

Effect of Transit Capacity onto Morning Commute Problem with Competitive Modes and Distributed Demand

Hiroshi Shimamoto

Department of Civil and Environmental Engineering, University of Miyazaki, Japan
Email: shimamoto@cc.miyazaki-u.ac.jp

How to cite this paper: Shimamoto, H. (2022) Effect of Transit Capacity onto Morning Commute Problem with Competitive Modes and Distributed Demand. *Journal of Transportation Technologies*, 12, 80-95. <https://doi.org/10.4236/jtts.2022.121005>

Received: November 14, 2021

Accepted: January 7, 2022

Published: January 10, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

A lane exclusively for transit is one measure to reduce urban congestion and intelligent transportation systems (ITSs) are expected to make this policy more effective. However, the advantage may be limited if a transit agency does not provide sufficient vehicles. This study evaluates the effect of transit capacity on a dedicated transit lane policy by examining the morning commute problem, by defining user equilibrium and system optimisation considering transit capacity. It was confirmed that reduced transit capacity under user equilibrium smoothly increases the transit cost per person, which decreases the duration period of the dedicated transit lane and increases the number of early and delayed car commuters. However, the effect of reducing transit capacity under a system optimum depends on the initial solution of the system optimum.

Keywords

Dedicated Transit Lane, Transit Capacity, User Equilibrium, System Optimum, System-Optimal Pricing

1. Introduction

Congestion due to the morning commute to a central business district (CBD) is a common problem globally, which may cause delays in the transport systems [1]. The congestion is caused by the concentration of demand. To mitigate such congestion, many cities have taken measures to increase the levels of service offered by public transit systems. Dedicated transit lanes are one example of increasing the level of service of public transit. Intelligent transportation system (ITS) technology is expected to make dedicated transit lane policies more effec-

tive. However, because dedicated transit lanes also affect road-traffic congestion, it is essential to evaluate how they will affect commuters' behaviour.

The morning commute problem is one of the methods used to describe commuters' behaviour in passing a bottleneck. Vickrey [2] first proposed a model to describe commuters' choice of departure time for a single route with a single bottleneck. Each commuter was assumed to have the same desired departure time and to choose that time so as to minimise the sum of travel cost plus the penalty for early or delayed arrival. Hendrickson and Kocur [3] expanded the model to consider a distribution of desired departure times via a bottleneck. Smith [4] showed the existence of an equilibrium of commuters' departure time choice when no commuters were motivated to change their departure time to reduce their travel cost under the condition that the desired arrival time distribution was S-shaped and the penalty function was smooth and convex. Arnott *et al.* [5] examined the user equilibrium of departure time decisions considering heterogeneous commuters. Arnott *et al.* [6] further compared the welfare effects between an optimal time-varying toll, capacity expansion, and toll-financed capacity expansion. The optimal time-varying toll was identical to the difference between the user equilibrium and system-optimal costs. Liu and Nie [7] proposed a morning commute problem that considered simultaneous route and departure time choices of heterogeneous users. They showed that a time-based system-optimal toll that minimised the total travel time provided a Pareto-improving solution to the social inequality issue associated with the system optimum, which minimises the system cost in the unit of monetary value. Sakai *et al.* [8] examined the morning commute problem for a two-to-one merge bottleneck. They showed a Pareto improving pricing scheme in such a network.

In the studies mentioned above, only cars have been considered in the morning commute problem. Gonzales and Daganzo [9] extended the problem for a single bottleneck to consider mode choice between cars and public transit. They showed both user equilibrium and the system optimum under competition of two modes. Because they assumed that cars and public transit experience the same bottleneck but with a fully segregated dedicated public transit lane, their model is useful for evaluating the effect of introducing a dedicated transit lane. Gonzales and Daganzo [10] further considered both the morning-commute and evening-commute problems and showed that user equilibrium for the two commutes in isolation was asymmetric but the system optimum for the two was symmetric. Note that these papers did not explicitly consider the capacity of each public transit vehicle. However, especially in developing countries, even if a dedicated transit lane is introduced, the level of service of public transportation may not be sufficiently improved if there is insufficient transit capacity. In such cases, it is important to evaluate the effects on both users and society of increasing and decreasing transit capacity so as to maximise the effect of introducing dedicated transit lanes.

Based on these considerations, this paper analyses a morning commute problem considering competing modes of transport and finite transit capacity, and

then evaluates the effects of insufficient transit capacity on commuters' travel patterns and social welfare.

2. Morning Commute with Cars and Transit

For readers' convenience, this section reviews the morning commute problem for a single bottleneck that includes a dedicated transit lane and mode choice, as outlined by Gonzales and Daganzo [9], factors that are required to describe our proposed model in the next section. After describing the assumptions, we first review the user equilibrium model in Section 2.2 and then describe the system-optimal model in Section 2.3.

