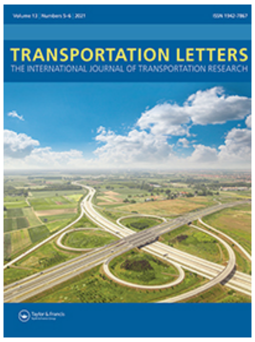

Transportation Letters
The International Journal of Transportation Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ytrl20

How did the COVID-19 pandemic impact traveler behavior toward public transport? The case of Athens, Greece

Athanasios Kopsidas, Christina Milioti, Konstantinos Kepaptsoglou & Eleni I. Vlachogianni

To cite this article: Athanasios Kopsidas, Christina Milioti, Konstantinos Kepaptsoglou & Eleni I. Vlachogianni (2021) How did the COVID-19 pandemic impact traveler behavior toward public transport? The case of Athens, Greece, Transportation Letters, 13:5-6, 344-352, DOI: 10.1080/19427867.2021.1901029

To link to this article: https://doi.org/10.1080/19427867.2021.1901029

TRANSPORTATION LETTERS
2021, VOL. 13, NOS. 5–6, 344–352
https://doi.org/10.1080/19427867.2021.1901029

Taylor & Francis
Taylor & Francis Group

How did the COVID-19 pandemic impact traveler behavior toward public transport? The case of Athens, Greece

Athanasios Kopsidas[a], Christina Milioti[b], Konstantinos Kepaptsoglou[a] and Eleni I. Vlachogianni[c]

[a]School of Rural and Surveying Engineering, National Technical University of Athens, Athens, Greece; [b]Department of Civil Engineering, University of West Attica, Athens, Greece; [c]School of Civil Engineering, National Technical University of Athens, Athens, Greece

ABSTRACT

The COVID-19 outbreak led to significant changes in daily commuting. As lockdowns were imposed to metropolitan areas throughout the globe, travelers refrained heavily from using public transport, to maintain social distancing. Based on data from Athens, Greece, this paper investigates the anticipated, post-pandemic behavior of travelers with respect to public transport use. Focus is given on analyzing those factors that affect post-pandemic recovery time of public transport users, i.e. the time travelers would refrain from using public transport, following a gradual exit from the pandemic outbreak and relaxation of lockdowns. The analysis is performed using both a clustering algorithm and a discrete duration model. Both methodologies highlighted the fact that the frequency of using public transport before the pandemic along with the travelers' age, influence their behavior in terms of recovery time. Results from the discrete duration model suggest also that self-employed and travelers who mostly use private vehicles, are less likely to use public transport after the outbreak. Concerning the psychological factors that shape COVID-19 safety-related perceptions that affect public transport use, travelers who would be willing to use protection gear when traveling with are also less likely to return to public transport. Findings of this study could be useful for policy making, suggesting that efficient marketing strategies toward promoting public transport usage in a post-pandemic era should focus on travelers with specific socio-demographic and travel characteristics.

KEYWORDS

COVID-19; Public transport; Traveler behavior; Discrete duration model; Cluster analysis; Travelers groups

Introduction and background

The fight against the COVID-19 pandemic drives countries into deploying mobility restrain policies, such as lockdowns, social distancing measures, tele-working, and travel bans (Tu et al. 2020; Hale et al. 2020; Roosa et al. 2020; Dalton, Corbett, and Katelaris 2020; Quilty et al. 2020; Boldog et al. 2020). Public transport is among those sectors strongly affected by the pandemic: services are either restricted, forced to adapt to social-distancing mandates, or even temporarily terminated (Jung et al. 2020; Sahin et al. 2020; Anjum 2020; Rodríguez-Morales et al. 2020). Published evidence on the effectiveness of service restrain measures are, however, region/country specific and as such, found to be contradicting across studies (Douglas et al. 2020; Lee and You 2020; Tian et al. 2020; Muller et al. 2020; Zheng et al. 2020; Almagro and Orane-Hutchinson 2020; Askitas, Tatsiramos, and Verheyden 2020; Musselwhite, Avineri, and Susilo 2020). Indeed, the level of public transport usage, the culture of travelers and so on, probably may or may not affect the performance of service restraints in mitigating the pandemic. The impact on ridership is on the other hand found to be rather clear: travelers fear infection, spend less on public transport, and tend to abandon transit in favor of private vehicles and taxis (Gerhold 2020; Honey-Roses et al. 2020; De Vos 2020; Belot et al. 2020; Baker et al. 2020, 2020; Andersen et al. 2020). Factors explaining that attitude include safety/risk perception, frequency of use, traveler profession and psychological attributes (Belot et al. 2020; Tan and Ma 2020; Dong et al. 2021). Most importantly though, many of the abovementioned studies report that COVID-19 aftermath is expected to affect post-pandemic traveler behavior; for example, according to Przybylowski, Stelmak, and Suchanek (2021), one out of four Polish commuters will refrain from riding public transport in the future.

Worldwide, social distancing, protection of public health, and the fear of infection, unavoidably hinder the use of public transport. Indeed, there is a clear and straightforward contradiction between promoting social distancing and the nature of public transport (often referred to as 'mass transport'). Nevertheless, a return to normal life activities will alleviate social-distancing and therefore should be accompanied by a gradual recovery in the demand for public transport, to support both sustainability and the viability of public transport systems. It is therefore critical to understand the anticipated behavior and attitude of public transport travelers, in the post-pandemic era.

In this context, this paper investigates the post-pandemic behavior of travelers toward public transport, using data from an online survey, undertaken in Athens, Greece, during the early stages of the country's first COVID-19 lockdown period (March 2020 – May 2020). At that time, strict, nationwide measures were imposed, and led to a national, seven – week lockdown. During that lockdown, social interactions and non-essential travel were banned and teleworking was imposed to several sectors of the economy. The lockdown led to a reduction of daily urban trips, while allowed occupancy of public transport was drastically reduced and the government suggested that travelers use their private vehicles in cases of essential trips. As such, the paper focuses on identifying groups of public transport travelers with similar, post-pandemic behavior and on analyzing those factors that affect recovery time of public transport users, i.e. the time in which travelers would refrain from using public transport, following the return to some sort of normal life activities.

