Exploring Changes in Public Transport Usage during the COVID-19 Pandemic using Smart Card Data: A Case Study from Hiroshima, Japan

Xuan Canh Do[*1], Thi Anh Hong Nguyen[2], Myo Min Than[3], Akimasa Fujiwara[4], Makoto Chikaraishi[5]

[1,2]Researchers, Hiroshima University, Japan

[3]Research student, Hiroshima University, Japan

[4]Professor, Hiroshima University, Japan

[5]Associate Professor, Hiroshima University, Japan

SHORT SUMMARY

There was a sharp decline in public transport (PT) ridership during the pandemic. It would be worth studying the changes in PT mobility during the pandemic to deeply understand PT users' behavior. This study contributes to the growing literature on PT-related studies based on smart card data. As a result, first, we explored the differences in the decrease in PT usage between regular and non-regular public ridership groups. Second, we confirmed a moderately higher reduction in ridership for the non-regular group than for the regular group regarding spatiotemporal dimensions. Third, few active cardholders in the non-regular group and few trips per day per cardholder in the regular group caused the ridership reduction. Finally, higher morning and afternoon traffic peaks are observed in the regular group during workdays.

Keywords: public transport**,** smart card data, COVID-19

1. INTRODUCTION

The COVID-19 pandemic has brought about unprecedented changes in human mobility, both within countries and across borders. Not only have many government-imposed policies to restrict and discourage people's movements, but fear about the risks of transmission has also influenced individual decision-making behaviors and actions. As a result, many researchers have focused on changes in individual mobility before and during the pandemic to understand travel behavior (Bohman et al., 2021; Ehsani et al., 2021).

Public transport (PT) has been one of the hardest-hit travel modes among transportation modes, as seen in the rapid decline in PT ridership levels when COVID-19 started to spread (Aloi et al., 2020). The main reasons are that PT is widely regarded as having a high risk of infection and the spread of diseases (Tan and Ma, 2021). For example, Brazil's number of bus users fell by more than 80%, with just 200 thousand bus passengers per day relative to the pre-pandemic level of 1.36 million bus passengers per day during the pandemic peak (Fumagalli et al., 2021). Therefore, focusing on the behavior of PT users during the pandemic is of great importance.

The pandemic poses challenges for PT on both supply and demand sides in terms of current and future travel behaviors and the environmental efficiency of public transport systems (Vitrano, 2021). Although many studies examine mobility changes during the pandemic, only a few studies focus on

PT users' behavior during and after the COVID-19 pandemic. Excepted for review studies, some studies are based on big data extracted from public transit systems (Almlöf et al., 2021; Fumagalli et al., 2021).

Although existing work has examined some facets of the pandemic's effects on PT, this paper aims to address these four research gaps, including (i) how to identify a breakpoint or milestone to split data into pre and post COVID-19, (ii) differences between traveler groups (regular and non-regular public ridership groups), (iii) detailed decrease in ridership derived from a decrease in the number of daily active PT users versus the number of trip per day by each active PT user, (iv) and changes in the morning and evening peak traffic from a temporal perspective.

2. DATA

Dataset used in this study is PASPY smart card data (including all buses, streetcars, ferries, and the Astram line) in Hiroshima.

After excluding invalid trip records, 153,762,115 trips remained with 2,949,764 cardholders in the study period from January 1, 2019, to August 31, 2020 (i.e., 406 workdays and 203 holidays, including weekends, national holidays, and Obon summer holidays). Every registered daily PT trip record trip is measured between tap-in time and tap-out time by a cardholder in a PT vehicle for a specific time on a specific day. It is generated and counted every time a passenger makes a tap-out to get out of a PT vehicle or a rail station gate, even if she/he makes a transfer between different PT modes. It contains several attributes used in this study (e.g., unique card ID number, boarding and alighting date and time, boarding and alighting stops, basic fare, actual fare, and the number of riders on that trip, to name just a few).

3. METHODOLOGY

The first research gap was tackled by investigating seasonal variations across years to overview changes before and after COVID-19. Furthermore**,** to get more insight, we split trip counts into different groups by a set of features, such as holiday/workday and before/during COVID-19.

First, a set of time change-points was generated to be potentially chosen as a breakpoint between before and after COVID-19. In this study, the optimal number of change-points was automatically determined using an approach proposed by Zou et al. (2014). With the given breakpoints, a proper breakpoint was chosen to split our PASPY dataset to assess the relative change during COVID-19 (in 2020) with the normal seasonal variations in 2019 (a reference base). This split would evaluate to what extent the overall decrease in trip demand is caused by a fall in active cardholders or fewer trips per cardholder (i.e., the third and four research gap).