2.1. Problem Statement

Assume there is a single bottleneck between commuters' homes and the CBD and the capacity of the bottleneck is μ . Consider the morning commute problem with a population of commuters who are identical, except for their desired departure times. The total number of N commuters who wish to depart from the bottleneck by t is described by a wish curve, $W(t)$, which is S-shaped as shown in **Figure 1**. It is further assumed that the demand per time unit exceeds the bottleneck capacity during the period between t_1 and t_2 as shown below.

$$\begin{cases} \dot{W}(t) > \mu & \text{for } t \in (t_1, t_2) \\ \dot{W}(t) \leq \mu & \text{otherwise} \end{cases} \quad (1)$$

(Note that $\dot{X}(t)$ denotes the time-derivative of a function $X(t)$.)

Suppose that each commuter experiences a penalty for schedule deviation from his/her desired departure time, $W(t)$; each minute of earliness is associated with a penalty of e equivalent minutes of travel time such that $0 < e < 1$, and each minute of lateness is equivalent to L minutes of travel time such that $L > 0$.

Both a car and public transit are assumed to be available. It is assumed that when transit is operating, it is fully segregated in its own lane (dedicated transit lane) so that transit services are not subject to traffic congestion. Thus, the automobile capacity of the bottleneck during transit operation decreases to $m\mu$ with $0 \leq m < 1$. The transit agency is assumed to provide sufficiently large vehicles with a regular schedule; the assumption of sufficient capacity will be relaxed later in this paper.

2.2. User Equilibrium

Based on the above assumptions regarding the transit service, transit users can always choose to pass the bottleneck at their desired time. Therefore, each transit user has an identical generalised cost, Z_T (hours). A car journey with no delay also has a generalised cost, Z_C (hours), which is independent of the number of car users. Under the user equilibrium condition, when transit is provided, the maximum delay when travelling by car, T , satisfies the following.

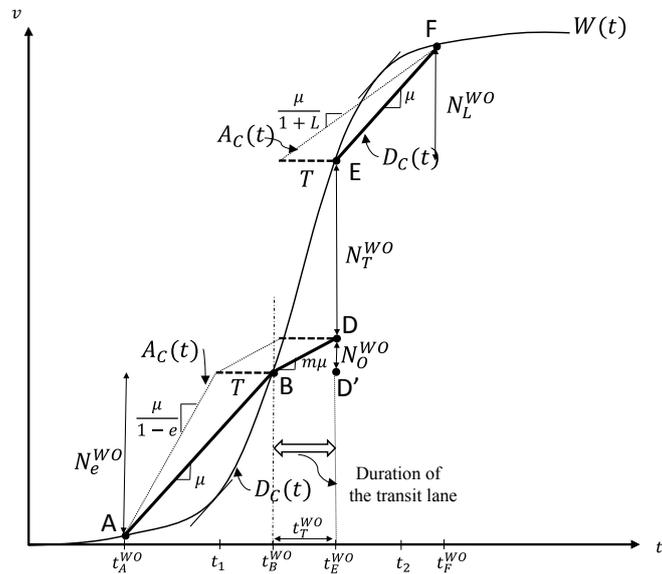


Figure 1. User equilibrium for a bottleneck with a transit alternative.

$$T = Z_T - Z_C < T_C^{WO} \tag{2}$$

where T_C^{WO} represents the maximum travel cost as the delay caused by queuing in the single-mode case, which can be formulated as:

$$T_C^{WO} = \frac{eL}{e+L} \frac{N_Q^{WO}}{\mu} \tag{3}$$

where N_Q^{WO} is the number of commutes delayed by the bottleneck. The superscript “WO” here indicates “without transit capacity constraints”. Note that we do not consider the case $Z_T < Z_C$ when all journeys will be made by transit or the case $Z_T > Z_C + T_C$ when all journeys will be made by car.

Figure 1 shows the user equilibrium travel pattern for a single bottleneck with both cars and transit. The dedicated transit lane is located between t_B^{WO} and t_E^{WO} , and no transit service is operated other than between t_B^{WO} and t_E^{WO} . $A_C(t)$ and $D_C(t)$ respectively represent the cumulative number of car arrivals and the cumulative number of car departures at the bottleneck by time t . According to Proposition 1 of Gonzales and Daganzo [9], there are four types of commuters at equilibrium:

- 1) Early car commuters who travel at the beginning of the rush, $t \in (t_A^{WO}, t_B^{WO})$;
- 2) Late car commuters who travel at the end of the rush, $t \in (t_E^{WO}, t_F^{WO})$;
- 3) On-time car commuters who travel in the middle of rush, $t \in (t_B^{WO}, t_E^{WO})$;
- 4) On-time transit commuters who travel in the middle of rush, $t \in (t_B^{WO}, t_E^{WO})$.