CONTACT Konstantinos Kepaptsoglou kkepap@central.ntua.gr School of Rural and Surveying Engineering, National Technical University of Athens, Iroon Polytechniou Str, 15773, Zografou Campus, Athens, Greece

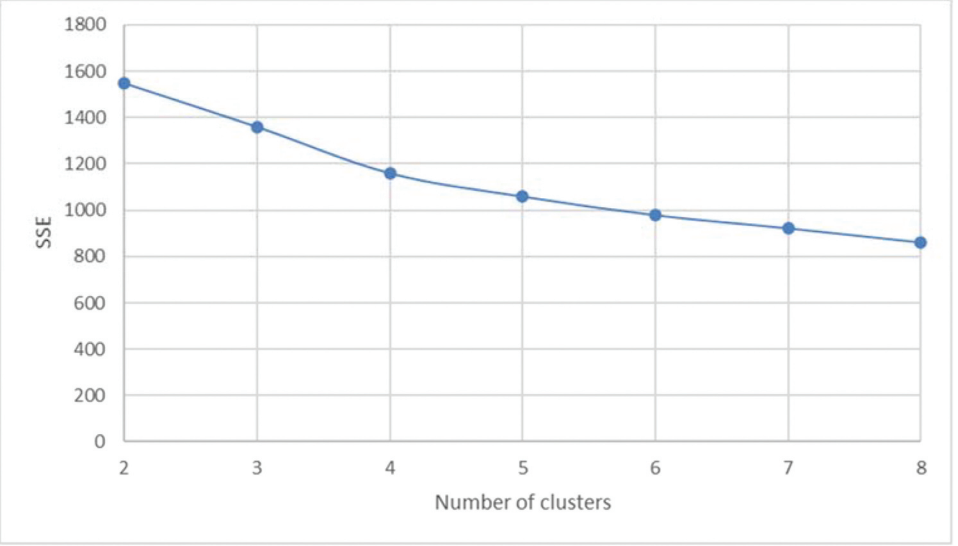

Figure 1. Elbow criterion for determining the optimal number of clusters.

The remainder of the paper is structured as follows: the data collection effort along with some fundamental descriptive statistics are presented in the next section. A cluster analysis follows, attempting to reveal groups of travelers expected to have a consistent behavior toward public transport in the post-pandemic era. Subsequently, post-pandemic recovery time of public transport travelers is explored using a discrete duration model. Finally, findings are discussed along with policy implications, and the conclusions derived from the analysis are put forward.

Data collection

Survey methodology

Data was collected through an on-line questionnaire survey, addressing Athens public transport travelers. It is noted that the Athens public transport system consists of three metro lines, two LRT lines and about 300 bus lines, and served about 3 million daily travelers before the pandemic. The questionnaire included 32 questions, structured in three groups, covering information on travel characteristics, respondents' perception about the pandemic and protection measures, and socio-demographic information (gender, age and so on). Most of the questions were based on multiple choice and Likert-scale rating, while there were a few open-type questions, such as age. The required time to complete the questionnaire was approximately ten minutes. Stratified random sampling was used for data collection, so that gender and age could be represented in line with the population census data of 2011 (Hellenic Statistical Authority 2020, July, 13). Answering to most survey questions was obligatory, except for those related to public transport usage during the pandemic, as many respondents could not or refrained from using public transport at that period, these questions were not applicable to them.

Survey implementation

The online survey was undertaken from April to May 2020; it was built on the SurveyMonkey™ platform and shared via e-mails, social media platforms and popular websites. In total, 422 valid questionnaires were collected, of a total of 463. Partially completed questionnaire was removed from the sample (41 in total), and a response ratio of 91.14% was achieved (422 valid out of a total of 463 questionnaires).

Methodology

Cluster analysis

A cluster analysis was implemented to identify profiles of distinct groups of respondents, with respect to their anticipated, post-COVID-19 behavior. A k-means clustering algorithm was exploited for that purpose: the criterion defining the assignment of the data to each cluster was the minimum Euclidean distance between the numerical values of the observations' attributes and the centers of the clusters (Wilkin and Huang 2007). For calculating this distance, an n-dimensional space was exploited, where n is the number of variables selected for clustering (Kanungo et al. 2002). Five clustering variables derived from the survey were used for clustering: frequency of using public transport prior to the pandemic, family income, age, household size and recovery time to the public transport system.

Before clustering, the data underwent pre-processing, for the quality of the results to be guaranteed: data were normalized so that Euclidean distances should be unbiased, and outliers were removed (Patel and Mehta 2011; Virmani, Taneja, and Malhotra 2015). Normalized values were calculated as follows: the difference between each observation and the mean of all observations was divided by the standard error. The Elbow criterion was then applied to determine the optimal number of clusters (Madhulatha 2012; Kodinariya and Makwana 2013): the sum of squared errors (SSE) of clustering was calculated for every outcome deriving from a clustering with k clusters ($k = 2, 3, \ldots, 8$) (Tan, Steinbach, and Kumar 2013). The optimal number of clusters, given that the fewer the better, is the one providing a relatively low SSE, meaning that a larger number of clusters reduces SSE at a slower pace. This effect is extracted by plotting SSE for the different numbers of clusters and determining the point, in which a sharper angle is created (like an elbow). In this analysis, the optimal number of clusters was found to be 4, according to the Elbow criterion, as depicted in Figure 1.