Additionally, many methods that can classify cardholders into regular or non-regular groups are considered to fill the second research gap. However, we followed a study by Fujiwara et al. (2021) and employed a simple binary logistic regression model considering sampling weight w_i to serve the above purpose. The objective variable is a binary choice of whether a cardholder i made a trip(s) at the day t. Thus, there are 2,949,764 models estimated for all cardholders. Each model with $n = 6$ explanatory dummy variables (holiday and weekdays) and 609 samples was estimated using the formulation in **Equation 1**:

$$\log \frac{{p_{it}}^{w_i}}{{p_{it}}^{w_i} + (1 - p_{it})^{w_i}} = \beta_0 + \sum_{j=1}^{n} \beta_j x_{jt} \quad (1)$$

where

x_j : the explanatory variable j (j = 1:n).
β_0 : the estimated constant coefficient.
β_j : the estimated coefficient j corresponding to the explanatory variable j.

The significance of explanatory variables will not be discussed because nearly 3 million models were estimated, and it is impossible to show all models' estimation results in this paper.

After completing these model estimations, we predicted the individual probabilities of trip generation for 609 days and took an average probability for each cardholder. The cardholder will be treated as a "non-regular" or "regular" PT user whether she/he has a probability less than 0.5 or not. As a result, the "non-regular" group consists of 2,840,492 cardholders (approximately 92.51% of total cardholders), while that of the "regular" group is 213,306 (approximately 7.49% of total cardholders), accounting for 76,940,690 and 76,821,425 trips (i.e., both are approximately 50% of total trips), respectively.

4. RESULTS

Seasonal variation in public transport usage across years

Since PASPY data contains users who may only use PT service for a short period, it is insufficient to review seasonal variations in public transport usage before and during the pandemic. Thus, we investigated the total ridership seasonal variation between 2019 and 2020 in **Figure 1**.

The upper part of **Figure 1** shows the weekly seasonal trend, with demand on weekends dropping around 50% in the regular group and around 30% in the non-regular group. Moreover, the regular group's trip demand was a bit lower in the first quarter of 2019, and it became stable up to February 2020 before starting to jump down from the beginning of March 2020. In contrast, in the non-regular group, it was stable in the first quarter of 2019, followed by a gradual decrease from the middle of April 2019 and then a plunge from March 2020. There are substantial decreases because of long holiday periods (e.g., New Year holidays, Golden week at the beginning of May, and Obon holidays in the middle of August).

Notably, there is an increase in the regular group's trip demand from the beginning of April 2019. In particular, students, who belong to the regular group, start to come back to their school after the spring school holidays. The non-regular group also reduced their activities as all spring vacation events had been gradually stopped. In line with public transport usage patterns worldwide, we also see a big plunge following the confirmation of the first infected coronavirus case in Hiroshima City on March 6, 2020. However, the decrease was already noticed a few days before that time, probably because a part of the population heard the news about the first COVID-19's death in Japan (February 13, 2020) and the government's request for all schools across the nation to close temporarily (February 27, 2020). This significant decrease at this period was confirmed using the change-point detection method. Then we denoted the day of March 1, 2020, as a breakpoint to split our study period into two periods: before and during COVID-19.