The number of early car commuters, N_e^{WO} , and the number of late car commuters, N_l^{WO} , can be analytically obtained respectively as $N_e^{WO} = \mu T / e$ and $N_l^{WO} = \mu T / L$. The number of on-time car commuters, N_o^{WO} , has been proven to be a strictly decreasing function of T^{WO} , and the number of on-time transit users, N_T^{WO} , has also been proven to be a strictly decreasing function of T^{WO} . The number of commuters during the rush, N_Q^{WO} , is given as:

$$N_Q^{WO} = N_e^{WO} + N_L^{WO} + N_O^{WO} + N_T^{WO} \quad (4)$$

2.3. System Optimum

So far, the user equilibrium travel pattern under the competitive mode has been reviewed. It is also possible to define a system-optimal travel pattern that minimises the total system cost associated with the bottleneck. Gonzales and Daganzo [9] segmented a system-optimal solution into three phases: an early phase $t \in (-\infty, t_B)$ in which transit is not used, cars flow at capacity, and all departures are early; a middle phase $t \in (t_B, t_E)$ in which transit is used, all departures are punctual and cars flow either at capacity or not at all; and a third phase $t \in (t_E, \infty)$ in which transit is not used, cars flow at capacity, and all departures are late. In the optimal system, the transit service and dedicated transit lane are provided only during the middle phase.

The duration of dedicated transit lane provision, $t_T (= t_E - t_B)$ and the transit headway are chosen to minimise total cost. Then, the number of transit users in the optimal system can also be determined. The system-optimal transit cost $Z_T(N_T, t_T)$ can be the analytic function of N_T and t_T , and this function is proven to be concave with respect to N_T and t_T . Then, one can obtain a system-optimal solution by minimising the total cost $Z(N_T, t_T)$ as follows:

$$Z(N_T, t_T) = S(t_T) + Z_C[N - N_T] + Z_T(N_T, t_T) \rightarrow \min \quad (5)$$

Subject to

$$\max\{0, M(t_T) - m\mu t_T\} \leq N_T \leq M(t_T) \quad (6)$$

Equation (5) represents the total cost, where $S(t_T)$ is the schedule penalty, $Z_C[N - N_T]$ is the total cost for car users, and $Z_T(N_T, t_T)$ is the system-optimal transit cost. Equation (6) represents the constraint condition that the number of transit users must satisfy. $M(t_T)$ in Equation (6) represents the number of users travelling during the period the dedicated transit lane is provided.

Because both $Z_C[N - N_T]$ and $Z_T(N_T, t_T)$ are proven to be concave with regard to N_T , the variable N_T should take an extreme point at an optimum. Therefore, the optimisation problem in Equation (5) and Equation (6) can be reduced to the 1-varibale optimisation problem with regard to t_T . To find the optimum, one can solve two optimisation problems by substituting $\max\{0, M_T(t_T) - m\mu t_T\}$ and $M(t_T)$ for N_T in Equation (5), and then adopt the solution with the smaller of the two objective functions. If the number of transit users N_T takes the value $\max\{0, M_T(t_T) - m\mu t_T\}$, the number of car users during the dedicated transit lane period is at capacity. Alternatively, if the number of transit users N_T takes the value $M(t_T)$, the number of car users during the dedicated transit lane period is zero. Therefore, the number of car users during the dedicated transit lane period is either zero or at capacity under a system optimum.

Alternatively, a system-optimal travel pattern can be graphically obtained as

shown in **Figure 2**. Points B and E can be determined so as to satisfy the following first order condition:

$$\frac{(\dot{W}_B - \mu)t_{AB}}{(\dot{W}_E - \mu)t_{EF}} = \frac{L}{e} \quad (7)$$

Note that t_{AB} indicates the time between points A and B, which is equivalent to $t_B - t_A$ and so on. Also, \dot{W}_B indicates the slope of the wish curve at point B, and so on.

2.4. System Optimum Pricing

It has been proven that the following time-dependent prices for car and transit result in the user equilibrium travel pattern being the system-optimal travel pattern:

$$\begin{cases} \dot{C}_C(t) = e & \text{for } t \in (t_A, t_B) \\ \dot{C}_C(t) = -L & \text{for } t \in (t_E, t_F) \\ -L < \dot{C}_C(t) < e, & \text{otherwise} \end{cases} \quad (8)$$

$$\begin{cases} C_T(t) = Z_C - Z_T + C_C(t) & \text{for } t \in (t_B, t_E) \\ C_T(t) > Z_C - Z_T + C_C(t), & \text{otherwise} \end{cases} \quad (9)$$

Note that any vertical translation of the transit and car price curves satisfies Equation (8) and Equation (9) and will result in the same system-optimal travel pattern.

3. User Equilibrium of Cars and Transit in the Morning Commute, Considering Transit Capacity

This section considers how a change in transit capacity affects the user equilibrium travel pattern with the general S-shaped wish curve. Note that the transit capacity is described as ω hereafter.

When transit capacity is limited, transit users will have to pay an additional cost, τ . Therefore, the maximum acceptable delay when travelling by car increases by τ as below.

$$(Z_T + \tau) - Z_C = T + \tau \quad (10)$$

The additional cost τ is caused by a decrease in the transit capacity and can be interpreted as an early arrival penalty, a queuing delay (at a stop) or a late arrival penalty. However, a queuing delay for transit users does not affect the travel time of car users. The user equilibrium travel pattern can be illustrated in **Figure 3**. Note that the superscript “ W ” here means “with transit capacity constraints”. Because the value of τ is determined such that the slope of GE is equal to the transit capacity, τ varies according to the shape of the desired arrival distribution and the transit capacity. Then, the following proposition holds. From the equilibrium condition, the maximum early arrival time and the maximum delay time are respectively τ/e and τ/L as shown in **Figure 3**.

