

Transit Usage in Social Shocks: A Case Study of Station-Level Metro Ridership in Anti-Extradition Movements in Hong Kong

Ho-Yin Chan

Department of Urban Planning and Design
The University of Hong Kong, Hong Kong, Hong Kong
Email: u3558916@connect.hku.hk

Hanxi Ma

Department of Urban Planning and Design
The University of Hong Kong, Hong Kong, Hong Kong
Email: mhx18@connect.hku.hk

Jiangping Zhou

Department of Urban Planning and Design
The University of Hong Kong, Hong Kong, Hong Kong
Email: zhoujp@hku.hk

Submitted [XX August 2021]

ABSTRACT

As a form of social movements aiming to effect social change, protests could bring about unintended impacts on all walks of life. In other words, expenses of protests can be incurred by those who might not be protesters. The protests triggered by an extradition bill in Hong Kong since 2019 are not an exception. This article focuses on the impacts on the ridership of metro system on protest days. We synthesize and hypothesize factors influencing the distribution of the ridership changes and conduct an empirical study in the context of Hong Kong to study its possible influencers and spatial dependence. We find that across metro stations, land use type (especially for commercial and open space), population age and income, as well as transit/road network characteristics and inter-modal connectivity significantly influence the ridership of metro stations during protest days. In addition, the mixed regressive spatial autoregressive model has higher explanatory power than the ordinary least square model, suggesting the need for a spatial lag and error specification. The results could also have significant policy and planning implications for operating metro services and managing metro stations before, amid and post social shocks.

Keywords: Public transit ridership; Social movement; Spatial regression; Smart card data; Hong Kong

1. Introduction

Social movements can be in a wide array of formats: protest, march, confrontation, vandalism, sabotage, and even revolution. They are generally theorized as collective efforts that the governed undertake to acquire “the right to the city” (e.g., 1, 2). Cities are prime sites for these events (3). From Occupy movement to Yellow Vest Movement and from Arab Revolutions to Color Revolutions, one after the other such events upholds this. A series of protests triggered by an extradition bill since June 2019 in Hong Kong provide additional and most recent evidence on this in the East Asia context. The direct aftermaths and indirect impacts of these protests are so wide-ranging and far-reaching that a wide array of topics is worth (further) investigation (e.g., 4, 5). To most local inhabitants in Hong Kong, one significant change due to the social movement they can observe is whenever protests occur, many protesters depend on public transit, metro in particular, to travel to and from protest sites. This certainly disturbs, strains, and even disrupts transit systems, in terms of the sudden increase in ridership and temporary suspensions and closures of certain entries, stations and lines due to safety concerns or unexpected vandalism by rioters. Researchers have produced in-depth knowledge on why, how, and when social movements occur, escalate, and spread (e.g., 3, 5, 6). Little work has been done, however, on where protests occur most frequently in cities, how many protesters travel to and from a common protest site, and how the systems could/should be managed efficiently given various disturbances due to the protests.

Hong Kong is a transit-oriented metropolis. In social movements, public transportation becomes a focus for different stakeholders, who expect public transportation to play different roles from normal days without any social movements. Since June 2019, Hong Kong has experienced numerous impactful protests triggered by an extradition bill (8). For transit operators and analysts, these protests and their impacts on the metro’s operations have offered them unprecedented “testbeds”, i.e., empirical cases and real-world data to rediscover the metro’s ridership in protests and explore how the system could/can be managed efficiently to meet safety and mobility demand of riders before, during, and after protests. Parallel to our previous works (9), of which resilience framework recognizes the importance of investigating demand shock during social shock, in this article, we further provide quantitative evidence with (spatial) regression models by quantifying the impacts of different influencers on the change in ridership and singling out prospective spatial error and spatial lag effects. We fit linear regression models (LRMs) to quantify the relationships between the change in ridership and its influencers and running spatial regression models (SRMs) to deal with spatial dependence, if any. Given the fairly understudied and somewhat nascent nature of the proposed study, the LRMs and SRMs will be more exploratory than hypo-deductive in nature. However, the results from these models could help formulate more explicit and relevant hypotheses for future research.

2. Literature review

Transportation systems face both nature- and human-induced incidents. Mega social events such as major sport events (e.g., Olympics) are examples of human-induced disruptions. But such disruptions involve significant proactive pre-event planning and strategies, which could help us manage the scale, duration, and scope of the disruptions (10, 11). Many travelers also are informed of and adapt to the change in travel conditions. For instance, an increase in acceptance of transit trips was witnessed during the 2000 Olympics in the car-oriented city of Sydney (11, 12). For the 2012 Olympics in London, trips were reduced and retimed to cope with the significant increase in transport demand (10). Nevertheless, the short-term adaptive behaviors are often unsustainable. In the case of Sydney 2000, the public were unable to transfer willingness for multi-modality into everyday travel in a longer term (12). For London 2012, reducing or re-timing journeys as adaptive behaviors were not sustained when the surrounding context returns to business as usual (10). However, these studies overlook the impacts of (transportation) infrastructure on travel behavioral change. The major sport events could be considered as opportunities to enhance transportation access, thanks to significant preparation and pledges; while (re)development of the Olympics site after the major sport events provides green areas for cycling and walking (13). Zhou et al. suggested that the change in transportation infrastructure, and thus housing affordability might induce a gradual change in demographic composition of local neighborhoods. This can overhaul

the overall travel patterns of a city in the long run (14). While these documents portray the disturbance from major sport events as a catalyst, bringing forward urban improvements via increased funding from the government, in general there is a lack of studies on the lasting impact on travel behaviors or policy to understand the need for change. More specifically, we still have little knowledge about crucial topics like who was most affected, for how long, and where they were during the interruptions.

Man-made disruptions could also be unpredictable events such as terrorist attacks. Although air travel have gotten a lot of attention in the past, terrorists start understanding how broad usage of such transportation infrastructure in the event of a terrorist strike may generate panic, disruption, and anxiety (15). Well-known examples include the Tokyo subway sarin attack in 1995, Madrid bombings in 2004 and the London bombings in 2005. Despite of its low probability of occurrence, people's behaviors may be strongly affected by the threat of terrorist attacks (16). Overall travel, in particular metro travel, can decrease after the above-mentioned events (17, 18). Meanwhile, people with different cultures, socio-demographic attributes and living in different neighborhoods may respond differently to terrorist events. Perceptions of risks and safety depends on the frequency of the attacks and whether people become desensitized (19), which may influence people long after the attack (20). Availability of other options can also affect substitution behaviors, either as instrumental constraints (e.g., availability of cars, bicycles, sidewalks, etc.) (17), or as personal constraints (e.g., budgetary and schedule constraints) (20). The occasional but severe disruptions highlight the importance of improvements in the security infrastructure for rail travel (15). Similarly, the disturbance from major threats can be regarded as a catalyst for behavioral changes and their underlying motivations, many of them are based on conventional samples with reported qualitative data. With emerging sources of large-scale mobility data, we are able to model human behaviors in response to disruptions, in particular, social shocks.

The abovementioned case studies provide examples of how a range of social events could resemble social protests in terms of disrupting normal transit operations (21). Transit infrastructure can enable and also suffer from social movements (9, 22), in other words, social movements could induce both predictable and unpredictable disruptions. On the one hand, political protests as one form of social activities can strain transit networks that are necessary for mobilizing large numbers of people; on the other hand, protesters disturb and even disrupt mass transit systems to gain political bargaining powers against the governments. Transit system could lose patronage from the intimidation of potential customers due to the fear of and concerns about the transit usage after service interruption or even targeted attacks on transportation infrastructure.

3. Research questions and method

In this study, we focus on the change in ridership by analyzing smart card data on protest days and answer the question:

- What are the influencing factors of transit ridership in social shock?
- What are the relationships between transit ridership and influencing factors?

We answer the question by fitting ordinary least squares (OLS) regression model to quantify the relationships between the ridership and their influencers and running spatial autoregressive (SARs) models to deal with spatial dependence, if any. Given the fairly understudied and somewhat nascent nature of the proposed study, the OLS and SARs will be more exploratory than hypo-deductive in nature. However, the results from these models could help formulate more explicit hypotheses for future research.

We hope to show that (a) regardless of their respective roots, various disruptions in transit operations are not exceptions; (b) we still know little about important issues such as who were affected the most, how long they were affected, and where they were located in the disruptions; (c) emerging smart card data can help us quantify the scale, scope and duration of both those disruptions and their heterogenous impacts on different rider groups; (d) when predicting the occurrence and impacts of those disruptions

based on smart card and other data, there could be some transferable procedures, methods and algorithms that we should develop or adapt to make better sense of the data—in this study, we have illustrated some.

4. The demand shock in social movement and its influencers

To the best of our knowledge, very little research has been conducted on the factors that influence the distribution of change in ridership during protests. We hypothesize that such variables can be grouped into four categories to guide our subsequent literature review and empirical studies: land use characteristics, transport infrastructure and network, socio-economic/ demographic characteristics, and political environment.