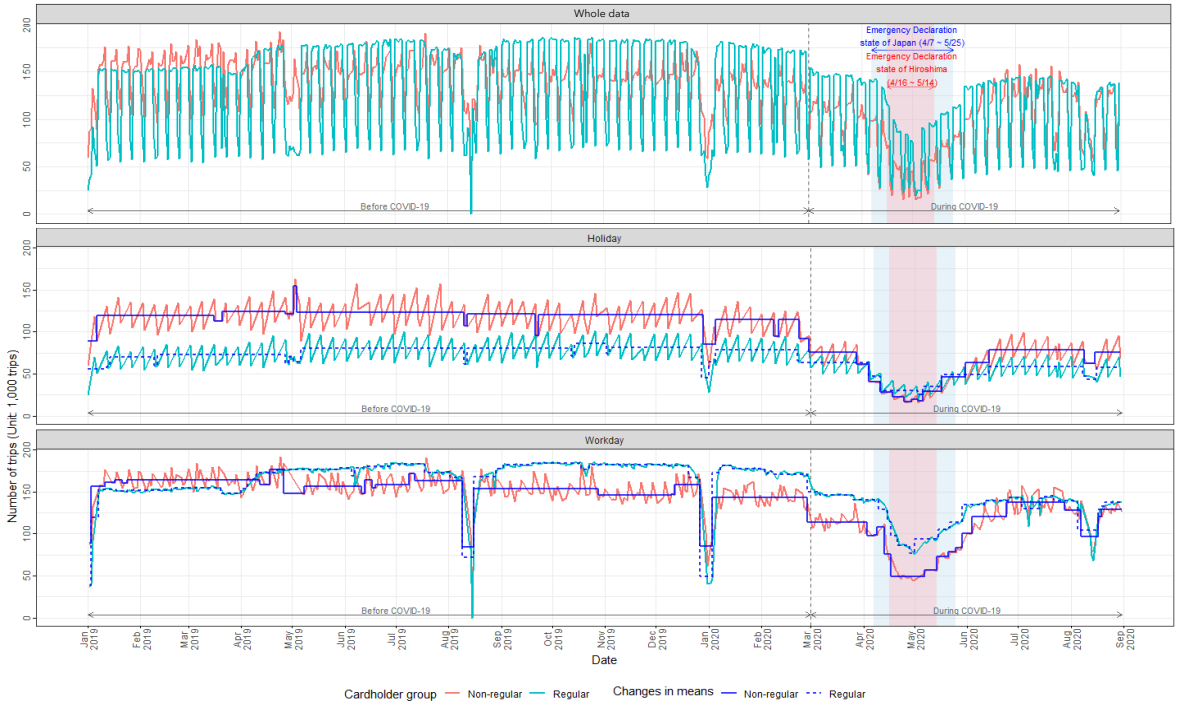

Figure 1 Seasonal variation in public transport usage for the whole dataset (upper), holidays (middle), and workdays (lower)

More specifically, in contrast to holidays, the ridership levels on workdays in the regular group were higher than the non-regular group, except in the first quarter of 2019. Regarding the demand during the holidays, the regular group is lower than the non-regular group, except during Hiroshima and Japan's emergency declaration period.

Public transport usage before and during COVID-19 periods

The ratio of the number of trips between the two compared periods (March 1 to August 31, 2020, and March 1 to August 31, 2019) is calculated to examine the relative changes in public transport usage.

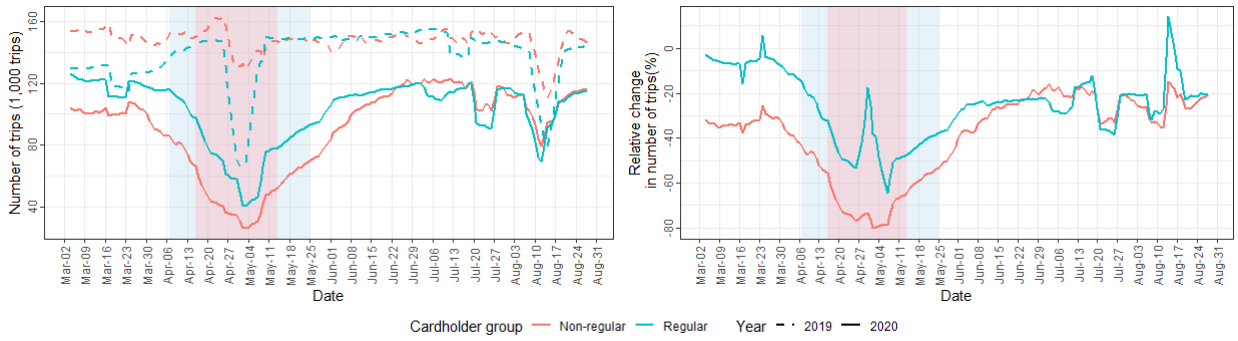

Figure 2 Day-to-day trip variation (left) and its relative change compared to 2019 (right)

Day-to-day variation comparison

Figure 2 compares the day-to-day variation of regular and non-regular groups between 2020 and 2019 by considering the 7-day centered moving average calculation which helps avoid the effect of seasonal variation on our assessment.

In the emergency declaration period, the non-regular group's demand steadily decreased to around 20-30% of the previous year in the first half. It increased gradually to 50% of the 2019 reference values in the second half. The regular group's demand fluctuated with a decrease to around 50%, a big jump to approximately 20%, and a plunge to 35% before slowly recovering and remaining at 50%.