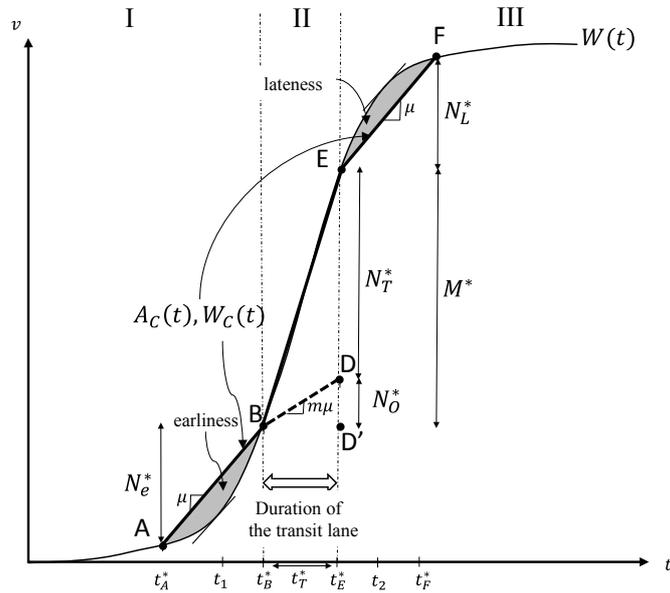


Figure 2. System optimum for a bottleneck with a transit alternative.

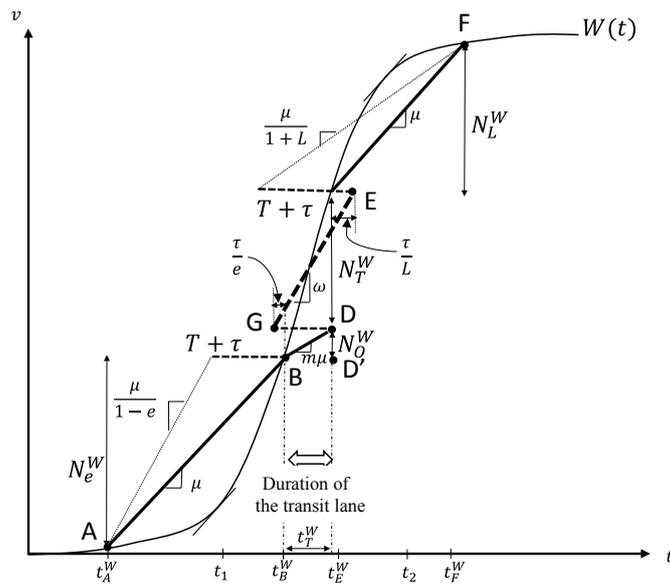


Figure 3. User equilibrium for a bottleneck with a transit alternative in case of Z-shaped wished curve.

The following proposition justifies the transit cost increase when transit capacity is limited.

Proposition 1

Under a user equilibrium travel pattern, the maximum delay when travelling by car, T , increases if and only if the number of transit users per unit time decreases.

Proof

1) Proof of “if” part of the proposition

From Proposition 1 of Gonzales and Daganzo [9], N_T^W decreases as T in-

creases, which implies that the slope of GE becomes smaller. This means that the number of transit users per unit time decreases.

2) Proof of “only if” part of the Proposition

From geometrical consideration, if the number of transit users per unit time decreases; *i.e.* the slope of GE becomes smaller, the length of GD should increase, which means the maximum delay by car, T , increases.

Q.E.D.

From the above Proposition and Proposition 1 of Gonzales and Daganzo [9], as transit capacity decreases, both the number of early car commuters N_e^W and the number of late car commuters N_L^W increase, while both the number of on-time car commuters N_O^W and the number of on-time transit commuters N_T^W decrease. Also, from geometrical consideration, we can confirm that as the transit capacity decreases, point A moves left and point F moves right, which means that the rush starts earlier and ends later. Similarly, as the transit capacity decreases, point B moves right and point E moves left, which means that the duration of the dedicated transit lane period is shorter.

Based on the above considerations, a decrease in transit capacity leads to an increase in transit cost and consequently a decrease in the duration of the dedicated transit lane period and a decrease in social welfare; *i.e.* an increase in the total schedule penalty.

4. System Optimal for Morning Commute with Cars and Transit: Considering Transit Capacity

This section considers how a change in transit capacity affects the system-optimal travel pattern. As described in Section 2, a system-optimal travel pattern can also be obtained graphically, as shown in **Figure 2**. However, in the case of considering transit capacity, the solution shown in **Figure 2** is not feasible if the slope of BE exceeds the transit capacity. In such a case, because we assume the restricted case that users are served in the phase they desire [9], there is no option other than to adjust the duration of the dedicated transit lane period so that the slope of BE becomes smaller than the transit capacity.

