sharply increased discomfort makes the value of riding time present an exponential growth pattern. In this case, the threshold of uncrowded level is in line with the prevailing standards for physical vehicle capacity design (LF = 140%). The major results, in terms of optimal frequencies, cost components, and optimal LFs, are graphically represented in Figure 3.

Figure 3(a) shows the simulated optimal frequencies for three scenarios. Unlike Case 1, the frequency difference between Scenario 1 and Scenario 3 is apparent in this case. This divergence can be attributed to the fact that incorporating the two-step function into system optimization causes frequencies to vary proportionally with demand from the initial tested demand rate. As graphically illustrated by Figure 3(b), Scenario 2 implies the highest total system costs, user costs, and operating costs. However, the inclusion of the nonlinear function in total system cost minimization incurs slightly lower user costs than Scenario 1 does. This can be explained by the fact that although considering crowding effect generally increases the relative weight of riding time costs in the total cost function, the different function forms of crowding cost deliver different weights. A plot of the optimal



(a) Optimal Frequency versus Demand (Veh/Hour)

(b) The Costs Comparison ($/Hour, $Q = 1100$)

(c) Load Factors versus Demand (%)

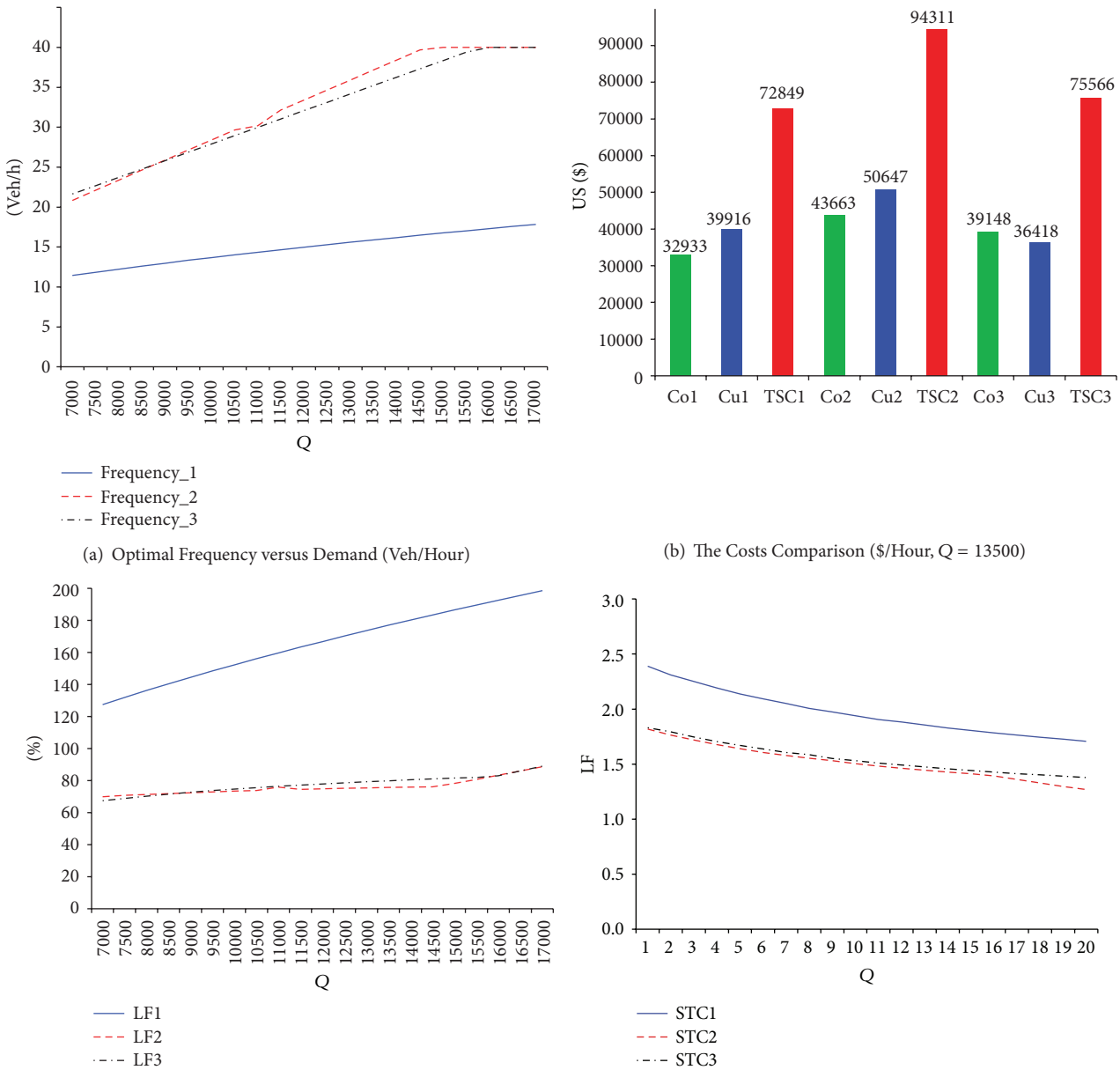
(d) The Existence of Economies of Scale

Figure 3: The Optimal Outputs for Numerical Case Two.

load factors versus demand is given in Figure 3(c). Due to the ignorance of crowding externality, Scenario 1 intends to fill vehicles with more passengers so that the optimal load factors range from 130% to 200%. In such a case, passengers consider discomfort to be impaired by high load factors. On the contrary, Scenario 2 has an incentive to run more frequent services with many empty seats, suggesting that the load factor is below 100%. The optimal load factors in Scenario 3 are around 100%. Thus, from a cost efficiency point of view, the most promising scenario would be Scenario 3. Concerning the degree of scale economies, Figure 3(d) depicts that all three scenarios involve pronounced scale economies. Furthermore, a closer look at Scenario 2 and Scenario 3 shows that no matter which type of crowding cost function is implemented the diseconomies of scale in users' riding costs resulting from introducing the crowding externality are dominated by the coexistence of economies in producer cost and waiting time cost, which finally generates the magnitude of scale economies in total system cost.