During the post-emergency declaration period, demand rose slowly to around 80%. It fluctuated to around 70% compared with the reference values in regular and non-regular groups before the Obon summer period. Finally, the regular group's demand rose from 70% to 110% in the middle of the Obon summer holidays. One explanation could be that the non-regular group reduced their discretionary activities and refrained from using PT service due to the fear of coronavirus infection. At the same time, the regular group also reduced their demand but maintained their necessary activities as usual.

Active public transport cardholders and trips per cardholder comparison

Table1 Overall number of trips, active cardholders, trips per card, and relative changes

	No. of trips		No. of active cards		No. of trips per card		Relative change (compared with 2019)		
							No. of trips	No. of active cards	No. of trips per card
Year	**2019**	**2020**	**2019**	**2020**	**2019**	**2020**	**2020**	**2020**	**2020**
Regular	25,095,863	18,790,675	107,682	105,624	233	178	-25%	-2%	-24%
Non-regular	27,044,528	16,760,600	1,709,166	924,041	16	18	-38%	-46%	15%
Total	52,140,391	35,551,275	1,816,848	1,029,665	n/a	n/a	-32%	-43%	n/a

In **Table1**, with the baseline is trip demand before the COVID-19 period (2019), trip demand during COVID-19 (2020) decreased by an average of approximately 32%, whereby the regular group and non-regular group demand dropped about 25% and 38%, respectively. In contrast, the number of active cards fell an average of 43%, with 2% and 46% in the regular and non-regular groups, respectively. Thus, the main reason for the reduced trip demand derives from fewer active cardholders in the non-regular group and fewer trips per card in the regular group.

Hourly trip variation by boarding and alighting time slots

Briefly, **Figure 3** shows an average trip demand per day per hour of the day. Additionally, in both regular and non-regular groups, the change could only be seen in reducing the average number of trips per day rather than making changes in the behavior of boarding/alighting time. In other words, the morning and evening peak hours did not change during the pandemic.

On workdays, non-regular groups saw the most considerable decrease in trip demand, about 26-28%, and 34.8-40%, during the morning (6 AM-9 AM) and evening peak (3 PM-7 PM), respectively. This difference was not prominent as the regular group, in which morning and evening peaks fell 25-27.2% and 26-32.8% in ridership, respectively. This decrease was even worse on holidays, which may be because the non-regular group cut down unnecessary trips going out (e.g., going to festival events, parties, or visiting their friends) as much as possible during the COVID-19 period. The remaining trips were necessary daily life trips for survival, such as going to supermarkets or visiting hospitals.

Notably, on the workdays of the regular group, there was no distinct evening peak (5 PM and 6 PM) of the boarding time. In the non-regular group, the evening traffic peak seems to disappear during the pandemic as non-regular PT users are less likely to be restricted by working hours and can adapt boarding times in the evening to return home earlier because of the reduced operating hours of some business facilities which was another coronavirus prevention measure.

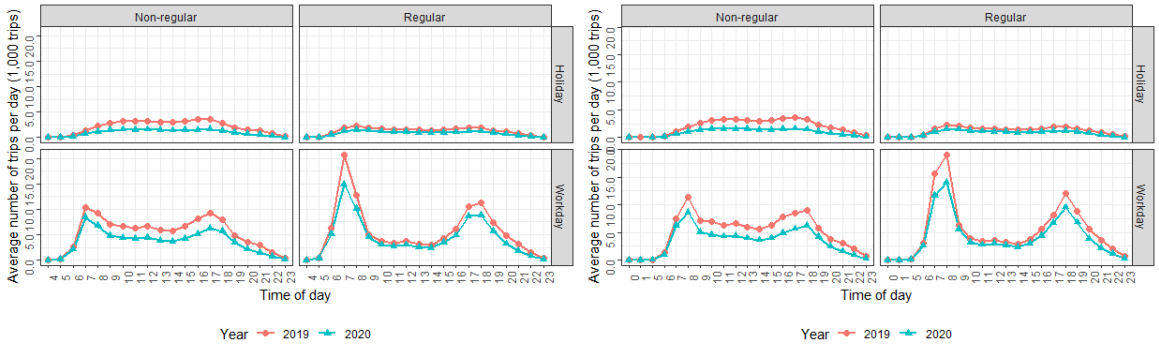

Figure 3 Hourly trip variation by boarding time (left) and alighting time (right)

5. CONCLUSION

The COVID-19 pandemic has changed mobility patterns and affected many sectors, one of which is public transport. Many recent studies have examined the impacts of the COVID-19 pandemic on travel mobility, and almost all literature shows a sharp decline in PT ridership levels. Understanding PT users' behavior amid the pandemic is helpful to examine changes in PT mobility before and during the pandemic. This study contributes to the growing literature on PT-related studies based on SCD during the pandemic.