4.1 Land use characteristics

Transit ridership is strongly affected by population density. At the station level, there is evidence for a positive relationship between population density and transit ridership (23). The significance of urban density is that the more people who live and/or work near public transportation, the more likely the service is used.

Land-use type and mix are also a critical driver of transit ridership, though to a lesser degree than density (23–25). The land-use mix (diversity) results in a more balanced demand for public transportation both in terms of time (reducing the distance between the peak and off-peak periods) and in terms of space (in terms of the direction of flow). In general, especially for suburban areas, the mixed-use area has a higher transit use rate than the average area.

4.2 Transportation infrastructure and station characteristics

The transport infrastructure and service play a critical role in transit passengers to select transport mode, while transport diversity/inter-connectivity is crucial when there are adverse events due to demand shocks/ shutdown of stations (26, 27). The presence of feeder modes (bus stops) near the station can significantly affect ridership since some riders reach the station by public transportation (28–30). Moreover, the roadway infrastructure seems to be significantly important for activities in social movements. A dense network with a higher capacity and connectivity may favor different activities (29, 30). In other way, protesters may prefer an area with fewer number/length of road so that it is easier to take control of the road in relative safety (31). Thus, this study explores how the transport infrastructure and network affect the ridership of stations.

The characteristics of the stations are also relevant for explaining ridership. The number of riders is related to the type of station (i.e., intermediate, terminal, interchange, or intermodal) (28, 32). Terminal stations are the nearest stations for residents of a large area beyond the end of the line and people are willing to walk longer to reach this type of station. Interchange stations are more desirable to passengers than intermediate stations, and they appear to draw more riders, whereas intermodal stations, which accept riders from other modes of transportation, have higher boarding rates. The centrality of stations within the network is also important because people use public transportation more often in central areas than in peripheral areas (33, 34).

4.3 Socio-economic/ demographic characteristics

According to a survey about people's views on political development in Hong Kong (35), young people and people with higher education levels are more likely they support the pro-democracy movement. Moreover, as indicated in another survey by Lai et al. (36), participants in the anti-extradition bill protests reported higher education level and household income relative to nonparticipants. Thus, socio-economic/ demographic characteristics such as age, education level, and income influence engagement in political activities, e.g., march and assembly. Thus, the population attributes of metro-served areas can indirectly increase or decrease ridership (protest trips) in protest days.

4.4 Political environment

We quantify the local political environment from the results of the Election of District Councils (ECE) in Hong Kong. The electoral competition of seats was one of the political battlegrounds in the anti-extradition movement for the pro-democracy and the pro-government camps. Recognizing the political significance of District Councils, political parties of the two camps paid serious efforts to gain seats and votes in District Councils elections. The extent of competition in the 2019 District Council election was the fiercest since 1997 (37): On the one hand, the number of candidates taking part in the election was the highest in the election history of the District Council, reaching 1,090. There was no uncontested seat in all constituencies, the first time in history. On the other hand, the number of registered voters was a record high at 4.12 million, and the voter turnout rate was 71.2%, which was record-breaking in Hong Kong's election history. The abovementioned figures indicate the representativeness of the data to quantify the political environment of an area, i.e., whether the majority of voters are pro-democracy and the pro-government.

We used data available on the political party affiliations of candidates to determine their political stances (i.e., pro-democracy versus pro-government). We set the percentage of votes to candidates from two camps as the political spectrum of districts. Our hypothesis is that the more pro-democracy the area is, the higher the ridership in protest days.

4.5 Law and order

In Hong Kong, a public procession consisting of more than 30 persons is required to apply for the Letter of No Objection (LoNO) from the Police Commissioner (38). The Commissioner can object to the public procession, but only if s/he reasonably considers that the objection is necessary in the interests of national security or public safety, public order or the protection of the rights and freedoms of others. Organizers are required to be present at the procession to maintain good order and public safety and prohibit the unreasonable use of amplification devices for ensuring compliance with the directions given by police officers. We hypothesize that people's choice to use public transit on protest days may be influenced by their perceived extent to which police law enforcement can manage potential violent outbreaks. These outbreaks may impact the safety and security of transit users. Not surprisingly, protests can take many forms and can often turn violent.

5. Model specification

Ridership prediction models use the abovementioned variables as predictors at the station level. They estimate ridership as a function of station environments and transit service features, using multiple regression because of its ability to simultaneously evaluate the effects of a bunch of selected factors. Multiple regression is flexible, widely used and easily understood by a broad audience (23, 24, 28, 29, 32–34). In general, most studies used transit stations as the unit of analysis, and the resulting model can be used for predictive purposes. We also adopt the regressive spatial autoregressive (SAR) model to test possible spatial autocorrelation and minimize its impacts on coefficient estimation, which is in line with many exploratory studies (32, 33, 39).

5.1 Ordinary least squares (OLS) regression model

Generally, the OLS regression model is used to explore the global relationship between dependent and independent variables, which assumes independent variables have the same effect on the dependent variable across the study area. The OLS regression model is denoted by Eq. (1).

$$y = \alpha_0 + \sum_k \alpha_k x_k + \varepsilon \quad (1)$$

where y represents the dependent variable, α_k is the estimated coefficient of the independent variable x_k , and ε is the residual error.

5.2 Spatial autoregressive (SAR) model

Spatial autocorrelation is a frequent phenomenon in transport geographical data and can affect estimates of model coefficients and inference from statistical models. However, the estimated results obtained from OLS regression models do not account for any possible spatial dependence. Therefore, a spatial autoregressive (SAR) model is used to investigate if there is any spatial dependence.

In this article, we test the performance of three common SAR model types (40), namely the spatial error (SAR_{err}), spatial lag (SAR_{lag}) and mixed (SAR_{mix}) with OLS regression model when accounting for spatial autocorrelation in ridership distribution data. The SAR models are denoted by Eq. (2-4).

SAR_{err}:

$$\begin{aligned} y &= X\beta + u \\ u &= \lambda Wu + \varepsilon \end{aligned} \quad (2)$$

SAR_{lag}:

$$y = \rho Wy + X\beta + \varepsilon \quad (3)$$

SAR_{mix}:

$$\begin{aligned} y &= \rho Wy + X\beta + u \\ u &= \lambda Wu + \varepsilon \end{aligned} \quad (4)$$

where W is an $N \times N$ spatial-weighting matrix describing the spatial arrangement of the spatial units in the sample, the variable Wy denotes the endogenous interaction effects among the dependent variables and Wu the interaction effects among the disturbance terms of the different spatial units. ρ is called the spatial autoregressive coefficient, λ is the spatial autocorrelation coefficient, while β represents the unknown parameter.

To be conservative, in this article, we use GeoDa to examine spatial dependence and adopt the rules from Anselin's study (41) for the SAR model selection (**Figure 1**).

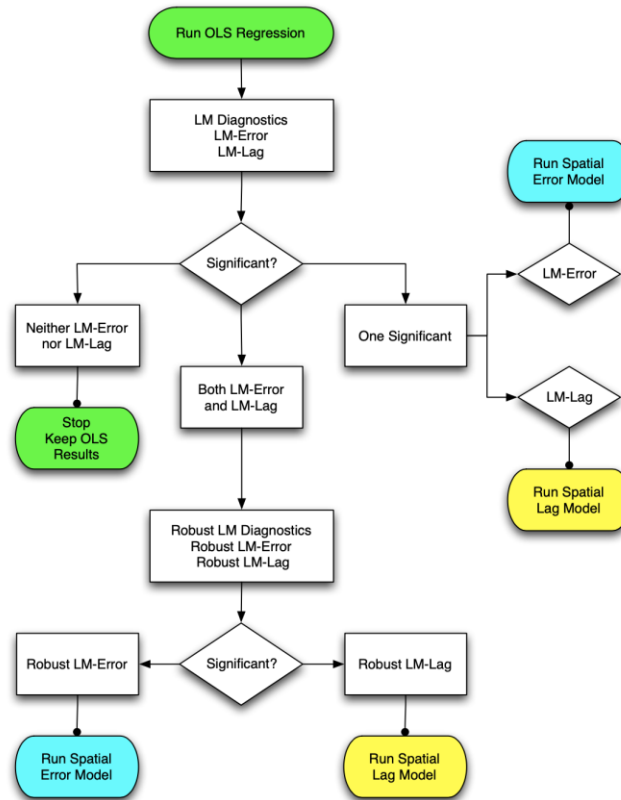


Figure 1 Spatial regression decision process (adopted from (41))

6. Empirical studies

6.1 The site and protest data

Public transportation, particularly the metro system of Hong Kong, is chosen as the study subject. Hong Kong adopts a transit-oriented development approach and provides one of the most efficient public transportation systems in the world. Approximately 90% of the 12.9 million daily motorized trips in Hong Kong are made by public transportation (42), one of the highest rates of any developed region worldwide. With the government's stated policy of "using railways as the backbone of Hong Kong's public transportation system," the Hong Kong metro network accounts for 43.4% of the average daily public transportation trips (42). As of 2019, 90 metro stations and ten metro lines (excluding the Airport Express Line) were operating. In Hong Kong, thousands of public gatherings, including protests occur each year (43). In between May and August 2019, some impactful protests have attracted so many participants that they have strained and disrupted MTR, posing various new challenges to the operator and riders (44).