Therefore, we can graphically investigate the effect of transit capacity on the system-optimal solution by analysing how transit capacity changes when the duration of the dedicated transit lane period changes. Section 4.1 considers a general S-shaped wish curve and Section 4.2 considers a Z-shaped wish curve as a special case. Section 4.2 further shows numerical examples by solving the optimisation problem.

4.1. General S-Shape Wish Curve

As described in Section 2, the number of car users during the period of the dedicated transit lane under a system optimum is either zero or at capacity. Hence, we investigate changes in transit capacity by expanding the duration of the dedicated transit lane period (*i.e.* let the start time of the dedicated transit lane Δt_1

be earlier and the end time Δt_2 be later) in the above two system-optimal cases.

4.1.1. In the Case That the Number of Car Users during the Dedicated Transit Lane Period Is at Capacity

Figure 4 shows the transit capacity before and after the duration period of the dedicated transit lane is extended. Let t_B^* and t_E^* , respectively, be the start and end times of the dedicated transit lane period under the system-optimal travel pattern before increasing the duration of the dedicated transit lane period. From geometrical considerations, $\overline{GE} = (t_T, N_T)$,

$\overline{G'G} = (\Delta t_1, \dot{W}_B \Delta t_1 - m\mu \Delta t_1 - m\mu \Delta t_2)$, $\overline{EE'} = (\Delta t_2, \dot{W}_E \Delta t_2)$. Because

$\overline{G'E'} = \overline{G'G} + \overline{GE} + \overline{EE'}$, we have

$$\overline{G'E'} = (t_T + \Delta t_1 + \Delta t_2, N_T + (\dot{W}_B - m\mu) \Delta t_1 + (\dot{W}_E - m\mu) \Delta t_2).$$

Then, the difference in the transit capacity before and after the increase in the dedicated transit lane period is given as the difference in the slope of BE and BE', which is equivalent to:

$$\begin{aligned} & \frac{N_T}{t_T} - \frac{\{N_T + (\dot{W}_B - m\mu) \Delta t_1 + (\dot{W}_E - m\mu) \Delta t_2\}}{t_T + \Delta t_1 + \Delta t_2} \\ &= \frac{\{M(t_T) - \dot{W}_B t_T\} \Delta t_1 + \{M(t_T) - \dot{W}_E t_T\} \Delta t_2}{t_T (t_T + \Delta t_1 + \Delta t_2)} \end{aligned} \tag{11}$$

where $M(t_T) \equiv N_T + m\mu t_T$ is the number of users travelling when the dedicated transit lane is provided. Because we assume a general S-shaped wish curve, the gradient in B and E (\dot{W}_B and \dot{W}_E) may be greater or smaller than the slope of BE, depending on the initial positions of B and E. Therefore, the difference in transit capacity before and after extending the period of the dedicated transit lane can take both positive and negative values depending on the initial system-optimal solution.

4.1.2. In the Case That the Number of Car Users during the Dedicated Transit Lane Period Is at Capacity

Figure 5 shows the transit capacity before and after the duration of the dedicated transit lane period is extended. Let t_B^* and t_E^* , respectively, be the start and end times of the dedicated transit lane under the system-optimal travel pattern. From geometrical considerations, $\overline{BE} = (t_T, N_T)$ and

$\overline{B'E'} = (t_T + \Delta t_1 + \Delta t_2, N_T + \dot{W}_B \Delta t_1 + \dot{W}_E \Delta t_2)$. Then, the difference in the transit capacity before and after the increase in the dedicated transit lane period is given by:

$$\begin{aligned} & \frac{N_T}{t_T} - \frac{(N_T + \dot{W}_B \Delta t_1 + \dot{W}_E \Delta t_2)}{t_T + \Delta t_1 + \Delta t_2} \\ &= \frac{(N_T - \dot{W}_B t_T) \Delta t_1 + (N_T - \dot{W}_E t_T) \Delta t_2}{t_T (t_T + \Delta t_1 + \Delta t_2)} \end{aligned} \tag{12}$$

Then, similar to the discussion above, this value can take both positive and negative values depending on the initial system-optimal solutions.

