Journal of Transport Geography 121 (2024) 104035

Contents lists available at ScienceDirect

Journal of Transport Geography

journal homepage: www.elsevier.com/locate/jtrangeo

A comparative analysis of cross-sectional study and natural experiment in rail transit-travel behavior research: A case study in Wuhan, China

Jingjing Wang [a], Yi Lu [b,*], Mi Diao [c,d], Ye Liu [e,f,g]

[a] School of Urban Design, Wuhan University, Wuhan, China
[b] Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong, China
[c] College of Architecture and Urban Planning, Tongji University, Shanghai, China
[d] Key Laboratory of Spatial Intelligent Planning Technology, Ministry of Natural Resources of China, Shanghai, China
[e] School of Geography and Planning, Sun Yat-Sen University, Guangzhou, China
[f] Guangdong Provincial Engineering Research Center for Public Security and Disaster, Guangzhou, China
[g] Guangdong Key Laboratory for Urbanization and Geo-simulation, Sun Yat-Sen University, Guangzhou, China

ARTICLE INFO

Keywords:
Urban rail transit
Travel behavior
Natural experiment
Cross-sectional study

ABSTRACT

There has been a global increase in investment in rail transit, driven by its potential to enhance transportation efficiency, reduce air pollution, and stimulate economic growth. Both cross-sectional studies and natural experiments have contributed to the growing body of evidence supporting these claims. While natural experiments are commonly preferred for evaluating the impact of rail transit, cross-sectional studies remain popular due to their ease of data collection. However, there is a scarcity of studies that compare these two approaches using the same dataset to assess the robustness of cross-sectional studies. Using a two-wave panel dataset from Wuhan, China, this study used both cross-sectional and natural experimental analyses to examine the relationship between urban rail transit and travel behavior. The study attempted to enhance the credibility of the cross-sectional analysis by controlling for confounding variables and by combining it with the propensity score matching (PSM) method, respectively. The results revealed that the cross-sectional analyses could produce similar results, when setting a more stringent significance level. The findings suggested that well-designed cross-sectional studies can be reliable and represent a cost-effective alternative to resource-intensive natural experiments.

1. Introduction

In recent years, many local and regional governments create new or expend existing rail transit systems, which are often regarded as a sustainable solution to alleviate traffic congestion and environmental problems stemming from automobile use (Nasri and Zhang, 2014). However, given the substantial investment required for rail transit infrastructure and its irreversible impact once built, it is critical to accurately assess the effects of rail transit, particularly its effectiveness in improving residents' travel behavior as expected.

Currently, the academic community favors the use of natural experiment methods to evaluate the impact of rail transit on the travel behavior of residents (Kärmeniemi et al., 2018). Natural experiments provide an opportunity to establish causal relationships between real-world interventions and outcomes. By leveraging the introduction of new rail transit, researchers can isolate the effects of the transit improvement from other confounding factors, enhancing the validity of causal claims (Leatherdale, 2019).

While natural experiments offer many advantages, their use in empirical research is often hampered by data collection challenges, e.g., extensive time and funding resources. Consequently, much research relies on cross-sectional approaches to study the effects of rail transit on travel behavior. Despite the limitations of cross-sectional studies, such as the lack of time-varying parameters and the issue of residential self-selection, they remain popular in rail transit and travel behavior research due to their ease of data acquisition and analysis. Establishing correlations that align with our hypothesized causal relationship serves as a crucial initial step in investigating potential causal links. The implementation of a typical cross-sectional design requires significantly fewer resources compared to conducting natural experiments. Therefore, employing cross-sectional approaches to lay the groundwork for further research on causal relationships is a sensible course of action.

* Corresponding author.
E-mail addresses: wangjingjing79@whu.edu.cn (J. Wang), yilu24@cityu.edu.hk (Y. Lu), diaomi@tongji.edu.cn (M. Diao), liuye25@mail.sysu.edu.cn (Y. Liu).

https://doi.org/10.1016/j.jtrangeo.2024.104035
Received 9 December 2023; Received in revised form 13 July 2024; Accepted 15 October 2024
Available online 25 October 2024

Furthermore, carefully designed cross-sectional studies, particularly those incorporating control variables and rigorous research methods, have the potential to provide reliable and valuable contributions to our understanding of the subject matter (Spector, 2019).