Recovery time modeling

A discrete duration model is exploited for modeling recovery time, which is the time-period, during which passengers will refrain from using public transport. Hazard-based duration models are often used for the analysis of the elapsed time till the occurrence of a specific event (Haque and Washington 2015). In such models, time could be a continuous variable taking any non-negative value or a discrete one, that is, only specific values can describe the variable of time (Tekle and Vermunt 2012). In practice, there is limited knowledge on the exact point in time when an event occurs, although there is confidence about an interval within which that event takes place. Thus, discrete time models are more convenient to use on such occasions (Steele 2011). In the context of transportation modeling, discrete duration models have been exploited by Milioti et al. (2019) and Rashidi and Mohammadian (2011). Assuming that T is a discrete variable representing time, f is a probability function of T and

$$\mathrm{F(t)} = \mathrm{P(T < t)} = 1 - \mathrm{P(T \geq t)} \tag{1}$$

Then the hazard function is given by Eq.2 (Hojati et al. 2013).

$$\mathrm{h(t)} = \frac{\mathrm{f(t)}}{1 - \mathrm{F(t)}} \tag{2}$$

or

$$\mathrm{h(t)} = \mathrm{P(T = t/T \geq t)} \tag{3}$$

where h(t) is the probability an event occurs within a time interval, given that this has not occurred until the beginning of the interval (Washington et al. 2020). When there are other possible time dependent covariates, for example x, Eq. 3 is formulated as (Fahrmeir and Wagenpfeil 1996):

$$\mathrm{h(t/x)} = \mathrm{P(T = t/T \geq t, x)} \tag{4}$$

These covariates, representing factors that affect the probabilities of the incidents, tend to multiply the baseline hazard function h0(t) by a factor equal to exp(βX) where β stands for a vector of parameters to be estimated (Washington et al. 2020). A common specification for the hazard function is the logistic one (Jenkins 1995). Taking logs gives:

$$\log \mathrm{h(t/x)} = \alpha + \beta \mathrm{X} \tag{5}$$

where a = log(h0(t)) is the logit of the underlying hazard and βX is the effect of the covariates on hazard logit. Discrete Hazard-based duration models can be estimated using logistic regression (Milioti et al. 2019). Dummy variables are created for each time interval taking binary values, which demonstrate whether this time interval is pointed or not. At the same time, another dummy indicator y is created taking the value of 1 if the event occurred in the interval pointed or otherwise, they take the value of 0. For each observation occurring at ti, a number of i pseudo-observations are created, unless the event happened during the last interval, and therefore y has already been censored, using a right-censoring technique (Steele and Washbrook 2013). Along with the pseudo-observations, covariates vector x is duplicated each time. Consequently, the model can be estimated treating yij, where j is the number of actual observations, as Bernoulli observations with a probability given by the hazard function (Milioti et al. 2019).

Data analysis and modeling

Descriptive survey results

The socio-demographic attributes of the respondents were gender, age, annual family income, car availability, marital status, profession, and household size. Descriptive statistics are presented in Table 1. The distribution of all socio-demographic attributes of the sample is in general consistent with that of Athens public transport users. As can be seen, over 80% of the respondents had the option of using a car for their daily transportation needs.

Travel characteristics of the respondents included their (pre-pandemic) primary mode of travel, trip purpose, frequency of public transport usage and the type of fare used. Private cars were the most popular choice among all modes used for commuting, followed by the metro system and buses. Respondents stated that work was by far the main, pre-pandemic trip purpose. As for the frequency of public transport use, that was in general found to be high. Last, single ride fares were used by most respondents. More details on travel characteristics are given in Table 2.

Survey participants were asked to respond (a) on how informed they were about the COVID-19 disease, (b) to what extent they trusted other travelers for following anti-pandemic rules and protection measures, (c) whether they feared of getting infected, (d) how safe they felt using public transport, (e) their perception on the possibility of getting infected while using public transport, and (d) to what extent they adopted social distancing and other protection measures. The respondents claimed that they were rather informed and quite fearful of getting infected by COVID-19. About 64% of them considered themselves informed as giving a four out of five rating in terms of the Likert-scale, while another 23% gave a top rating. Concurrently, a little less than half of them (43.6%) scored a 4/5 with respect to fear of COVID-19 infection. However, fear was a significant factor among the respondents, since 90% of the answers belong to the upper levels of the 1–5 scale. At the same time, few respondents trusted fellow travelers for maintaining COVID-19 safety rules. Specifically, 84.3% of them trusted other travelers to an average extent or did not trust them at all. Overall, respondents did not feel comfortable using public transport whatsoever, as they considered it very probable to get infected during a public transport ride. Most participants felt in danger or not very safe in using public transport.

Furthermore, a group of three questions, on the extent that respondents obey safety rules against a COVID-19 infection provided an index of the level of compliance with these rules. A 5-point Likert scale was used for that purpose (Weijters, Cabooter, and Schillewaert 2010). Specifically, related questions concerned the level of daily trips limitation, social-distancing and personal hygiene. For evaluating the internal consistency of this set, the Cronbach-a value was used and demonstrated a minimally accepted consistency, reaching the value of 0.668 (Cortina 1993). Cronbach-a value constitutes a metric for evaluating the internal consistency of a test consisting of different parts (Tavakol and Dennick 2011). Internal consistency indicates the level at which all parts serve the purpose of measuring the same concept. It can take values between 0 and 1. Table 3 presents statistical properties of this group of questions.

Recovery time, that is the time duration that respondents stated that they would refrain from using public transport after a lockdown relaxation, was collected as part of the survey. The time cluster of the highest frequency (31.28% of the respondents)

Table 1. Socio-demographic characteristics.

Personal information	Attributes	Frequency	Percentage
Gender	Female	213	50.5%
	Male	209	49.5%
Age groups	15–18	7	1.7%
	19–30	128	30.3%
	31–45	152	36%
	46–65	123	29.1%
	>65	12	2.8%
Income	0–10.000€	84	19.9%
	10.000–20.000€	170	40.3%
	20.000–40.000€	110	26.1%
	>40.000€	58	13.7%
Car availability	Yes	335	79.4%
	No	87	20.6%
Marital status	Married-engaged	200	47.4%
	In a relationship	79	18.7%
	Divorced	15	3.6%
	Single	101	23.9%
	Widower	4	0.9%
	Other	23	5.5%
Profession	Public sector employee	76	18.0%
	Self-employed	90	21.3%
	Part time private sector employee	18	4.3%
	Full time private sector employee	114	27.0%
	Student	88	20.9%
	Retired	15	3.6%
	Unemployed	9	2.1%
	Other	12	2.8%
Household size	1	47	11.1%
	2	109	25.8%
	3	108	25.6%
	4	116	27.5%
	5	33	7.8%
	6 or more	9	2.1%

Table 2. Pre-pandemic travel characteristics.