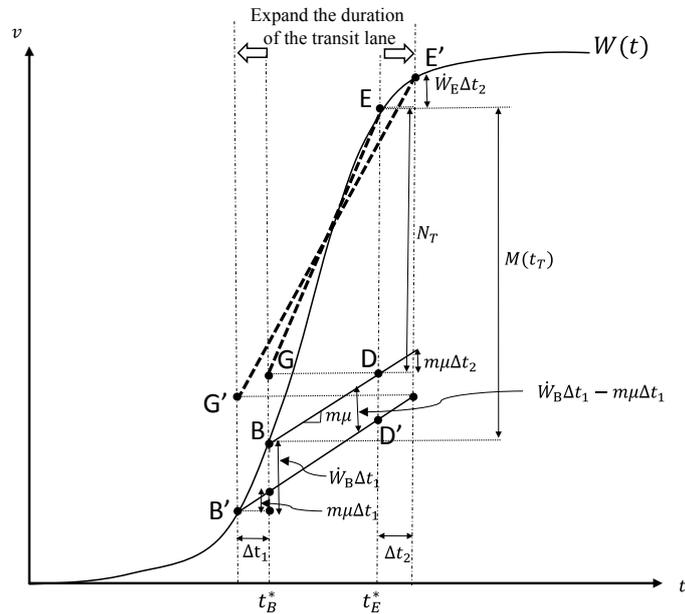


Figure 4. Change in system optimum by extending the duration of the dedicated transit lane period (in case the number of car users during the dedicated transit lane period is at capacity).

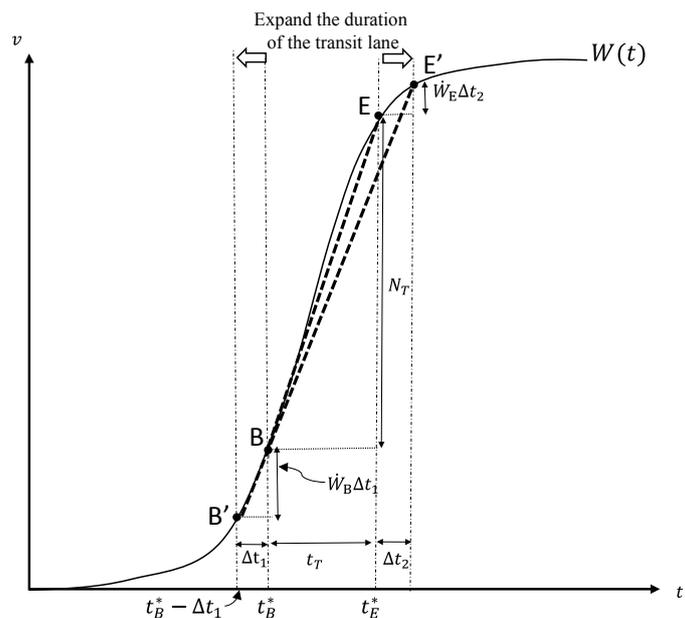


Figure 5. Change in system optimum by extending the duration of the dedicated transit lane period (in case that the number of car users during the dedicated transit lane period is zero).

To summarise so far, when the wish curve is assumed to be a general S-shape, the increase or decrease in transit capacity resulting from increasing the duration of the dedicated transit lane period differs depending on the initial system optimum. In other words, the effect on the dedicated transit lane of reducing the transit capacity differs depending on the initial system-optimal solutions.

4.2. Z-Shaped Wish Curve

Let us assume that the wish curve is Z-shaped with N users and the demand rate is λ as shown in **Figure 6** as the special case of an S-shaped wish curve.

4.2.1. Qualitative Consideration

When the wish curve is Z-shaped with a fixed demand rate λ , the difference in the transit capacity described in Sections 4.1.1 and 4.1.2 can be simplified. Therefore, the following propositions are established.

Proposition 2

Under a system optimum with the number of car users during the dedicated transit lane period at capacity, the duration of the dedicated transit lane period and the number of transit users are stable, even if transit capacity is slightly decreased.

Proof

When the wish curve is assumed to be Z-shaped, $\dot{W}_B = \dot{W}_E = \lambda$, then, from **Figure 6**, both $\dot{W}_B t_T$ and $\dot{W}_E t_T$ must be equal to the number of users traveling during the period that the dedicated transit lane is provided, $M(t_T)$. Therefore, the difference in transit capacity between before and after increasing the period of the dedicated transit lane defined in Section 4.1.1 should be 0 because the denominator of Equation (11) is $M(t_T) - \dot{W}_B t_T = M(t_T) - \dot{W}_E t_T = 0$. This implies that the duration of the dedicated transit lane period is stable if the transit capacity is slightly decreased. Then, the number of transit users is also stable.

Q.E.D.

Proposition 3

Assume that the transit capacity ω is less than the demand rate λ . Under a system optimum with the number of car users during the dedicated transit lane period being zero, the duration period of the dedicated transit lane and the number of transit users decrease if the transit capacity decreases slightly.

Proof

When the wish curve is assumed to be Z-shaped, $\dot{W}_B = \dot{W}_E = \lambda$. Therefore, the difference in the transit capacity between before and after increasing the dedicated transit lane period, shown in the denominator of Equation (12), should be negative because $N_T - \dot{W}_B t_T = N_T - \dot{W}_E t_T = (\omega - \lambda)t_T \leq 0$. This implies that the transit capacity increases if the duration of the dedicated transit lane period is extended. In other words, if transit capacity is limited, the duration of the dedicated transit lane period should be decreased to adjust for the insufficient transit capacity. Consequently, the number of transit users decreases because both the transit capacity and the duration of the dedicated transit lane period decrease. Hence, Proposition 3 is established.