6.2 Data source and processing

Our primary data for this study came from four sources: smartcard swipe data from the local transit agency, land use data from the Hong Kong Planning Department, transport data from the Transport Department, census data from the Census and Statistics Department, and District Councils election data from the Electoral Affairs Commission. Details are discussed in the subsequent sections.

6.2.1 Explanatory variables: land use characteristics

Land use data are derived from the outline zoning plans (OZPs) of the Hong Kong Planning Department (45). The major statutory land-use types found near railway stations in Hong Kong include residential,

commercial, industrial, institutional, mixed, open space, green belts, new development area, and others. The area of each land-use type within an 800-m buffer is used as variables.

6.2.2 Explanatory variables: transport infrastructure and network

The transport data are adopted from the open data platform of Hong Kong Transport Department. The road network dataset (46) is available in spatial data file format, allowing us to derive the network characteristics, including the road length. The transit dataset (47) is available in General Transit Feed Specification (GTFS), which includes stations/stops and service details. We evaluate the nearby transport infrastructure at station level with the total road length and bus/minibus stops that fall into M-SAs. We also derive station centrality within the metro network, namely the degree, betweenness, and closeness centrality by NetworkX package in Python (48).

6.2.3 Explanatory variables: socio-economic/ demographic characteristics

The 2016 census data are adopted from the Hong Kong Census and Statistics Department (49) at the Tertiary Planning Unit (TPU) level. The Hong Kong Planning Department uses these regional units for the fine-scale regional planning. The 291 TPUs in Hong Kong are aggregated by the Census and Statistics Department into 154 TPU groups to protect personal data privacy of census data. These data include demographic, educational, economic, household, and housing characteristics. All the data within the TPUs were transformed from the number of persons into ratio variables, which indicate the percentage of the population within the TPU with specific socio-demographic characteristics. Besides, each TPU group's total population is also involved in this study to evaluate the influences of these factors on metro ridership.

6.2.4 Explanatory variables: political environment

We obtain the results of the 2019 Election of District Councils in eighteen districts composed of 452 constituencies. The number of votes for each candidate and the geographical boundaries for each constituency is obtained at the website of Hong Kong Electoral Affairs Commission (50, 51). The percentage of votes to pro-democracy is calculated for each constituency and an average value is adopted for all constituencies within the metro served areas (M-SAs), which will be defined below.

6.2.5 Explanatory variables: law and order

We set the approval of the police for different protest sites/routes as a dummy variable. For our case study on June 9 and 16, 2019, police issued the LoNO to the organizer of the public meetings and procession, and the approved route with a Hong Kong metro map is shown in **Figure 1**. This is a common route for mass demonstrations in Hong Kong that starts from Victoria Park (between Causeway Bay and Tin Hau stations) and ends at the Legislative Council Complex near the Admiralty station. A total of seven metro stations, namely Fortress Hill, Tin Hau, Causeway Bay, Wan Chai, Admiralty, Central, and Hong Kong stations, are considered potentially affected by the ridership increase caused by the protests based on being situated within 800 meters. For those stations we record as “1” in the model.

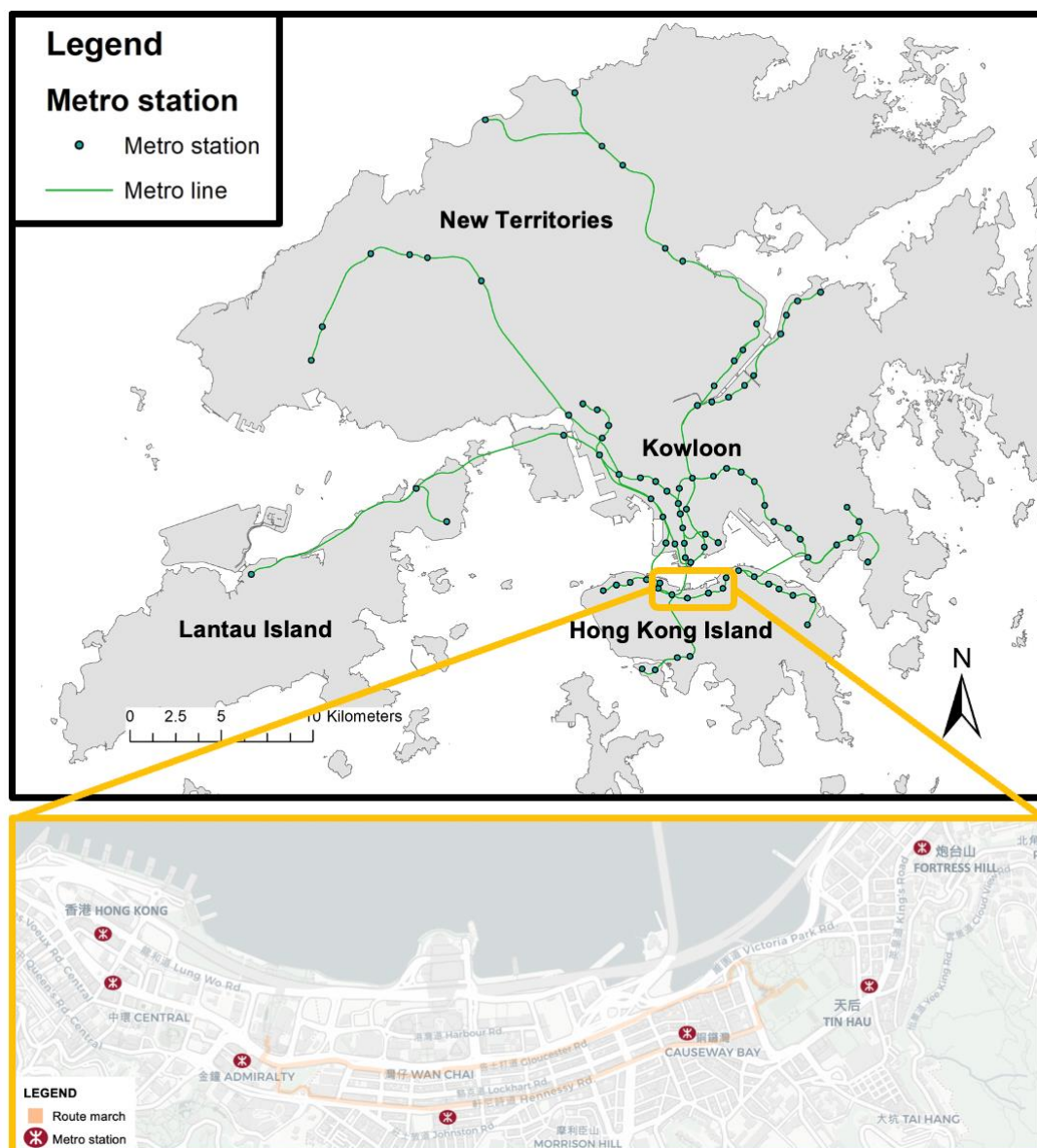


Figure 2 Route march on June 9 and 16, 2019

6.2.6 Explanatory variables: data processing

Lastly, a critical operational decision underlying station ridership analysis is how to define the M-SA. This decision influences M-SA variables, such as population, road density, bus stations, and land use mix. The M-SAs used in previous studies are normally determined using an empirical value, and researchers usually accept that an 800 m Euclidean distance or network distance that people are willing to walk to a station (23, 28, 32, 39, 52, 53). The detailed information of independent variables adopted in this study is provided in Table 2.

6.2.7 Dependent variable: change in ridership

Our primary data for ridership come from the smartcard swipe data from the Hong Kong Mass Transit Railway Corporation, which is called MTR locally. We define impactful protests as those with at least 50,000 protesters simultaneously gathering in one site well served by MTR. We choose 50,000 as the threshold because it is the typical per-hour capacity of a metro line (54). Such capacity can serve as a benchmark to gauge the severity of mass transit service pressure or disruption due to protests. To this end, we choose two extremely impactful protests, as shown in **Table 1** for our case study: we use the

June 9 data for model estimation while the June 16 data for validation. We use June 2, which is the same day of the week (Sunday) without any protests, as a baseline for estimating probable protest trips and affected recurrent and choice trips on different protest days-the latter can allow us to see how protests might have influenced different types of trips. We use the difference in daily ridership as the dependent variable in our models.