Travel characteristics	Attributes	Frequency	Percentage
Major travel mode	Bike	19	4,5%
	Car	172	40.8%
	Bus	77	18.2%
	Metro	124	29.4%
	Bicycle/on foot	19	4.5%
	Taxi	6	1.4%
	Tram	5	1.2%
Trip purpose	Work	292	69,2%
	Leisure	22	5,2%
	Education	80	19,0%
	Shopping	15	3,6%
	Non home based	2	0,5%
	Other	11	2,6%
Frequency of public transport usage	Never	8	1,9%
	Rarely	61	14,5%
	Up to once a month	33	7,8%
	2–4 times a month	53	12,6%
	2–4 times a week	69	16,4%
	Once a week	35	8,3%
	Daily	163	38,6%
Fare	Single fare	187	44,3%
	Discounted single fare	46	10,9%
	Monthly pass	103	24,4%
	Discounted monthly pass	51	12,1%
	Other	35	8,3%

was the one including recovery time of 1 week to 1 month. Nevertheless, a similar number of travelers would use public transport in less than a week, demonstrating a positive attitude toward public transport in an era of a pandemic. On the other hand, about one out of seven participants stated that at least six months would have to pass until they would again use public transport. The distribution of answers is illustrated in Figure 2.

Clustering results

Four groups of travelers were identified, providing with a view of respondents' profiles:

(i) The first group involves middle-aged, frequent public transport users, with low or average family income, living alone or in couples and willing to use public transport in no more than a week, after pandemic restrictions are relaxed.
(ii) The second group consists of younger, frequent public transport users (mostly women and students), with low or average family income, living in households of an average size of three individuals and willing to refrain from public transport for about a month after a relaxation of pandemic restrictions.
(iii) The third group includes middle-aged, married respondents, who mostly use private cars instead of public transport, with average or high income, living in households of 3 or 4 members and willing to abstain from using public transport for about three months.
(iv) The fourth group consists of middle-aged respondents (mostly men), who mostly prefer private cars instead of public transport, having low or average income, members of nuclear families and who are not willing to use public transport no earlier than six months.

Table 3. Level of compliance with safety rules.

Statement	Minimum (1 = Totally disagree)	2	3	4	Maximum (5 = Totally agree)	Mean	Std. Deviation
I have limited my daily trips as much as possible	2.84%	5.92%	9.48%	18.72%	63.03%	4.33	1.05
I keep a safe distance from the others	1.66%	2.37%	10.66%	31.04%	54.27	4.34	0.88
I wash my hands regularly	0.95%	1.42%	6.64%	19.43%	71.56%	4.59	0.76
Level of compliance (a = 0.668)						4.42	0.70

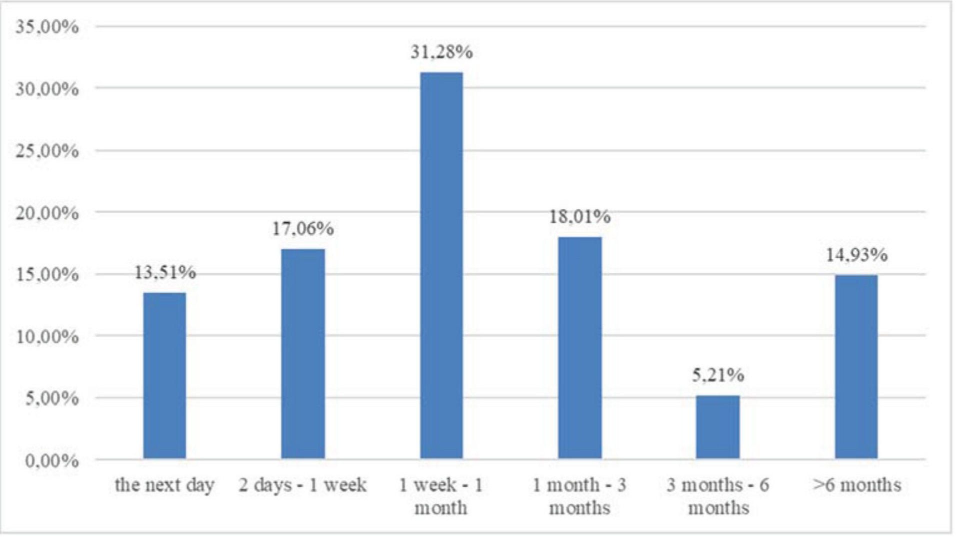

Figure 2. Stated recovery time distribution.

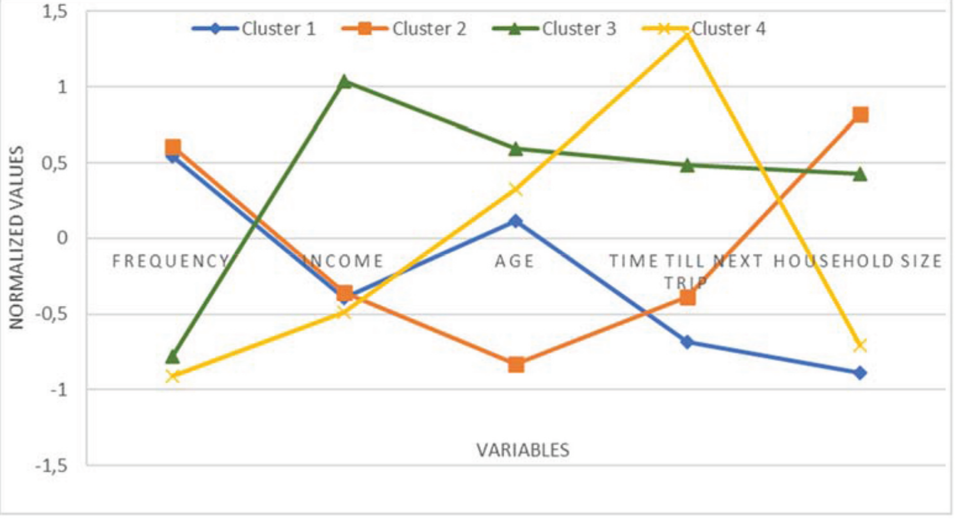

Figure 3. Travelers groups during the pandemic.

The abovementioned groups are presented in Figure 3, in which the lines correspond to different clusters, the clustering variables lie in the horizontal axis and their normalized values are presented in the vertical axis.