Q.E.D.

Based on the above two propositions, the effects of a change in transit capacity on the duration of the transit lane period and the number of transit users are not uniform based on an initial system-optimal solution.

Let us assume a system in which $N = 10000$ desired journeys must pass a single bottleneck with capacity $\mu = 6000$ veh/h and $m = 2/3$. The car and transit cost functions are assumed to be:

$$Z_C(N - N_T) = 0.45(N - N_T) \tag{18}$$

$$Z_T(N_T, t_T) = 0.4N_T + \sqrt{45t_T N_T + 20N_T} \tag{19}$$

The other parameters are assumed to be identical to Gonzales and Daganzo [9]: $\lambda = 10000$ veh/hours, $e = 0.5$, $L = 2.0$. The value of time for all users is assumed to be $\beta = 20$ \$/h. We compare the system optimum travel pattern with several transit capacities. We compare the system-optimal travel patterns for several transit capacities. We compare system-optimal solutions by decreasing the transit capacity from 10,000 ($= \lambda$) to 6000 ($= \lambda - m\mu$).

Figure 7 shows the differences in the system-optimal charge and the duration of the dedicated transit lane period. Note that the difference in the system-optimal charge is defined as the difference between maximum and minimum system-optimal pricing, which can be obtained by $\beta L(t_F - t_E)$ according to Equation (8). As transit capacity decreases to 7000, the duration of the dedicated transit lane period decreases. On the other hand, the difference in the system-optimal charge increases as the transit capacity decreases to 7000. Furthermore, the slope of the duration of the transit lane period changes from over 7500 to 7000 to 7500. This is because the optimum value of N_T shifts from $\min(\omega t_T, \lambda t_T)$ to $(\lambda - m\mu)t_T$ when the transit capacity falls below 7500. Note again that the number of transit users should be either $\min(\omega t_T, \lambda t_T)$ or $(\lambda - m\mu)t_T$ at an optimum. For the same reason, the duration of the dedicated transit lane period and the difference in the system-optimal charge are stable if the transit capacity is less than 7000. Consequently, as shown in Figure 8, the percentage of transit users decreases and that of car users who arrive on time increases as the transit capacity decreases to 7000. However, when the transit capacity is less than 7000, the number of transit users is stable. Again, this is because the optimum value of N_T shifts from $\min(\omega t_T, \lambda t_T)$ to $(\lambda - m\mu)t_T$ when the transit capacity falls below 7500. These facts coincide with Propositions 2 and 3. Finally, Table 1 shows the cost component for each transit capacity. As the transit capacity decreases to 7000, the transit cost decreases and the car cost increases due to the changes in the respective numbers of users. Also, the schedule cost and the total cost increase slightly as the transit capacity decreases to 7000.

Table 1. Cost components for each transit capacity.

		10,000	9500	9000	8500	8000	7500	7000	6500	6000
Cost	Transit	4483.8	4265.3	4048.4	3832.8	3618.6	3405.6	2772.0	2772.0	2772.0
	Car	309.9	531.8	751.7	969.7	1185.8	1400.0	2032.7	2032.7	2032.7
	Schedule Penalty	6.3	6.9	7.4	7.9	8.4	8.9	9.9	9.9	9.9
	Total	4800.0	4804.0	4807.5	4810.4	4812.8	4814.4	4814.7	4814.7	4814.7

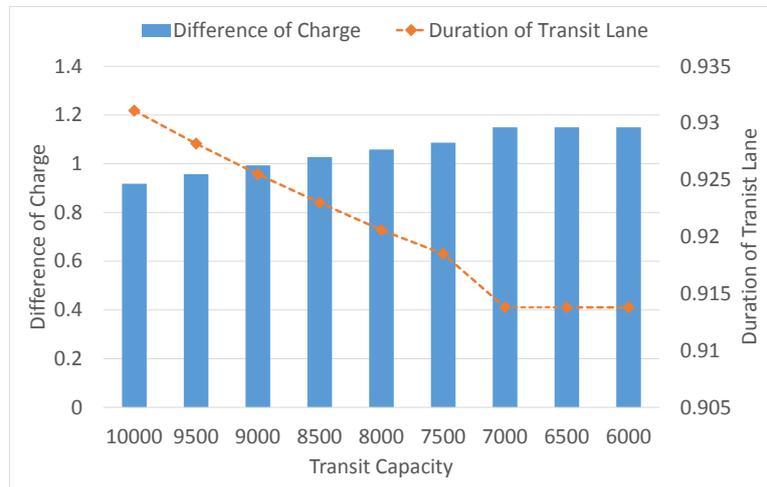


Figure 7. Difference in system optimum charge and the duration of the transit lane period for each transit capacity.