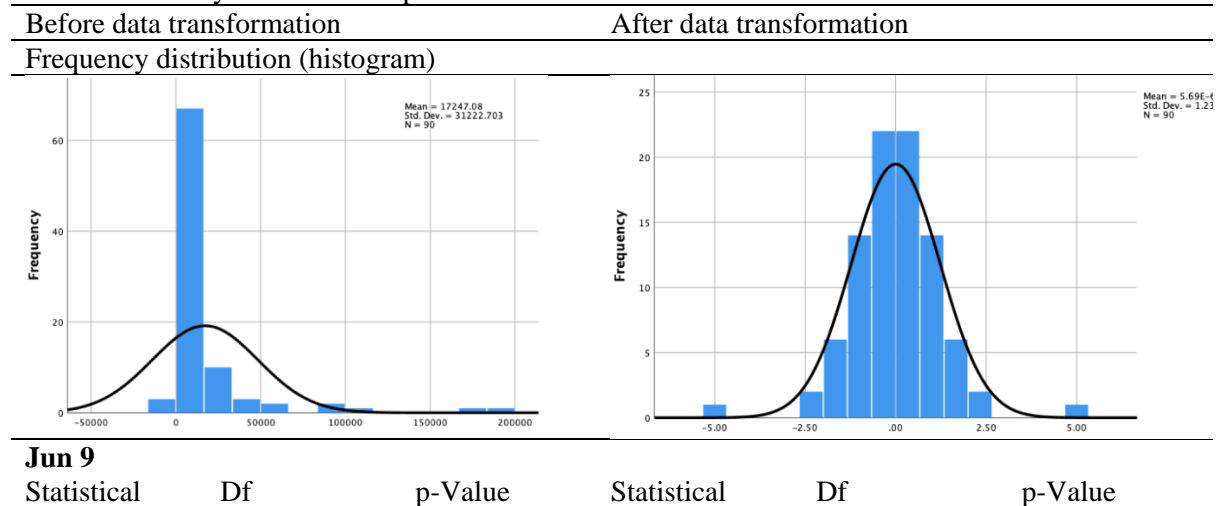
Table 1 Summary of two extremely impactful protests in Hong Kong in 2019

Date of 2019	Day of the week	Est. Start time	Est. End time	Type of activity	Epicenter	Est. size (ten thousand)	
						Organizer	Police
Jun 9	Sun	1400	2200	March	Victoria Park (Tin Hau) - Legislative Council Complex (Admiralty)	103	24
Jun 16	Sun	1400	2300	March	Victoria Park (Tin Hau) - Legislative Council Complex (Admiralty)	200	33.8

6.2.8 Dependent variable: data processing and assumption test

Data normality is verified based on Shapiro-Wilk and Kolmogorov-Smirnov (KS) tests (55), which are useful for small-to-medium-sized sample populations (i.e., $n < 300$) and therefore is suitable for our model ($n = 90$). Large Kolmogorov-Smirnov test values with a small p-value indicate non-normality, while skewness at or near zero implies a normal distribution (56). From **Table 2**, both skewness and KS coefficient indicate the data non-normality of the distribution. To ensure normality, we transform the data by quantile transformation, which can map a variable's probability distribution to another probability distribution, in this case, the normal distribution (57). The quantile transform function is available in the scikit-learn Python machine learning library, the "PowerTransformer" function in "sklearn.preprocessing" package (58). The quantile-transform provides an automatic way to transform a numeric input variable to have a different data distribution, which in turn, can be used as input to a predictive model. To visually inspect the shape and distribution of the dependent variables, **Table 2** shows the seasonal model dependent variables distribution curves overlaid on histograms before and after data transformation. As we see in the figure, after the quantile transformations, the dependent variable approaches normal distribution, with a KS coefficient of close to 0 with $p = 0.2$. Therefore, the transformed value of the changes in patronage is used as the dependent variable instead of directly using the original value. We summarize the statistics of dependent and independent variables adopted in **Table 3** and geo-visualize some key variables in **Figure 3**.

Table 2 Normality statistics of dependent variables before and after data transformation.



Kolmogorov–Smirnov					
0.325	90	<0.001	0.051	90	0.200
Shapiro–Wilk					
0.478	90	<0.001	0.927	90	<0.001
Jun 16					
Statistical	Df	p-Value	Statistical	Df	p-Value
Kolmogorov–Smirnov					
0.310	90	<0.001	0.057	90	0.200
Shapiro–Wilk					
0.511	90	<0.001	0.927	90	<0.001

Table 3 Summary of independent and dependent variables

Variables	Mean (stdev)	Notes
<u>Dependent variable</u>		
Change in station-level ridership on Jun 9 (transformed)	0 (1.23)	Derived from the Octopus data as the total boarding and alighting of a station. The reference day is June 2, 2019.
Change in station-level ridership on Jun 16 (transformed)	0 (1.23)	
<u>Independent variable</u>		
<i>Land use characteristics</i>		
Agriculture	6.08 (46.46)	The total area (m ²) within M-SAs from stations, derived from the latest (last updated: November 11, 2019) Outline Zoning Plans (OZP).
Commercial	94.21 (115.86)	
Industrial	33.89 (104.31)	
Institutional	225.38 (146.62)	
Mixed	50.24 (82.67)	
New Development Area	69.17 (122.44)	
Open Space	188.03 (101.54)	
Others	54.10 (118.18)	
Recreational	29.728 (109.85)	
Residential	400.81 (222.54)	
Transport	369.96 (162.99)	
Rangelands	229.20 (232.04)	
<i>Socio-economic/ demographic characteristics</i>		
Median of age	42.07 (4.74)	An aggregated value for M-SAs, derived from population census 2016 and by ArcGIS 10.6.
Median of income (thousand)	16.09 (6.52)	
Population (thousand)	40.76 (26.95)	
<i>Transport infrastructure and network</i>		
Interchange station	0.23 (0.43)	Whether the station is an interchange station: 1=yes 0=no
Terminus station	0.23 (0.43)	Whether the station is a terminus station: 1=yes, 0=no
Betweenness centrality (metro)	0.09 (0.08)	Station level, derived by NetworkX package in Python.
Closeness centrality (metro)	0.09 (0.08)	
Degree centrality (metro)	0.02 (0.00)	Road links within an M-SA, derived by ArcGIS 10.6. Numbers of bus/minibus stop within 800 m of a metro station, derived from General Transit Feed Specification (GTFS) data and by ArcGIS 10.6.
Total road length (km)	68.96 (53.07)	
Inter-modal connectivity: Bus	49.39 (28.37)	
Inter-modal connectivity: Minibus	66.49 (61.48)	
<i>Political Environment</i>		
Percentage of votes to Pro-democracy Camp	0.56 (0.05)	Derived from 2019 District Council election.
Epicenter of protests	0.08 (0.27)	Whether the station is located within an 800 m catchment from approved protest routes. 1: yes; 0: no

Legend

- MTR_Station_2019
- MTR_Line_2019
- TPU_2011_Population**
 - 0
 - 1 - 4556
 - 4557 - 14554
 - 14555 - 51030
 - 51031 - 287901

0 2.5 5 10 15 20 Kilometers

(a) Population

Legend

- MTR_Station_2019
- MTR_Line_2019
- ECE2019_Result**
 - 0.00 - 0.20
 - 0.20 - 0.40
 - 0.40 - 0.60
 - 0.60 - 0.80
 - 0.80 - 1.00

0 2.5 5 10 15 20 Kilometers

(b) Political environment (Percentage of votes to Pro-democracy Camp in ECE2019)

Figure 3 Geo-visualization of key independent variables

7. Results and discussions

7.1 Influence on ridership

Using smart card data, ridership is assessed as the entry and exit from a station, and the spatial distribution of positive and negative changes in ridership as dependent variables are shown in **Figure 4**.

For the positive change, there appears to have been significant numbers of trips to/from Hong Kong Island and the latter direction is of major demand. This can be reasonably explained by some protesters traveling to the protest sites by road transport but being unable to return in this manner because of road blockages. Specifically, we can witness a significant number of trips from the epicenter ends at some intermodal interchange stations (e.g., Tsim Sha Tsui station). The results reveal some overlooked stations that might come under pressure in some occasions, as bus–metro interchanges are not common in some areas on regular days. This implies our hypothesis that the characteristics of transport infrastructure and network, such as degree centrality and inter-modal connectivity, could be useful for predicting potential trips arising from protests.

For the negative or insignificant change, a significant decrease in the number of trips between other stations apart from the stations located in the epicenter. Specifically, an intense decrease is observed for trips to/from cross-border trips at the northern New Territories (e.g., Lok Ma Chau, Sheung Shui), and no significant changes are observed in some densely populated areas in New Territories and Kowloon (e.g., Kowloon Bay, Tsuen Wan West, Wu Kai Sha, Lohas Park). While the origin/destination are often dense residential area (i.e., residential new towns), we might expect a negative correlation between population and ridership in this sense, and a closer look at the socio-economic characteristics of the population (i.e., income and age) is necessary.