As expected, frequent, lower income travelers will return to public transport early after the outbreak of the pandemic; these are mostly captive travelers with hardly any alternatives for their daily trips. On the contrary, travelers who have access to private vehicles, and especially those who use public transport less often (but still use it), will refrain from using public transport for a longer period, even though some of them declared low to average incomes.

Recovery time modeling

Before the statistical analysis, the dataset was re-arranged in such a way for the discrete duration model to be estimated by logistic regression. Prior to this procedure, the dataset contained 422 rows, each one of which accounted for a single respondent. A suitable form of the dataset required the number of rows to be equal to the rank of the interval in which the event, that is re-using public transport, occurred. For example, each respondent accounted for four rows in the reformed dataset if the answer were that he or she would re-use public transport during the fourth time interval given as an option. Due to this procedure, the new dataset included 1327 rows. The dependent variable of the logistic regression (Y) took the value of 1 each time the event occurred at the respective time interval, and the value of 0 otherwise. To explain, considering that for every Ti given as an answer corresponds to a number of rows equal to i. The last of these rows include the variable Y with a value of 1, since the event occurs during the time interval which creates the last row for every participant. However, due to right-censoring, the absolute last time interval (the last option given in the survey) did not correspond to a row, whatsoever. That is, the maximum number of rows could be 5, although there were six alternatives in the questionnaire to choose from. On such occasions, the dependent variable was equal to 0 in every row. For instance, were a participant to state that they would have been refraining from public transport for more than six months (duration that corresponds to the sixth time interval), five rows would be created, in any of which Y would be equal to 0.

Five dummy variables were created, corresponding to the time intervals given in the survey, apart from the last one (due to censoring). In specific, T1 was for re-using public transport the next day, T2 for 2 days to 1 week, T3 for 1 week to 1 month, T4 for 1 month to 3 months and T5 for 3 months to 6 months. For every respondent Ti took the value of 1 in the row i and the value of 0 elsewhere. Therefore, hazard rate was different in every period. A model, without a constant term, including all the dummy variables was chosen (Milioti et al. 2019). The dummy variables correspond to the part of the model related to time. Nevertheless, another part of the model included all the other factors affecting recovery time, as mentioned in the methodology section. Hence, the model included additional variables concerning attributes deriving from the survey, which could explain the behavior of the travelers to some extent. Explanatory variables included in the final model were profession and age, frequency of public transport usage, public transport mode, and factors representing perception on public transport and COVID-19. The latter were the level of feeling safe when using public transport, their opinion about the efficiency of the government's measures for mitigating the virus spread as well as their willingness to take precautions, such as wearing a face mask, when using public transport. Results are presented in Table 4.

It must be noted that all the explanatory variables are statistically significant for a level of significance of 10%. However, seeking stricter limits, the independent variables are also significant for a level of 5%, apart from the one regarding the age group of 46 to 65, which demonstrates a slightly higher p-value of 7.1%.

According to the results, it can be inferred that the likelihood for the passengers to re-use public transport during T4 (1 to 3 months), given that they have not already done so, is the highest among the other time intervals, as the coefficient of the logistic regression is the highest. T3 (1 week to 1 month) and T5 (3 months to 6 months) demonstrate some dynamics, as for their probability to be chosen, as well. On the other hand, it seems rather improbable for the passengers to return in public transport in a period up to one week following the pandemic. The negative signs indicate that a higher absolute value leads to a lower probability, and thus the variables with the lowest absolute values are the ones with the highest likelihood. Nevertheless, all the variables having a negative sign yield a lower baseline hazard. Therefore, the shape of the hazard function is concave, demonstrating a peak (Tekle and Vermunt 2012). As far as socio-demographic and travel characteristics are concerned, findings suggest that travelers who are self-employed or have an age between 46 and 65 years old are less likely to use again public transport, by 41.8% and 24.9% compared to the other professions and age groups, respectively (according to the odds ratio). Travelers who mostly use a private car are 33.6% less likely to use public transport soon. On the other hand, frequent travelers and travelers who mostly use buses more likely to return to public transport.

When it comes to safety perceptions and psychological factors, those who feel safer when using public transport are the ones who are 69.0% more likely to return to public transport than those who feel less safe. In addition, respondents who exhibited a positive attitude toward the government's reaction to the outbreak, even if they considered it as 'too much', had an odds ratio of 1.144, meaning that it was 14.4% more likely for them to re-use public transport. Finally, participants who were not willing to take precautions in the future while using public transport, were the ones with a 58.7% higher probability of using public transport after the lockdown.

Table 4. Discrete duration model.

Covariate	Attribute	B	Sig.	Exp(B)
T1	1 if T1, 0 otherwise	−3.757	0.000	0.023
T2	1 if T2, 0 otherwise	−3.058	0.000	0.047
T3	1 if T3, 0 otherwise	−1.370	0.000	0.254
T4	1 if T4, 0 otherwise	−0.787	0.039	0.455
T5	1 if T5, 0 otherwise	−1.757	0.000	0.172
Self-employed	1 if the participant's profession was a freelancer, 0 otherwise	−0.541	0.002	0.582
Age between 46_65	1 if the participant's age was between 46 and 65, 0 otherwise	−0.287	0.071	0.751
Frequency of using public transport	0 = never to 6 = daily	0.318	0.000	1.374
Mostly use a bus	1 if the participant uses a bus, 0 otherwise	0.601	0.004	1.825
Mostly use of a private car	1 if the participant uses a private car, 0 otherwise	−0.410	0.021	0.664
Perceived safety level	1 = minimal safety when using public transport to 5 = maximal safety	0.525	0.000	1.690
Opinion about the government's measures	1 = hardly any to 5 = too many	0.134	0.049	1.144
Willingness to take precautions	1 = Yes, 0 = No	−0.462	0.031	0.630
Goodness of fit measures				
Cox & Snell R square		0.371		
Nagelkerke R square		0.494		
Classification Table				
Observed		Predicted		
Y	0	91.2%		
	1	38.3%		
Overall		76.9%		
Number of observations		1327		