However, few studies have compared the results of cross-sectional studies and natural experiments with a single dataset to assess the reliability of cross-sectional studies. To the best of our knowledge, there is only one study that has used both cross-sectional and longitudinal analyses to explore the effect of metro transit on driving and found conflicting results (Huang et al., 2019). Additional empirical evidence is needed to investigate the disparities between the findings of cross-sectional analyses and natural experiments. Furthermore, it is crucial to explore the potential utilization of statistical methods in aligning and refining the outcomes of cross-sectional analyses. This would enable cross-sectional analyses to serve as a cost-effective alternative to more resource-intensive natural experiments, particularly in situations where research resources are limited.

In this study, we used both cross-sectional and natural experiment analyses to estimate the relationship between urban rail transit and travel behavior in Wuhan, China. To strengthen the reliability of the cross-sectional analysis, this study employed two strategies: controlling for additional variables and combining it with the propensity score matching (PSM) technique. Then, we conducted a comparison between the results obtained from cross-sectional analyses and the natural experiment study. The aim was to assess the extent to which cross-sectional analyses can yield consistent findings with the natural experiment, and to explore methodological optimizations that can enhance the reliability of cross-sectional analyses as a valuable complement to natural experiment studies.

2. Literature review

A substantial body of literature has demonstrated that the construction of urban rail transit has a great influence on individuals' travel behavior, which can adjust their trip frequency, distance, duration, travel mode, and trip distribution. The availability of rail transit options provides individuals with a convenient and efficient alternative to other travel modes, leading to changes in their travel patterns (Cao and Schoner, 2014). Researchers have found that proximity to rail transit stations or access to rail lines can encourage individuals to reduce their reliance on private vehicles, resulting in a shift to more sustainable travel modes, including walking, cycling, or using public transit (Xie, 2016). Furthermore, the presence of rail infrastructure has been observed to improve the overall accessibility and connectivity of urban areas, enabling residents to travel more frequently, cover longer distances, and access a wider range of destinations (Deng and Zhao, 2022; Wang et al., 2023b). However, contrary arguments have been put forth by numerous scholars, suggesting that the newly built rail infrastructure may not necessarily effectively promote individuals' sustainable travel behavior (Chatman, 2013). This is due to the influence of various confounding factors, such as the characteristics of the station's built environment, personal travel preferences, and attitudes. Additionally, the relationship between rail transit and travel behavior may exhibit nonlinear patterns and vary spatially, further complicating the association (Cheng et al., 2022; Tao et al., 2023). For instance, a study conducted in Hong Kong revealed a significant threshold effect of MTR and bus accessibility on daily trip duration specifically for low-income groups (Tao et al., 2023). Similarly, another study conducted in Beijing identified significant spatial heterogeneity in the influence of metro station accessibility on metro ridership (Du et al., 2022). These findings indicate that the association between rail transit and individuals' travel behavior is complex, calling for further high-quality and in-depth research (Wang et al., 2023a).

Existing studies in this field have mainly used two methods: cross-sectional study and natural experiment study. In a cross-sectional study, data are collected from a population at a single time point and used to explore associations or patterns (Wang and Cheng, 2020). Fig. 1 shows the framework of a typical cross-sectional study design in this area. In a cross-sectional study, researchers divide the sample into two or more distinct groups: the treatment group and the control group (Wang et al., 2023b). The treatment group refers to the participants who received an active intervention, which involves the introduction of newly constructed rail infrastructure. The control group represents the participants who did not receive such an intervention. Scholars then compare the travel behavioral outcomes between these groups, attributing any differences in outcomes to the effects of the transit intervention. In contrast to longitudinal studies that track respondents over a period of time, cross-sectional analyses capture a momentary snapshot of information at a given moment (Melissa and Morrison, 2009). This makes cross-sectional research easy, quick, and inexpensive to conduct. It serves as a valuable tool for hypothesis generation and examination of multiple outcomes and exposures that can inform other studies. However, due to the inherent methodological limitations (e.g., lack of temporal sequence and presence of selection bias), cross-sectional studies remain hard to establish the causal relationship between urban rail transit and travel behavior (Wang and Cheng, 2020).