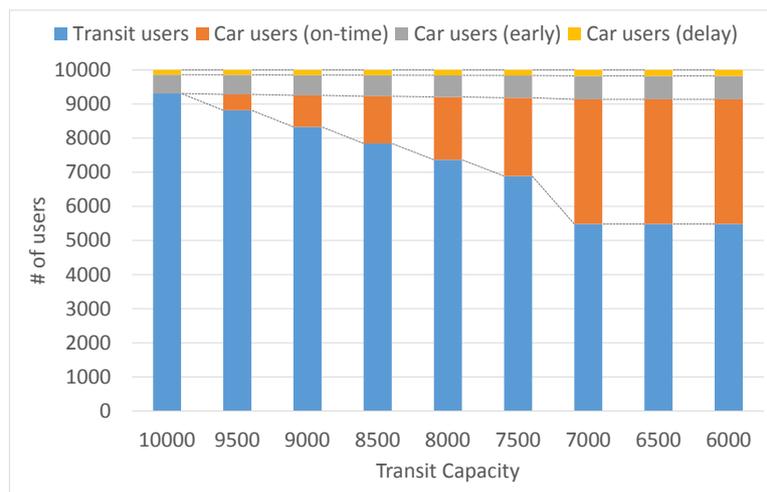


Figure 8. The number of each type of user for each transit capacity.

5. Conclusions

This paper analysed the effects of insufficient transit capacity on a dedicated transit lane measure based on the morning-commute problem, considering competitive modes of transport and finite transit capacity. We defined user equilibrium and a system optimum when the transit capacity was limited. As a result of qualitative analysis, the following findings were obtained:

- 1) Under user equilibrium, if the transit capacity decreases, the transit cost per person increases, which decreases the duration period of the dedicated transit lane and increases the number of early and delayed car commuters. This effect is smooth with regard to the reduction in transit capacity.
- 2) Under a system optimum, the effect of reducing transit capacity depends on the initial solution of the system optimum, in a general wish curve.
- 3) Under a system optimum, if the wish curve is assumed to be Z-shaped as a special case, the duration of the dedicated transit lane period and the number of

transit users decrease as the transit capacity decreases. However, if the transit capacity is below a certain level, the duration of the dedicated transit lane period and the number of transit users are stable.

We further numerically confirmed the effect of transit capacity under a system optimum with a Z-shaped wish curve. These findings can be useful in considering measures to increase transit capacity in an area where transit capacity is insufficient.

We considered only heterogeneity in the desired time to leave a bottleneck. For future study, there is room to consider additional heterogeneity among users, such as captive users and the heterogeneity of the value of time.

Acknowledgements

This work was partially supported by JSPS KAKENHI Grant Number 18H01562 and 21H01459.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Ohshima, K. and Yamamoto, K. (2021) Contributing Factors for Delays during the Morning Commute Hours and the Impact of the Spread of COVID-19 for Metropolitan Train Lines in Japan. *Journal of Transportation Technologies*, **11**, 519-544. <https://doi.org/10.4236/jtts.2021.114033>
- [2] Vickrey, W.S. (1969) Congestion Theory and Transport Investment. *The American Economic Review*, **59**, 251-260.
- [3] Hendrickson, C. and Kocur, G. (1981) Schedule Delay and Departure Time Decisions in a Deterministic Model. *Transportation Science*, **15**, 62-77. <https://doi.org/10.1287/trsc.15.1.62>
- [4] Smith, M.J. (1984) The Existence of a Time-Dependent Equilibrium Distribution of Arrivals at a Single Bottleneck. *Transportation Science*, **18**, 385-394. <https://doi.org/10.1287/trsc.18.4.385>
- [5] Arnott, R., de Palma, A. and Lindsey, R. (1988) Schedule Delay and Departure Time Decisions with Heterogeneous Commuters. *Transportation Research Record* (1197), 56-67.
- [6] Arnott, R., Palma, A. and Lindsey, R. (1994) The Welfare Effects of Congestion Tolls with Heterogeneous Commuters. *Journal of Transport Economics and Policy*, **28**, 139-161.
- [7] Liu, Y. and Nie, Y. (2011) Morning Commute Problem Considering Route Choice, User Heterogeneity and Alternative System Optima. *Transportation Research B*, **45**, 619-642. <https://doi.org/10.1016/j.trb.2010.11.003>
- [8] Sakai, K., Liu, R., Kusakabe, T. and Asakura, Y. (2017) Pareto-Improving Social Optimal Pricing Schemes Based on Bottleneck Permits for Managing Congestion at a Merging Section. *International Journal of Sustainable Transportation*, **11**, 737-748. <https://doi.org/10.1080/15568318.2017.1312646>

- [9] Gonzales, E.J. and Daganzo, C.F. (2012) Morning Commute with Competing Modes and Distributed Demand: User Equilibrium, System Optimum, and Pricing. *Transportation Research B*, **46**, 1519-1534. <https://doi.org/10.1016/j.trb.2012.07.009>
- [10] Gonzales, E.J. and Daganzo, C.F. (2013) The Evening Commute with Cars and Transit: Duality Results and User Equilibrium for the Combined Morning and Evening Peaks. *Transportation Research B*, **57**, 286-299. <https://doi.org/10.1016/j.trb.2013.06.009>