Discussion and limitations

In this paper, two different methodologies were applied to analyze post-pandemic behavior of travelers with respect to public transport. First, using cluster analysis, the post-COVID-19 behavior of different respondent groups was investigated. Students, frequent passengers with low income, infrequent travelers with high income and infrequent ones with low-income form four closed groups and demonstrate a common post-COVID-19 attitude toward public transport. Students, and frequent travelers with low income are not willing to refrain from using public transport for a long period of time, after the pandemic. On the other hand, infrequent users of public transport belong to the same cluster with those who demonstrate high values of their recovery time. The same also applies to the age group between 46 and 65 years old, since this age group dominates the clusters with high recovery duration. Second, a discrete duration model was developed to investigate the factors affecting the decision of travelers to refrain from using public transport for some time. Findings suggest that socio-demographic characteristics as well as psychological factors affect public transport recovery time. Self-employed, those who mostly use private cars for their daily trips, travelers with ages between 46 and 65 years and people who prefer to use protection gear in a public transport ride, are less likely to return to normal public transport use after the outbreak. On the other hand, frequent public transport users, bus riders, travelers who feel more comfortable and safer using public transport, as well as people who consider that the government has applied a decisive policy against the pandemic are more likely to return sooner in public transport. Consequently, the findings deriving from clustering with respect to age and frequency of public transport use before the pandemic follow a similar trend with the ones of the discrete duration model, supporting the robustness of the results obtained.

A comparison with similar studies in the literature is considered crucial for a constructive screening of results. Based on data from US cities, Hotle, Murray-Tuite, and Singh (2020) found that men were less likely to change their behavior regarding public transport. However, according to this study, gender is not a significant estimator of recovery time. Beck and Hensher (2020) claimed that travelers in Australia demonstrate significant concerns toward public transport, which is in line with the findings of this study, which suggest that a large group of people (corresponding to 38.15% of the respondents) would be willing to refrain from public transport for at least one month. Pawar et al. (2020) also found that a high percentage (41.65%) of the commuters in India stopped traveling during lockdown. Nevertheless, they indicate that safety perception is not a significant factor affecting passengers' travel behavior. Contrarily, evidence from this study shows that safety perception is significant for passengers to return to public transport. The negative attitude toward public transport during a pandemic is also highlighted by De Haas, Faber, and Hamersma (2020) in their study conducted in the Netherlands. A trend for shifting from public transport to private modes emerges, which although is not directly considered in this research, it is consistent with the large group of people requiring over a month to 'recover', as well as their safety perceptions and the levels of fear. Specifically, travelers not feeling safe in public transport along with the ones considering precautions crucial, need more time to return into using public transport. Safety feelings were highlighted by Dong et al. (2021) as significant factors for passengers' perception on public transport use in China, while Tan and Ma (2020) proposed several attributes that influence travelers' behavior in China, with respect to rail transit. Interestingly, some of their findings are similar this study; self-employed passengers along with the risk-averse are less likely to use public transport in the future.

Limitations of the study derive from the web nature of the survey due to lockdown restrictions, that make impossible a face-to-face data collection effort. First, a web-based survey is extensively unsupervised, in terms of guidance and clarifications when needed. Second it is well known that an online questionnaire may suffer from accuracy and validity issues. To eliminate such problems, answers given were checked carefully those that appear to be problematic or have missing answers were eliminated. Moreover, in online surveys, the sample may be biased as the elderly and the technologically illiterate might have been excluded to large extent. Nevertheless, the age and gender distribution of the sample is quite representative of the Greek population. Finally, it should be noted that access to private car appears to be rather high (79.4%) in the sample. Therefore, responses may be somewhat biased in favor of refraining from public transport, due to the existence of an alternative travel mode. Future endeavors should focus on an enhanced sample, since the underlying sample, albeit representative, comes with all the constraints deriving from a survey during lockdown.

Conclusions

Public transport is the epitome of sustainable mobility in urban areas. Therefore, it is important that a positive stand, translated to a frequent yet safe use of public transport, be maintained by travelers as much as possible. This paper attempts to investigate traveler behavior toward public transport, following a pandemic outbreak, using survey data collected in Athens, Greece. Clustering and a discrete duration model are exploited for the purposes of the analysis. Results from both methods highlight the fact that the frequency of using public transport before the pandemic, along with the travelers' age, influence their behavior in terms of the time they will refrain from using public transport, after the pandemic. The findings of this study can be useful for policy making in the aftermath of the pandemic, for public transport ridership to recover in the shortest time period possible. According to discrete duration model results, self-employed travelers, and travelers with ages between 46 and 65 are the most likely to refrain from public transport, following a gradual exit from the pandemic. As such, marketing strategies targeting these groups may be necessary for recapturing some of these travelers. Also, there is evidence that specific psychological factors affect willingness to use public transport after the pandemic, such as safety perception and the desire to take precautions against an infection. Thus, efficient strategies should include campaigns and measures, aiming at promoting a feeling of safety among travelers, who use public transport. At the same time, social distancing should be maintained, by keeping an upper threshold on vehicle occupancy and improving frequencies where possible.

The duration of the COVID crisis, beyond the time-period initially foreseen, continues to affect travel behavior along two conflicting directions. First, the need for safety precautions during travel activities negatively affects public transport use and, second, the need for resuming social and economic activities affects positively public transport use. The full effect of these conditions on public transport use may be assessed, when the pandemic will be over. Useful insights could be also obtained by conducting a longitudinal study that could show how these findings evolve over time. It is expected that this study can provide a useful basis for comparison for future studies in this field.

Disclosure statement

No potential conflict of interest is reported.

ORCID

Eleni I. Vlachogianni http://orcid.org/0000-0002-2423-5475

Data availability statement

The data that support the findings of this study are available from the corresponding author, [KK], upon reasonable request.

References

Almagro, M., and A. Orane-Hutchinson. 2020. "JUE Insight: The Determinants of the Differential Exposure to COVID-19 in New York City and Their Evolution over Time." *Journal of Urban Economics* 103293. doi:10.1016/j.jue.2020.103293.

Andersen, A. L., E. T. Hansen, N. Johannesen, and A. Sheridan. 2020. "Pandemic, Shutdown and Consumer Spending: Lessons from Scandinavian Policy Responses to COVID-19." *arXiv Preprint arXiv* 2005: 04630.