4.3. Numerical Case Three for Massive Capacity Modes. The latest version of PDFH 5 [17] provides six crowding penalties for a certain range of standing densities and recommends adopting these values for metro planning. Following the suggestion of PDFH, a stepwise function was developed for massive capacity transit modes, particularly for heavy rail. In this case, the stepwise function was examined for its applicability. The optimal solutions of three scenarios yield some insights.

First, due to disregarding the crowding effect, Scenario 1 intends to provide less frequent services along the entire range of demand, as shown in Figure 4(a). However, a comparison between Scenario 2 and Scenario 3 shows that the frequency difference is relatively small before the demand



(a) Optimal Frequency versus Demand (Veh/Hour)

(b) The Costs Comparison ($/Hour, $Q = 13500$)

(c) Load Factors versus Demand (%)

(d) The Existence of Economies of Scale

Figure 4: The Optimal Outputs for Numerical Case Three.

exceeds 11,500 pax/h. Beyond that point, Scenario 2 starts to adjust the frequency up in order to reduce the increasing crowding, since the slope of a linear function is steeper than the slope of stepwise function. Second, to show the effectiveness of vehicle utilization, we plot the load factor curves in Figure 4(c). In Scenario 1, the optimal load factors increase from 130% at 7,000 pax/h to 200% at 17,000 pax/h. On the contrary, in Scenario 2 and Scenario 3, the loaded passengers never exceed the vehicle capacity, implying that enough capacity is afforded to help relieve overcrowding. Finally, Figure 4(d) clearly depicts that Scenario 1 yields the scale economies, resulting from the combined scale economies on the consumption and production sides. By contrast, at high demand rates, Scenario 2 and Scenario 3 are prone to be diseconomies of scales. The likely explanation is that the diseconomy of scale from "*Kraus effect*" dominates the economy of scale from "*Mohring effect*," which finally leads to the diseconomy of scale on consumption side.