Overall, the visualization helps us to observe the difference in ridership on protest days compared to a normal day and the descriptive analysis can be used as a basis for interpretations of the model results.

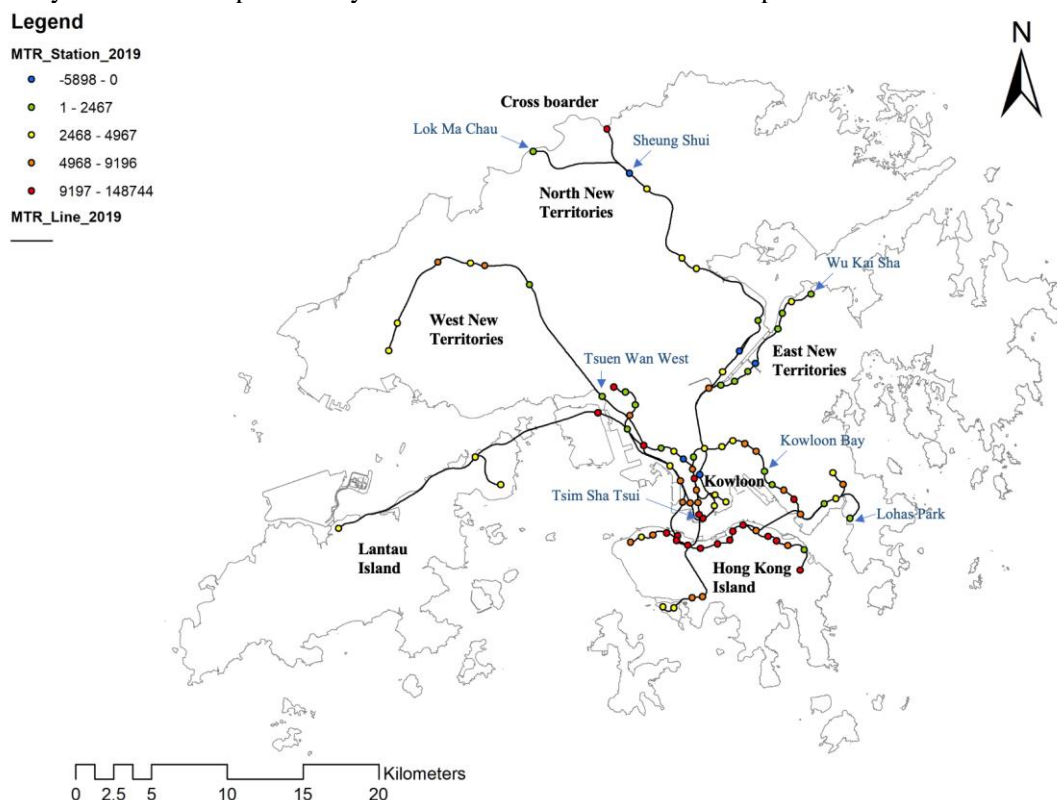


Figure 4 Change in station ridership on protest day (June 9) in reference to normal day (June 2)

7.2 Modeling results for influencing factors

First, Spearman's correlation analysis is first conducted to evaluate the relationship between variables, and the variables are removed if the correlation value was higher than 0.8 so that the variables with high correlation can be excluded. Then, the insignificant variables (with p-value smaller than 0.1), except for variables of special interests (see Section 2), are not included in the final model presented because there are a large number of variables. In addition, we check whether there are collinearity issues among the independent variables in all the regressions via the commonly used Variance Inflation Factor (VIF). To be conservative, we regard a VIF value larger than four as an indication of collinearity, in line with similar studies (39).

After testing different combinations of variables, the OLS results of the ridership change model are presented in **Table 4**. The results of the model indicate that among the four groups of factors of relevance to us, namely, population characteristics (measured by median age and income), land use type (measured by the area of commercial and open space land use), transport network (measured by the total length of roads, number of bus stops and betweenness/degree centrality of stations), are all statistically significant predictors for the ridership change. Together, these predictors explain 52% of the variance in the ridership change model for the selected protest day on June 9.

Across all independent variables, the intermodal connectivity (measured by the number of bus stops in M-SAs) and degree centrality of metro stations have the largest positive impacts on ridership change. It is reasonable since interchange and intermodal stations attract riders from other transportation modes, thus having higher ridership. Other variables that significantly influence the ridership change are the type of land use, including commercial and open space. It is understandable since protests in Hong Kong typically happen in business and shopping districts. At the same time, open spaces such as large-scale city parks are suitable places for mass gathering. Somewhat to our surprise, the population is not associated with ridership change, indicating it could be too simple factors to account for the change. This might also be because the protest occurred on Sunday when most workers no longer commuted. Interestingly, the political environment has a negative impact in the early stage of movement, indicating that the more pro-democracy M-SAs are the lower change in ridership can be observed. This might mean that election voters might mainly be conservative protesters. In the early stage of the movement, they might not be motivated for protests seeking radical change ($\beta = -0.20$). Finally, the results generally confirm that younger people and people with a higher income level are more likely to support/engage in the pro-democracy movement. Besides, it also suggests that protesters might prefer an area with fewer numbers/length of the road so that it is easier to take control of the area. In summary, the results confirm the reasonableness of our hypotheses about the ridership change and its influencers.

7.3 Spatial autocorrelation

To test possible spatial autocorrelation and minimize its impacts on coefficient estimation, we begin by examining Moran's I . Based on a weight of k-nearest neighbors ($k=2$ for neighbors being adjacent stations for most cases), the coefficient is 0.85. However, they are not statistically significant (i.e., $p > 0.1$), suggesting that further validation for the spatial autocorrelation is required. Thus, Lagrange Multiplier (LM) diagnostics are incorporated into the OLS regression (the baseline model) by GeoDa to examine the residuals and to determine the spatial model specification. The results again suggest the need for a spatial lag and error specification with $p < 0.1$ (see **Table 5**).

Table 4 presents the SAR and OLS regression results for the ridership change. The signs and magnitude of other coefficients are consistent with those in the OLS regression, while R^2 (the larger the better) and AIC (the smaller the better) imply a better fit of the SAR model than the OLS model. It also indicates the existence of spatial autocorrelation in the ridership change, which could be positively influenced by that of nearby stations, while the existence of spatial error could be negatively influenced by that of nearby stations.

Table 4 OLS and SAR regression results for station patronage change

	OLS		SAR		VIF
	Std coef (β)	Std.err	Std coef (β)	Std.err	
Population	-0.09	0.00	-0.11*	0.00	1.653
Income	0.22**	0.00	0.20**	0.00	1.431
Age	0.11	0.02	0.10	0.03	1.304
Metro_DC	0.19*	23.00	0.23	21.23	1.657
Metro_CC	-0.24**	8.67	-0.08	5.36	2.087
Bus_stop	0.13	0.01	0.10	0.00	3.092
Road_length	-0.03	0.00	-0.12	0.00	3.558
Commercial	0.23**	0.00	0.14*	0.00	1.637
OpenSpace	0.25***	0.00	0.13**	0.00	1.292
ProtestSite	0.30***	0.45	0.11	0.35	1.584
Polit_Env	-0.20**	2.23	-0.10*	1.68	1.108
Weight_lag			0.56***	0.21	
Weight_error			-0.57***	0.22	
Adjusted R ²	0.52		0.56		
AIC	249.28		191.32		
Log likelihood	-112.639				
Number of Observations	90				
Moran's I	0.85 (0.39)				

*** Indicators significance at the 99% level.

** Indicators significance at the 95% level.

* Indicators significance at the 90% level.

Table 5 Lagrange Multiplier Diagnostics

Lagrange Multiplier (lag)	0.94 (0.33)
Robust LM (lag)	3.87 (0.05)
Lagrange Multiplier (error)	0.02 (0.89)
Robust LM (error)	2.95 (0.09)

P values in bracket.

7.4 Model validation

We use the June 9 data for estimation and the June 16 data for validation, and compare the predicted vs observed outcomes for estimated models to illustrate the better predictive power and model superiority of SAR model. A plot of predicted vs observed outcomes for both models are shown in **Figure 5**. We can see both models are able to predict considerable number of samples, however, OLS generally has less predictive power for outliers. It is understandable since protests are dynamics that many underlying factors might not be considered/quantifiable in the models (e.g., temporary traffic/pedestrian flow management by police officers, impromptu actions by protesters, etc.). The SAR models try to account for some by considering the spatial-interaction effects of variables. For instance, protesters' actions likely to affect areas near the epicenters of protest sites, which could be to certain extent reflected in the ridership changes. The residuals' absolute maximum, median and minimum statistics further demonstrate a better performance of SAR models.