Anjum, N. A. (2020). "Good in the Worst: COVID-19 Restrictions and Ease in Global Air Pollution."

Askitas, N., K. Tatsiramos, and B. Verheyden. 2020. "Lockdown Strategies, Mobility Patterns and Covid-19." *arXiv Preprint arXiv* 2006: 00531.

Authority, H. S. 2020. July 13). 2011 "Population-Housing Census." https://www.statistics.gr/en/2011-census-pop-hous

Baker, S. R., R. A. Farrokhnia, S. Meyer, M. Pagel, and C. Yannelis (2020). "How Does Household Spending Respond to an Epidemic? Consumption during the 2020 Covid-19 Pandemic (No. W26949)." National Bureau of Economic Research.

Beck, M. J., and D. A. Hensher. 2020. "Insights into the Impact of COVID-19 on Household Travel and Activities in Australia–The Early Days of Easing Restrictions." *Transport Policy* 99: 95–119. doi:10.1016/j.tranpol.2020.08.004.

Belot, M., S. Choi, J. C. Jamison, N. W. Papageorge, E. Tripodi, and E. Van Den Broek- Altenburg (2020). "Six-Country Survey on Covid–19."

Boldog, P., T. Tekeli, Z. Vizi, A. Dénes, F. A. Bartha, and G. Röst. 2020. "Risk Assessment of Novel Coronavirus COVID-19 Outbreaks outside China." *Journal of Clinical Medicine* 9 (2): 571. doi:10.3390/jcm9020571.

Cortina, J. M. 1993. "What Is Coefficient Alpha? An Examination of Theory and Applications." *Journal of Applied Psychology* 78 (1): 98. doi:10.1037/0021-9010.78.1.98.

Dalton, C. B., S. J. Corbett, and A. L. Katelaris. 2020. "Pre-emptive Low Cost Social Distancing and Enhanced Hygiene Implemented before Local COVID-19 Transmission Could Decrease the Number and Severity of Cases." *The Medical Journal of Australia* 212 (10): 1.

De Haas, M., R. Faber, and M. Hamersma. 2020. "How COVID-19 and the Dutch 'Intelligent Lockdown'change Activities, Work and Travel Behaviour: Evidence from Longitudinal Data in the Netherlands." *Transportation Research Interdisciplinary Perspectives* 6: 100150. doi:10.1016/j.trip.2020.100150.

De Vos, J. 2020. "The Effect of COVID-19 and Subsequent Social Distancing on Travel Behavior." *Transportation Research Interdisciplinary Perspectives* 100121. doi:10.1016/j.trip.2020.100121.

Dong, H., S. Ma, N. Jia, and J. Tian. 2021. "Understanding Public Transport Satisfaction in Post COVID-19 Pandemic." *Transport Policy* 101: 81–88. doi:10.1016/j.tranpol.2020.12.004.

Douglas, M., S. V. Katikireddi, M. Taulbut, M. McKee, and G. McCartney. 2020. "Mitigating the Wider Health Effects of Covid-19 Pandemic Response." *Bmj* 369. doi:10.1136/bmj.m1557.

Fahrmeir, L., and S. Wagenpfeil. 1996. "Smoothing Hazard Functions and Time-varying Effects in Discrete Duration and Competing Risks Models." *Journal of the American Statistical Association* 91 (436): 1584–1594. doi:10.1080/01621459.1996.10476726.

Gerhold, L. (2020). "COVID-19: Risk Perception and Coping Strategies."

Hale, T., A. Petherick, T. Phillips, and S. Webster (2020). "Variation in Government Responses to COVID-19." Blavatnik School of Government Working Paper. 31.

Haque, M. M., and S. Washington. 2015. "The Impact of Mobile Phone Distraction on the Braking Behaviour of Young Drivers: A Hazard-based Duration Model. Transportation Research Part C." *Emerging Technologies* 50: 13–27. doi:10.1016/j.trc.2014.07.011.

Hojati, A. T., L. Ferreira, S. Washington, and P. Charles. 2013. "Hazard Based Models for Freeway Traffic Incident Duration." *Accident Analysis & Prevention* 52: 171–181. doi:10.1016/j.aap.2012.12.037.

Honey-Roses, J., I. Anguelovski, J. Bohigas, V. Chireh, C. Daher, C. Konijnendijk, and E. Oscilowicz (2020). "The Impact of COVID-19 on Public Space: A Review of Themerging Questions."

Hotle, S., P. Murray-Tuite, and K. Singh. 2020. "Influenza Risk Perception and Travel- Related Health Protection Behavior in the US: Insights for the Aftermath of the COVID- 19 Outbreak." *Transportation Research Interdisciplinary Perspectives* 5: 100127. doi:10.1016/j.trip.2020.100127.

Jenkins, S. P. 1995. "Easy Estimation Methods for Discrete-time Duration Models." *Oxford Bulletin of Economics and Statistics* 57 (1): 129–138. doi:10.1111/j.1468-0084.1995.tb00031.x.

Jung, S. M., A. R. Akhmetzhanov, K. Hayashi, N. M. Linton, Y. Yang, B. Yuan, . . . H. Nishiura. 2020. "Real-time Estimation of the Risk of Death from Novel Coronavirus (COVID-19) Infection: Inference Using Exported Cases." *Journal of Clinical Medicine* 9 (2): 523. doi:10.3390/jcm9020523.

Kanungo, T., D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Wu. 2002. "An Efficient K-means Clustering Algorithm: Analysis and Implementation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (7): 881–892. doi:10.1109/TPAMI.2002.1017616.

Kodinariya, T. M., and P. R. Makwana. 2013. "Review on Determining Number of Cluster in K-Means Clustering." *International Journal* 1 (6): 90–95.

Lee, M., and M. You. 2020. "Psychological and Behavioral Responses in South Korea during the Early Stages of Coronavirus Disease 2019 (COVID-19)." *International Journal of Environmental Research and Public Health* 17 (9): 2977. doi:10.3390/ijerph17092977.

Madhulatha, T. S. 2012. "An Overview on Clustering Methods." *arXiv Preprint arXiv* 1205: 1117.