5. Summary

Recently, means of public transit often ply in extremely overcrowded conditions in large metropolitan cities. Although crowding can be mitigated through costly infrastructure improvement and network expansion, it can also be avoided through low-cost crowding relief measures if crowding cost can be correctly formulated. To provide insights into the in-vehicle crowding effect, this paper has presented three

nonlinear crowding cost functions and discussed the implementation in the system optimization.

A synthesis of the theoretical and numerical analysis generates valuable insights. First, with nice mathematical properties, the proposed nonlinear functions can be successfully implemented in system optimization. From a methodology point of view, the advantage of the nonlinear crowding cost function over the simple linear one is that nonlinear formulas provide a more realistic and rational representation. Second, numerical case studies show that, irrespective of the crowding costs function form (linear or nonlinear), the inclusion of crowding effect has complex impacts on the level of service provision, the cost of the system, and the degree of scale economies. Specifically, incorporating crowding costs functions into the system optimization usually generates frequent transit services, low users costs, and more efficient vehicle usage. With high demand rates, the diseconomies of scale from the *Kraus effect* may dominate the economies of scale from the *Mohring effect*, which probably leads to the presence of scale diseconomies in total system cost. Thus, we should treat the *Kraus effect* as importantly as the *Mohring effect*.

In most transit assignment applicants, in-vehicle crowding is not taken into account for modeling the mode/route choices. The prevailing software, such as EMME/2 and Tube, usually uses the fixed in-vehicle travel costs. In order to enable the transit assignment module to reflect crowding costs, all original fixed travel cost functions should be multiplied by the item $*(1 + \rho)$. In terms of the proposed crowding cost function, the equilibrium assignment function can be separated in a linear travel time cost part and a nonlinear crowding part. During the assignment steps, the nonlinear crowding part can be iteratively fed back to the generalized cost for each route/mode until the whole network reaches equilibrium.

The contributions of this paper to the literature are twofold. Firstly, after observing the disadvantages of conventional linear crowding cost functions, this paper proposed three nonlinear function forms for different transit modes to help decision-makers better examine the impacts of in-vehicle crowding on transit service planning. Secondly, for the first time, two separate strands of research work closely related to the in-vehicle crowding effect (i.e., the empirical estimation work on crowding penalties and theoretical in-vehicle crowding cost modeling) are integrated to fill the gap between theory and practice. It is hoped that the nonlinear crowding cost functions proposed here will help planners make better use of existing line capacities in order to relieve in-vehicle crowding before costly capital works are required.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The author is very grateful to two anonymous referees whose comments have led to a significant improvement of the paper. Financial support from the National Natural Science Foundation of China (Grant no. 71301112) is gratefully acknowledged.

References

[1] R. Arnott and K. Small, "The economics of traffic congestion," *American Scientist*, vol. 82, no. 5, pp. 446–455, 1994.

[2] O. Johansson, "Optimal road-pricing: simultaneous treatment of time losses, increased fuel consumption, and emissions," *Transportation Research D*, vol. 2, no. 2, pp. 77–87, 1997.

[3] H. Mohring, "Optimization and scale economies in urban bus transportation," *American Economic Review*, vol. 62, pp. 591–604, 1972.

[4] R. Turvey and H. Mohring, "Optimal bus fares," *Journal of Transport Economics and Policy*, vol. 9, pp. 280–286, 1975.

[5] Q. Tian, H.-J. Huang, and H. Yang, "Equilibrium properties of the morning peak-period commuting in a many-to-one mass transit system," *Transportation Research B*, vol. 41, no. 6, pp. 616–631, 2007.