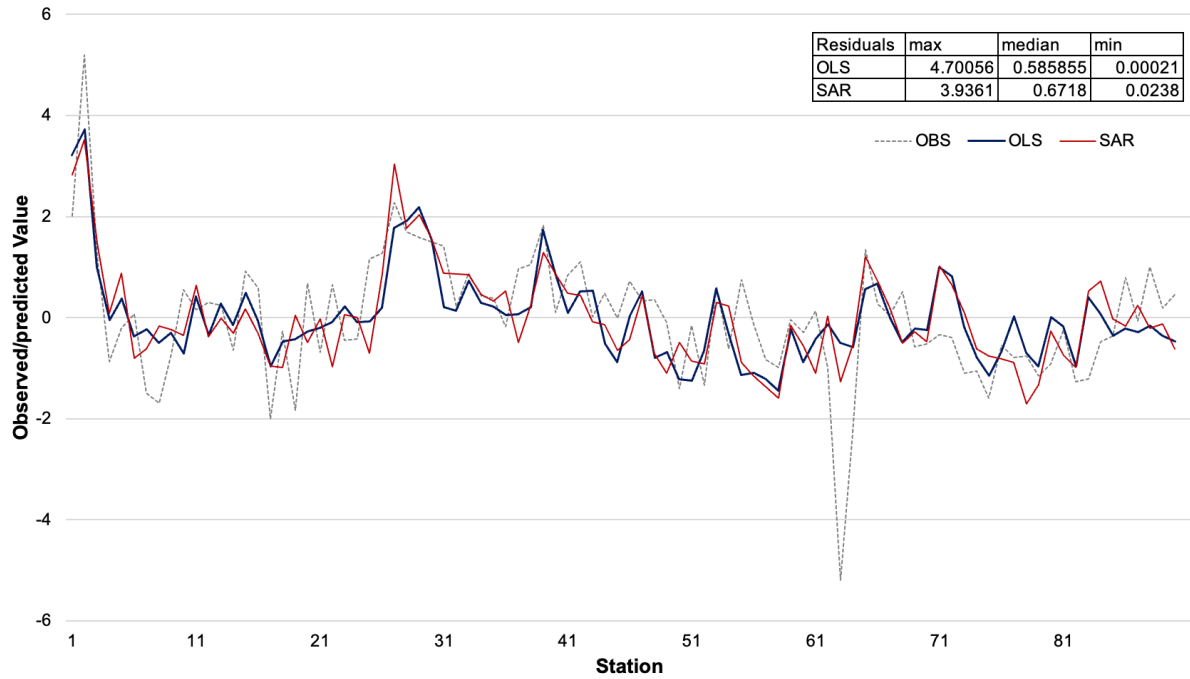


Figure 5 Validation using the plot of observed vs predicted value

7.5 Marginal effect analysis

In the marginal effect analysis for a linear regression model, an estimated coefficient represents the effect on the dependent variable by a unit change in the corresponding explanatory variable. While the dependent variables in both models are quantile transformed, the marginal effect of the independent variables is different from that in basic OLS models. Specifically, considering the slope (dy/dx) of each decile (i.e., each part represents 1/10 of the sample) is different (**Figure 6**), if the coefficient of an independent variable is β , an increase in the variable by one unit is correlated with a percentage increase of $\beta \times dy/dx$ in the dependent variables in each of the deciles. The estimated marginal effects of independent variables are shown in **Table 6** and **Table 7**. While the marginal effects are the product of the coefficient and the relevant dy/dx , the marginal effects generally increase for more strained stations (in terms of ridership changes). Examples of interpretation are the protest site being in a station will increase on median the ridership of that station by 21285; 1000 m² of commercial and open space results in an increase on median the ridership of that station by 0.04 and 0.05 respectively; and a 10⁻³ unit of increase in degree and closeness centrality of station contribute to changes in ridership by 685.09 and -284.96.

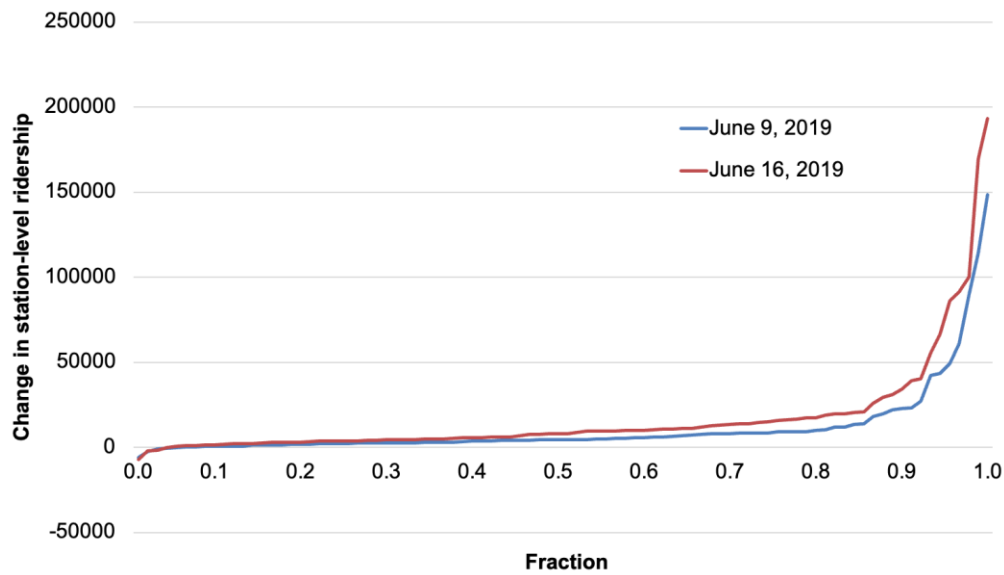


Figure 6 Change in station-level ridership by quantile

Table 6 Marginal effect of OLS and SAR regression results

Marginal effects (1st to 10th decile)			
OLS			
	Min	Median	Max
Population (thousand)	-27.12	-61.07	-4862.11
Income (thousand)	290.18	653.53	52033.27
Median of age	192.61	433.79	34537.38
Degree centrality (10^{-3})	304.20	685.09	54546.15
Closeness centrality (10^{-3})	-126.53	-284.96	-22688.32
Number of bus stop	40.93	92.19	7339.91
Total road length (km)	-0.99	-0.01	-0.01
Land use: Commercial (m^2)	0.02	0.04	3.01
Land use: Open space (m^2)	0.02	0.05	3.78
Protest site (binary)	9450.82	21284.53	1694644.29
Political environment (%)	-372.78	-839.55	-66844.09
SAR			
	Min	Median	Max
Population (thousand)	-56.86	-128.06	-10196.04
Income (thousand)	266.76	600.78	47833.26
Median of age	190.35	428.70	34132.18
Degree centrality (10^{-3})	370.02	833.34	66349.70
Closeness centrality (10^{-3})	-40.96	-92.25	-7345.21
Number of bus stop	29.87	67.27	5355.81
Total road length (km)	-0.02	-0.04	-3.40
Land use: Commercial (m^2)	0.01	0.02	1.89
Land use: Open space (m^2)	0.01	0.03	2.01
Protest site (binary)	-0.04	-0.08	-6.55
Political environment (%)	-190.03	-427.98	-34074.96

Notes: Values are in the original (i.e., not transformed) scale.

8. Conclusions

Our study has produced some promising outcomes and insights. Above all, by synthesizing discrete information in existing studies and by exploiting smart card data that we managed to access, we are able to validate our hypotheses about influencers of ridership change in social shocks brought by protests. Specifically, our model results show that the population's median age and income, land use type of commercial and open space, public transportation network, and road infrastructure are all statistically significant predictors of the ridership change. This informs transit operators about what to expect when protests happen and how to respond to them. For instance, the preference of protest sites (more open space for gathering, drawing more attention in commercial areas, less road length the site) can help predict the epicenter of political activities, which significantly affects the normal operations of transit services. Depending on the intensity, polarization, and political/social issue, different types of social movements are likely to have different sets of unintended consequences. The findings may not be generalizable across all forms of social movements since we focus on the influence on ridership on protest days, an aggregated dataset without analyzing individual behaviors.

Similar to many existing studies, this article examines the relationships between travel demand and its determinants at the local scale using linear models (39, 59), which are promising in exploratory studies. Given the fairly understudied and nascent issues in the interfaces of social movement and transportation, the linear models with additional cares on spatial correlations are exploratory than hypo-deductive in nature. However, the nonlinear nature of some of the relationships might involve, such as how a population across the political spectrum and with varying political opportunities (e.g., availability of space, points of interests from the political perspective) living near the stations along a metro line influence those relationships, could be overlooked. To fill these gaps, recent studies have tried to identify the nonlinear impacts of influencing factors on their respective predicted variables, adopting interpretable machine learning methods and using real-world big and/or open data. For instance, Xiao et al. (60) used graph convolutional neural networks to handle both linearity and nonlinearity issues for predicting metro station ridership. Enhanced models that can deal with the complexity of social movements might provide more accurate results.