Milioti, C., K. Kepaptsoglou, A. Deloukas, G. Prodromitis, and C. Iliopoulou. 2019. "Modeling Traveler Recovery Time following Man-made Incidents: The Case of the Athens Metro." *Journal of Transportation Security* 12 (3–4): 103–117. doi:10.1007/s12198-019-00205-y.

Muller, S. A., M. Balmer, A. Neumann, and K. Nagel. 2020. "Mobility Traces and Spreading of COVID-19." *medR* xiv.

Musselwhite, C., E. Avineri, and Y. Susilo. 2020. "Editorial JTH 16–The Coronavirus Disease COVID-19 and Implications for Transport and Health." *Journal of Transport & Health* 16: 100853. doi:10.1016/j.jth.2020.100853.

Patel, V. R., and R. G. Mehta. 2011. "Impact of Outlier Removal and Normalization Approach in Modified K-means Clustering Algorithm." *International Journal of Computer Science Issues (IJCSI)* 8 (5): 331.

Pawar, D. S., A. K. Yadav, N. Akolekar, and N. R. Velaga. 2020. "Impact of Physical Distancing Due to Novel Coronavirus (Sars-cov-2) on Daily Travel for Work during Transition to Lockdown." *Transportation Research Interdisciplinary Perspectives* 7: 100203. doi:10.1016/j.trip.2020.100203.

Przybylowski, A., S. Stelmak, and M. Suchanek. 2021. "Mobility Behaviour in View of the Impact of the COVID-19 Pandemic—Public Transport Users in Gdansk Case Study." *Sustainability* 13 (1): 364. doi:10.3390/su13010364.

Quilty, B. J., S. Clifford, S. Flasche, and R. M. Eggo. 2020. "Effectiveness of Airport Screening at Detecting Travellers Infected with Novel Coronavirus (2019-ncov)." *Eurosurveillance* 25 (5): 2000080. doi:10.2807/1560-7917.ES.2020.25.5.2000080.

Rashidi, T., and A. Mohammadian. 2011. "A Competing Hazard Model of Household Vehicle Transaction Behavior with Discrete Time Intervals and Unobserved Heterogeneity." *Transportation Letters* 3 (3): 219–229. doi:10.3328/TL.2011.03.03.219-229.

Rodríguez-Morales, A. J., K. MacGregor, S. Kanagarajah, D. Patel, and P. Schlagenhauf. 2020. "Going global–Travel and the 2019 Novel Coronavirus." *Travel Medicine and Infectious Disease* 33: 101578. doi:10.1016/j.tmaid.2020.101578.

Roosa, K., Y. Lee, R. Luo, A. Kirpich, R. Rothenberg, J. M. Hyman, P. Yan, and G. Chowell. 2020. "Real-time Forecasts of the COVID-19 Epidemic in China from February 5th to February 24th, 2020." *Infectious Disease Modelling* 5: 256–263. doi:10.1016/j.idm.2020.02.002.

Sahin, A. R., A. Erdogan, P. M. Agaoglu, Y. Dineri, A. Y. Cakirci, M. E. Senel, and A. M. Tasdogan. 2020. "2019 Novel Coronavirus (COVID-19) Outbreak: A Review of the Current Literature." *EJMO* 4 (1): 1–7.

Steele, F. 2011. "Multilevel Discrete-time Event History Models with Applications to the Analysis of Recurrent Employment Transitions." *Australian & New Zealand Journal of Statistics* 53 (1): 1–20. doi:10.1111/j.1467-842X.2011.00604.x.

Steele, F., and E. Washbrook (2013) "Discrete-time Event History, Lecture Notes for Course LEM-06-11/12-P-R." Centre for Multilevel Modeling, University of Bristol

Tan, L., and C. Ma. 2020. "Choice Behavior of Commuters' Rail Transit Mode during the COVID-19 Pandemic Based on Logistic Model." *Journal of Traffic and Transportation Engineering (English Edition)*. doi:10.1016/j.jtte.2020.07.002.

Tan, P. N., M. Steinbach, and V. Kumar. 2013. *Data Mining Cluster Analysis: Basic Concepts and Algorithms. Introduction to Data Mining*, 487–533.

Tavakol, M., and R. Dennick. 2011. "Making Sense of Cronbach's Alpha." *International Journal of Medical Education* 2: 53. doi:10.5116/ijme.4dfb.8dfd.

Tekle, F. B., and J. K. Vermunt (2012). "Event History Analysis."

Tian, H., Y. Liu, Y. Li, M. U. Kraemer, B. Chen, C. H. Wu, and B. Wang. 2020. "Early Evaluation of Transmission Control Measures in Response to the 2019 Novel Coronavirus Outbreak in China." *medR* xiv.

Tu, H., S. Tu, S. Gao, A. Shao, and J. Sheng. 2020. "Current Epidemiological and Clinical Features of COVID-19; a Global Perspective from China." *Journal of Infection* 81 (1): 1–9. doi:10.1016/j.jinf.2020.04.011.

Virmani, D., S. Taneja, and G. Malhotra. 2015. "Normalization Based K Means Clustering Algorithm." *arXiv Preprint arXiv* 1503.00900.

Washington, S., M. G. Karlaftis, F. Mannering, and P. Anastasopoulos. 2020. *Statistical and Econometric Methods for Transportation Data Analysis*. CRC press.

Weijters, B., E. Cabooter, and N. Schillewaert. 2010. "The Effect of Rating Scale Format on Response Styles: The Number of Response Categories and Response Category Labels." *International Journal of Research in Marketing* 27 (3): 236–247. doi:10.1016/j.ijresmar.2010.02.004.

Wilkin, G. A., and X. Huang. 2007. "K-means Clustering Algorithms: Implementation and Comparison." *Second International Multi-Symposiums on Computer and Computational Sciences (IMSCCS 2007)*. August 133–136. IEEE.

Zheng, R., Y. Xu, W. Wang, G. Ning, and Y. Bi. 2020. "Spatial Transmission of COVID-19 via Public and Private Transportation in China." *Travel Medicine and Infectious Disease* 34: 101626. doi:10.1016/j.tmaid.2020.101626.