[6] A. de Palma, M. Kilani, and S. Proost, "The location and pricing of mass transit stations," in *Proceedings of the 12th World Conference on Transport Research Society*, pp. 978–989, 2010.

[7] M. L. Fried and V. J. DeFazio, "Territoriality and boundary conflicts in the subway," *Psychiatry*, vol. 37, no. 1, pp. 47–59, 1974.

[8] E. S. Knowles, "The gravity model of crowding: application to social physics to the effects of others," in *Human Responses to Crowding*, A. Baum and Y. Epstein, Eds., Lawrence Erlbaum Associates, New York, NY, USA, 1978.

[9] D. Stokols, "On the distinction between density and crowding: some implications for future research," *Psychological Review*, vol. 79, no. 3, pp. 275–277, 1972.

[10] R. Baron and D. Richardson, *Human Aggression Edition*, Plenum, New York, NY, USA, 2nd edition, 2004.

[11] G. Whelan and J. Crockett, "An investigation of the willingness to pay to reduce rail overcrowding," in *Proceedings of the International Conference on Choice Modeling*, Harrogate, UK, 2009.

[12] Sinclair Knight Merz (SKM), "Critical Review of Transport Modelling Tools," 2011, http://www.bitre.gov.au/publications/34/Files.

[13] Accent Marketing and Research and Hague Consulting Group, "Rail Differential Pricing," Report Prepared for Rail Operational Research, 1996.

[14] W. H. K. Lam, C.-Y. Cheung, and C. F. Lam, "A study of crowding effects at the Hong Kong light rail transit stations," *Transportation Research A*, vol. 33, no. 5, pp. 401–415, 1999.

[15] Accent, "Rail Customer Valuations of Sitting, Standing and Crowding," Report Prepared for Transport for London, 2006.

[16] F. Maunsell and M. MacDonald, "Rail Overcrowding, Reliability and Frequency," Report Prepared for Demand Impacts Project for Centro, 2007.

[17] Association of Train Operating Companies (ATOC), "Passenger Demand Forecasting Handbook (Version 5)," Report prepared for British Railway Board, 2009.

[18] P. B. Goodwin, "Human effort and the value of travel time," *Journal of Transport Economics and Policy*, vol. 10, no. 1, pp. 3–15, 1976.

[19] A. Polydoropoulou and M. Ben-Akiva, "Combined revealed and stated preference nested logit access and mode choice model

for multiple mass transit technologies," *Transportation Research Record*, no. 1771, pp. 38–45, 2001.

[20] Z. Li and D. A. Hensher, "Crowding and public transport: a review of willingness to pay evidence and its relevance in project appraisal," *Transport Policy*, vol. 18, no. 6, pp. 880–887, 2011.

[21] M. Wardman and G. Whelan, "Twenty years of rail crowding valuation studies: evidence and lessons from British experience," *Transport Reviews*, vol. 31, no. 3, pp. 379–398, 2011.

[22] M. Kraus, "Discomfort externalities and marginal cost transit fares," *Journal of Urban Economics*, vol. 29, no. 2, pp. 249–259, 1991.

[23] K. Jansson, "Optimal public transport price and service frequency," *Journal of Transport Economics and Policy*, vol. 27, pp. 33–50, 1993.

[24] S. R. Jara-Diaz and A. Gschwender, "Towards a general microeconomic model for the operation of public transport," *Transport Reviews*, vol. 23, no. 4, pp. 453–469, 2003.

[25] A. Tirachini, D. A. Hensher, and S. R. Jara-Díaz, "Comparing operator and users costs of light rail, heavy rail and bus rapid transit over a radial public transport network," *Research in Transportation Economics*, vol. 29, no. 1, pp. 231–242, 2010.

[26] Australian Transport Council (ATC), *National Guidelines for Transport System Management in Australia, Canberra*, 2006.

[27] K. A. Small and E. T. Verhoef, *The Economics of Urban Transportation*, Routledge, London, UK, 2007.