As the socio-political context makes the travel environment and the decision-making processes of travelers and operators more complex, it also adds complexity to the assessment of transport resilience. To better understand complex social phenomena (e.g., socio-political movements with government counter-mobility measures and grassroots responses), a more comprehensive study with both qualitative (including interviews and questionnaires) and quantitative (including more complete travel datasets) methods is warranted. People's travel behavior is complex in itself, and in complex socio-political contexts, a thorough understanding of their perceptions and attitudes is needed. Qualitative (or mixed) methods would be a powerful tool for exploring these complexities because they allow the researcher to grasp the individuals' own reasons for their behavior and attitudes (61, 62).

Acknowledgements

The research underlying this paper is partially funded by a grant (Grant #: 17603220) by General Research Fund of Hong Kong. We would like to also thank the editor and three anonymous reviewers for their helpful comments that have led to the improvements of the manuscript. The views and any omissions in the paper are solely responsibilities of the authors.

Author Contributions

The authors confirm the contribution to this paper as follows: envisioning and securing grants in support of the study: J. Zhou; study conception and design: H. Chan, J. Zhou; data collection: H. Chan, H. Ma,

1 J. Zhou; analysis and interpretation of results: H. Chan, J. Zhou; draft manuscript preparation: H. Chan,
2 J. Zhou. All authors reviewed the results and approved the final version of the manuscript.
3

References

1. Brenner, N., P. Marcuse, and M. Mayer. Cities for People, Not for Profit. *City*, Vol. 13, No. 2–3, 2009, pp. 176–184. <https://doi.org/10.1080/13604810903020548>.
2. Harvey, D. The Right to the City. *International Journal of Urban and Regional Research*, Vol. 27, No. 4, 2003, pp. 939–941. <https://doi.org/10.1111/j.0309-1317.2003.00492.x>.
3. Uitermark, J., and W. Nicholls. How Local Networks Shape a Global Movement: Comparing Occupy in Amsterdam and Los Angeles. *Social Movement Studies*, Vol. 11, No. 3–4, 2012, pp. 295–301. <https://doi.org/10.1080/14742837.2012.704181>.
4. Yam, J. Hong Kong's Extradition Bill Protests: Dozens of Questions Remain to Be Answered. *South China Morning Post*, Jun 29, 2019.
5. Chan, H.-Y., and J. Zhou. Research Notes: Social Movement Revealing Opportunities for Grassroots Transport Initiatives: Lessons from Hong Kong. *Journal of the Eastern Asia Society for Transportation Studies*, Vol. 14, 2022, pp. 50–70. <https://doi.org/10.11175/easts.14.50>.
6. Salah Fahmi, W. Bloggers' Street Movement and the Right to the City. (Re)Claiming Cairo's Real and Virtual "Spaces of Freedom." *Environment and Urbanization*, Vol. 21, No. 1, 2009, pp. 89–107. <https://doi.org/10.1177/0956247809103006>.
7. Tufekci, Z. *Twitter and Tear Gas: The Power and Fragility of Networked Protest*. Yale University Press, New Haven and London, 2017.
8. Robles, P., D. Long, and D. Wong. 100 Days of Protests Rock Hong Kong. *SCMP*, Sep 17, 2019.
9. Chan, H.-Y., H. Ma, and J. Zhou. Public Transportation and Social Movements: Learning from the Hong Kong Anti-Extradition Bill Protests. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2676, No. 2, 2022, pp. 553–566. <https://doi.org/10.1177/03611981211044466>.
10. Parkes, S. D., A. Jopson, and G. Marsden. Understanding Travel Behaviour Change during Mega-Events: Lessons from the London 2012 Games. *Transportation Research Part A: Policy and Practice*, Vol. 92, 2016, pp. 104–119. <https://doi.org/10.1016/j.tra.2016.07.006>.
11. Hensher, D. A., and A. M. Brewer. Going for Gold at the Sydney Olympics: How Did Transport Perform? *Transport Reviews*, Vol. 22, No. 4, 2002, pp. 381–399. <https://doi.org/10.1080/01441640110121112>.
12. Mulley, C., and C. J. Moutou. Not Too Late to Learn from the Sydney Olympics Experience: Opportunities Offered by Multimodality in Current Transport Policy. *Cities*, Vol. 45, 2015, pp. 117–122. <https://doi.org/10.1016/j.cities.2014.10.004>.
13. McCarthy, M., R. J. Ravelli, and M. Sinclair-Williams. Health Impact Assessment of the 2012 London Olympic Transport Plans. *The European Journal of Public Health*, Vol. 20, No. 6, 2010, pp. 619–624. <https://doi.org/10.1093/eurpub/ckp154>.
14. Zhou, X., D. Hristova, A. Noulas, and C. Mascolo. Evaluating the Impact of the 2012 Olympic Games Policy on the Regeneration of East London Using Spatio-Temporal Big Data. *Social and Information Networks*, 2018.
15. Potoglou, D., N. Robinson, C. W. Kim, P. Burge, and R. Warnes. Quantifying Individuals' Trade-Offs between Privacy, Liberty and Security: The Case of Rail Travel in UK. *Transportation Research Part A: Policy and Practice*, Vol. 44, No. 3, 2010, pp. 169–181. <https://doi.org/10.1016/j.tra.2009.12.006>.
16. Elias, W., G. Albert, and Y. Shiftan. Travel Behavior in the Face of Surface Transportation Terror Threats. *Transport Policy*, Vol. 28, 2013, pp. 114–122. <https://doi.org/10.1016/j.tranpol.2012.08.005>.
17. Fasolo, B., Z. Ni, and L. D. Phillips. *A Study of the Impact of the July Bombings on Londoners' Travel Behaviour*. London, 2008.
18. López-Rousseau, A. Avoiding the Death Risk of Avoiding a Dread Risk: The Aftermath of March 11 in Spain. *Psychological Science*, Vol. 16, No. 6, 2005, pp. 426–428. <https://doi.org/10.1111/j.0956-7976.2005.01551.x>.
19. Litman, T. Terrorism, Transit and Public Safety: Evaluating the Risks. *Journal of Public Transportation*, Vol. 8, No. 4, 2005, pp. 33–45. <https://doi.org/10.5038/2375-0901.8.4.3>.

20. Hirsch, A. R., and A. Lee Levert. The Katrina Conspiracies: The Problem of Trust in Rebuilding an American City. *Journal of Urban History*, Vol. 35, No. 2, 2009, pp. 207–219. <https://doi.org/10.1177/0096144208327908>.
21. Loo, B. P. Y., and K. Y. K. Leung. Transport Resilience: The Occupy Central Movement in Hong Kong from Another Perspective. *Transportation Research Part A: Policy and Practice*, Vol. 106, No. April, 2017, pp. 100–115. <https://doi.org/10.1016/j.tra.2017.09.003>.
22. Badger, E. Why Highways Have Become the Center of Civil Rights Protest. *The Washington Post*, Jul 13, 2016.
23. Gutiérrez, J., O. D. Cardozo, and J. C. García-Palomares. Transit Ridership Forecasting at Station Level: An Approach Based on Distance-Decay Weighted Regression. *Journal of Transport Geography*, Vol. 19, No. 6, 2011, pp. 1081–1092. <https://doi.org/10.1016/j.jtrangeo.2011.05.004>.
24. Jun, M.-J., K. Choi, J.-E. Jeong, K.-H. Kwon, and H.-J. Kim. Land Use Characteristics of Subway Catchment Areas and Their Influence on Subway Ridership in Seoul. *Journal of Transport Geography*, Vol. 48, 2015, pp. 30–40. <https://doi.org/10.1016/j.jtrangeo.2015.08.002>.
25. Chan, H.-Y., Y. Xu, A. Chen, and X. Liu. Impacts of the Walking Environment on Mode and Departure Time Shifts in Response to Travel Time Change: Case Study in the Multi-Layered Hong Kong Metropolis. *Travel Behaviour and Society*, 2022. <https://doi.org/10.1016/j.tbs.2022.04.010>.
26. Chan, H.-Y., A. Chen, G. Li, X. Xu, and W. Lam. Evaluating the Value of New Metro Lines Using Route Diversity Measures: The Case of Hong Kong’s Mass Transit Railway System. *Journal of Transport Geography*, Vol. 91, 2021, p. 102945. <https://doi.org/10.1016/j.jtrangeo.2020.102945>.
27. Hsieh, C.-H., and C.-M. Feng. The Highway Resilience and Vulnerability in Taiwan. *Transport Policy*, Vol. 87, 2020, pp. 1–9. <https://doi.org/10.1016/j.tranpol.2018.08.010>.
28. Zhao, J., W. Deng, Y. Song, and Y. Zhu. Analysis of Metro Ridership at Station Level and Station-to-Station Level in Nanjing: An Approach Based on Direct Demand Models. *Transportation*, Vol. 41, No. 1, 2014, pp. 133–155. <https://doi.org/10.1007/s11116-013-9492-3>.
29. Zhao, J., W. Deng, Y. Song, and Y. Zhu. What Influences Metro Station Ridership in China? Insights from Nanjing. *Cities*, Vol. 35, 2013, pp. 114–124. <https://doi.org/10.1016/j.cities.2013.07.002>.
30. Gan, Z., M. Yang, T. Feng, and H. J. P. Timmermans. Examining the Relationship between Built Environment and Metro Ridership at Station-to-Station Level. *Transportation Research Part D: Transport and Environment*, Vol. 82, 2020, p. 102332. <https://doi.org/10.1016/j.trd.2020.102332>.
31. Endres, D., and S. Senda-Cook. Location Matters: The Rhetoric of Place in Protest. *Quarterly Journal of Speech*, Vol. 97, No. 3, 2011, pp. 257–282. <https://doi.org/10.1080/00335630.2011.585167>.
32. Zhou, J., Y. Yang, P. Gu, L. Yin, F. Zhang, F. Zhang, and D. Li. Can TODness Improve (Expected) Performances of TODs? An Exploration Facilitated by Non-Traditional Data. *Transportation Research Part D: Transport and Environment*, Vol. 74, 2019, pp. 28–47. <https://doi.org/10.1016/j.trd.2019.07.008>.
33. Shi, Z., N. Zhang, Y. Liu, and W. Xu. Exploring Spatiotemporal Variation in Hourly Metro Ridership at Station Level: The Influence of Built Environment and Topological Structure. *Sustainability*, Vol. 10, No. 12, 2018, p. 4564. <https://doi.org/10.3390/su10124564>.
34. He, Y., Y. Zhao, and K.-L. Tsui. Geographically Modeling and Understanding Factors Influencing Transit Ridership: An Empirical Study of Shenzhen Metro. *Applied Sciences*, Vol. 9, No. 20, 2019, p. 4217. <https://doi.org/10.3390/app9204217>.
35. Centre for Communication and Public Opinion Survey. *Public Opinion & Political Development in Hong Kong: Survey Results*. Hong Kong, 2014.
36. Lai, F. T. T., B. J. Hall, L. Liang, S. Galea, and W. K. Hou. Socioeconomic Determinants of Depression amid the Anti-Extradition Bill Protests in Hong Kong: The Mediating Role of Daily Routine Disruptions. *Journal of Epidemiology and Community Health*, 2020, p. jech-