This study provides deeper insights into reducing PT ridership during the pandemic by addressing unknown aspects. First, it explored the differences in the decrease in PT usage between regular and non-regular public ridership groups. Second, it examined the extent to which the decrease in PT ridership is represented by the number of daily active PT users versus the reduction in the number of trips per day by each active PT user. Third, it examined the dramatic reduction in PT ridership in more detail, taking into account peak traffic, morning and afternoon peaks, and other temporal and scale dimensions changes.

Also, much higher morning and afternoon traffic peaks are observed in the regular group than non-regular groups during workdays. Interestingly, the evening traffic peak in the non-regular group seems to disappear during the pandemic as non-regular PT users are less likely to be bound to work

hours at offices. They can return to their homes before peak hours and may also be responding to policies related to the early closure of restaurants and amusement facilities.

Given the contribution mentioned above of this study, some limitations and future research directions should be noted. First, as defined in this study, each "trip" may be an incomplete trip (e.g., transfer trip), so the boarding/alighting stops are not always "true" origins/destinations. Therefore, a complete trip consisting of a chain of trip segments could be identified and analyzed. Second, the literature shows that PT usage is highly correlated with socioeconomic attributes and geographic area. A further examination of this relationship should be done in future work.

ACKNOWLEDGMENTS

This study was funded by the research project "Short-term travel demand prediction and comprehensive transport demand management (principal investigator: Assoc. Prof. Makoto Chikaraishi)", supported by the Committee of Advanced Road Technology, the Ministry of Land, Infrastructure, Transport, and Tourism of Japan (MLIT). We thank Jenny Yamamoto for proofreading this paper. Finally, we thank Yu Fujiwara for explaining his research paper.

REFERENCES

Almlöf, E., Rubensson, I., Cebecauer, M., Jenelius, E. 2021. Who continued travelling by public transport during COVID-19? Socioeconomic factors explaining travel behaviour in Stockholm 2020 based on smart card data. *European Transport Research Review*, Vol. 13, No. 1, pp. 1-13.

Aloi, A., Alonso, B., Benavente, J., Cordera, R., Echániz, E., González, F., Ladisa, C., Lezama-Romanelli, R., López-Parra, Á., Mazzei, V., Perrucci, L., Prieto-Quintana, D., Rodríguez, A., Sañudo, R. 2020. Effects of the COVID-19 Lockdown on Urban Mobility: Empirical Evidence from the City of Santander (Spain). *Sustainability*, Vol. 12, No. 9, pp. 3870.

Bohman, H., Ryan, J., Stjernborg, V., Nilsson, D. 2021. A study of changes in everyday mobility during the Covid-19 pandemic: As perceived by people living in Malmö, Sweden. *Transport policy*, Vol. 106, No., pp. 109-119.

Ehsani, J. P., Michael, J. P., Duren, M. L., Mui, Y., Porter, K. M. P. 2021. Mobility Patterns Before, During, and Anticipated After the COVID-19 Pandemic: An Opportunity to Nurture Bicycling. *American Journal of Preventive Medicine*, Vol. 60, No. 6, pp. e277-e279.

Fujiwara, Y., Urata, J., Fujiwara, A., Chikaraishi, M. (2021). Exploring the Diversity of Public Transport Usage with Smart Card Data and Its Associations with Accessibility and Land Use Patterns. Presented at the The 14th International Conference of Eastern Asia Society for Transportation Studies. Hiroshima (Japan), September 12-15, 2021. Accepted.

Fumagalli, L. A. W., Rezende, D. A., Guimarães, T. A. 2021. Challenges for public transportation: Consequences and possible alternatives for the Covid-19 pandemic through strategic digital city application. *Journal of Urban Management*, Vol. 10, No. 2, pp. 97-109.

Tan, L., Ma, C. 2021. Choice behavior of commuters' rail transit mode during the COVID-19 pandemic based on logistic model. *Journal of Traffic and Transportation Engineering (English Edition)*, Vol. 8, No. 2, pp. 186-195.

Vitrano, C. (2021) COVID-19 and Public Transport: A Review of the International Academic Literature [Lecture]. Sweden's national centre for research and education on public transport (K2), unpublished.

Zou, C., Yin, G., Feng, L., Wang, Z. 2014. Nonparametric Maximum Likelihood Approach to Multiple Change-Point Problems. *The Annals of Statistics*, Vol. 42, No. 3, pp. 970-1002.