- 2019-213693. <https://doi.org/10.1136/jech-2019-213693>.
37. Lam, J. T. M. The 2019 District Council Election in Hong Kong: A Localism Perspective. *Asian Affairs: An American Review*, 2021, pp. 1–20. <https://doi.org/10.1080/00927678.2021.1903779>.
38. Hong Kong Government. Cap. 245 Public Order Ordinance. *Hong Kong e-Legislation*. <https://www.elegislation.gov.hk/hk/cap245>.
39. Zhou, J., Y. Yang, H. Ma, and Y. Li. “Familiar Strangers” in the Big Data Era: An Exploratory Study of Beijing Metro Encounters. *Cities*, Vol. 97, 2020, p. 102495. <https://doi.org/10.1016/j.cities.2019.102495>.
40. Kissling, W. D., and G. Carl. Spatial Autocorrelation and the Selection of Simultaneous Autoregressive Models. *Global Ecology and Biogeography*, 2007, pp. 070618060123007-??? <https://doi.org/10.1111/j.1466-8238.2007.00334.x>.
41. Anselin, L. *Exploring Spatial Data with GeoDaTM: A Workbook*. Spatial Analysis Laboratory, Department of Geography, University of Illinois, Urbana-Champaign, 2005.
42. Transport Department. *Monthly Traffic and Transport Digest*. 2020.
43. Hong Kong Police Force. Public Order Event Statistics. *The Government of Hong Kong Special Administrative Region*. https://www.police.gov.hk/ppp_en/09_statistics/poes.html. Accessed Jan. 17, 2021.
44. Duhalde, M., and H. Huang. History of Hong Kong Protests: Riots, Rallies and Brollies. *South China Morning Post*, Jul 04, 2019.
45. HKSARG Planning Department. Hong Kong Outline Zoning Plans Land Use Zonings. *Esri China (Hong Kong)*. <https://www.arcgis.com/home/item.html?id=5375a88ec76143ea974d5fb64efbec0d>. Accessed Apr. 2, 2021.
46. Transport Department. Road Network (2nd Generation). *data.gov.hk*. https://data.gov.hk/en-data/dataset/hk-td-tis_15-road-network-v2. Accessed Mar. 1, 2021.
47. Transport Department. Headway Information of Public Transport Services. *data.gov.hk*. https://data.gov.hk/en-data/dataset/hk-td-tis_11-pt-headway-en. Accessed Mar. 1, 2021.
48. Hagberg, A., D. Schult, and P. Swart. NetworkX Reference: Release 2.6rc1.Dev0. *NetworkX developers*. https://networkx.org/documentation/latest/_downloads/networkx_reference.pdf. Accessed Apr. 2, 2021.
49. Census and Statistics Department. 2016 Population By-Census. *Financial Services and the Treasury Bureau*. <https://www.bycensus2016.gov.hk/en/>.
50. HKSARG Electoral Affairs Commission. 2019 District Council Ordinary Election Constituency Boundaries. *data.gov.hk*. https://www.eac.hk/en/distco/2019dc_elect_map.htm. Accessed Apr. 2, 2021.
51. HKSARG Electoral Affairs Commission. 2019 District Council Election - Election Results. *District Council Election*. <https://www.elections.gov.hk/dc2019/eng/results.html>. Accessed Apr. 2, 2021.
52. Kuby, M., A. Barranda, and C. Upchurch. Factors Influencing Light-Rail Station Boardings in the United States. *Transportation Research Part A: Policy and Practice*, Vol. 38, No. 3, 2004, pp. 223–247. <https://doi.org/10.1016/j.tra.2003.10.006>.
53. Cardozo, O. D., J. C. García-Palomares, and J. Gutiérrez. Application of Geographically Weighted Regression to the Direct Forecasting of Transit Ridership at Station-Level. *Applied Geography*, Vol. 34, 2012, pp. 548–558. <https://doi.org/10.1016/j.apgeog.2012.01.005>.
54. Bernick, M. S., and R. B. Cervero. *Transit Villages for the 21st Century*. McGraw-Hill, New York, 1997.
55. Kim, H.-Y. Statistical Notes for Clinical Researchers: Assessing Normal Distribution (2) Using Skewness and Kurtosis. *Restorative Dentistry & Endodontics*, Vol. 38, No. 1, 2013, p. 52. <https://doi.org/10.5395/rde.2013.38.1.52>.
56. Ghasemi, A., and S. Zahediasl. Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. *International Journal of Endocrinology and Metabolism*, Vol. 10, No. 2, 2012, pp. 486–489. <https://doi.org/10.5812/ijem.3505>.
57. Bolstad, B. M., R. Irizarry, M. Astrand, and T. P. Speed. A Comparison of Normalization Methods for High Density Oligonucleotide Array Data Based on Variance and Bias.

- 1 *Bioinformatics*, Vol. 19, No. 2, 2003, pp. 185–193.
2 <https://doi.org/10.1093/bioinformatics/19.2.185>.
- 3 58. Brownlee, J. *Data Preparation for Machine Learning: Data Cleaning, Feature Selection, and*
4 *Data Transforms in Python*. Machine Learning Mastery, 2020.
- 5 59. Liu, X., J. Wu, J. Huang, J. Zhang, B. Y. Chen, and A. Chen. Spatial-Interaction Network
6 Analysis of Built Environmental Influence on Daily Public Transport Demand. *Journal of*
7 *Transport Geography*, Vol. 92, 2021, p. 102991.
8 <https://doi.org/10.1016/j.jtrangeo.2021.102991>.
- 9 60. Xiao, L., S. Lo, J. Zhou, J. Liu, and L. Yang. Predicting Vibrancy of Metro Station Areas
10 Considering Spatial Relationships through Graph Convolutional Neural Networks: The Case
11 of Shenzhen, China. *Environment and Planning B: Urban Analytics and City Science*, Vol. 0,
12 No. 0, 2020, pp. 1–22. <https://doi.org/10.1177/2399808320977866>.
- 13 61. Schwanen, T., K. Lucas, N. Akyelken, D. Cisternas Solsona, J.-A. Carrasco, and T. Neutens.
14 Rethinking the Links between Social Exclusion and Transport Disadvantage through the Lens
15 of Social Capital. *Transportation Research Part A: Policy and Practice*, Vol. 74, 2015, pp.
16 123–135. <https://doi.org/10.1016/j.tra.2015.02.012>.
- 17 62. Chan, H.-Y., A. Chen, W. Ma, N.-N. Sze, and X. Liu. COVID-19, Community Response,
18 Public Policy, and Travel Patterns: A Tale of Hong Kong. *Transport Policy*, Vol. 106, No.
19 March, 2021, pp. 173–184. <https://doi.org/10.1016/j.tranpol.2021.04.002>.
- 20