A natural experiment is a study design in which the experimental conditions are determined by nature or other factors rather than under the control of the researchers (Leatherdale, 2019). Natural experiments have been widely used as a study design when controlled experiments are difficult to conduct, such as in the fields of economics, political science, and public health, and have provided important evidence to support the effectiveness of some specific interventions (Craig et al., 2017). Fig. 2 shows the framework of a typical natural experiment study design. In a natural experiment, as in a cross-sectional study, researchers also divide the sample into a treatment group that received a rail transit intervention and a control group that did not. By collecting data before and after the intervention, researchers can compare changes in individuals' travel behavior between the treatment and control groups. In an ideal natural experiment design, other than the intervention, the other variables should remain homogeneous so that any disparities in outcomes can be seen as the effect of rail transit.

While natural experiments are often regarded as the gold standard for evaluating interventions of a rail transit due to their ability to establish causal relationships, they typically require more financial resources and time compared to cross-sectional studies (Zhong et al., 2021). This highlights the need to compare the two methods to determine whether cross-sectional studies can serve as a viable and cost-effective alternative to more expensive natural experiments, particularly when research resources are limited. However, there is limited empirical research that directly compares the two methods in the field of travel behavior. There was only one study that adopted both cross-sectional and natural experiment designs to investigate the impact of metro transit on driving behavior. The results of this study indicated that the quasi-experiment revealed a negative association between metro transit and driving, whereas the cross-sectional analysis found no significant effect (Huang et al., 2019). The authors further concluded that regardless of the sophistication of the modeling approaches applied to cross-sectional data, the quasi-experiment appears to yield more robust results than cross-sectional studies. However, additional empirical evidence is still needed to better understand how the findings of cross-sectional analyses differ from those of natural experiments. This comparative analysis of the two methods can help researchers and policymakers make more informed choices when designing studies and allocating research resources.

3. Methods

3.1. Study area

The study area is situated in Wuhan, which is a city located in Hubei Province, China. In 2021, Wuhan had an urban area of 8569 km^2 and an

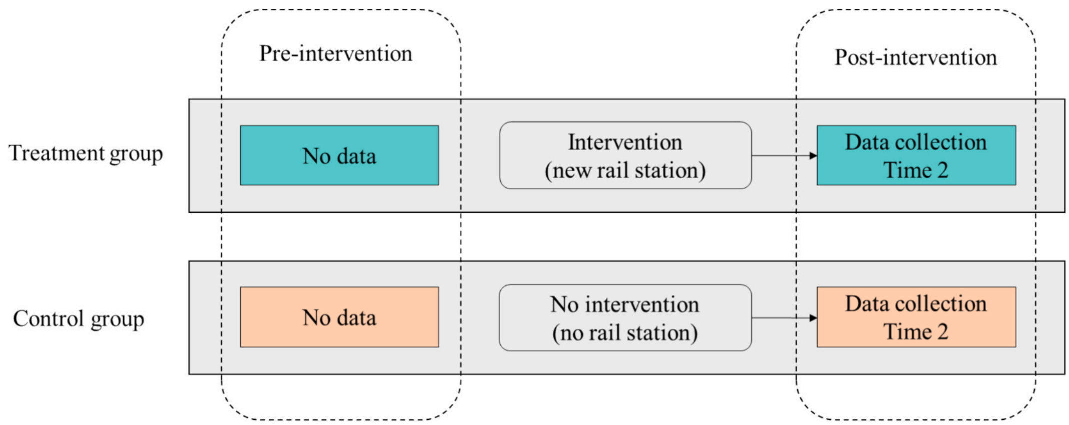

Fig. 1. The framework of a cross-sectional study design.

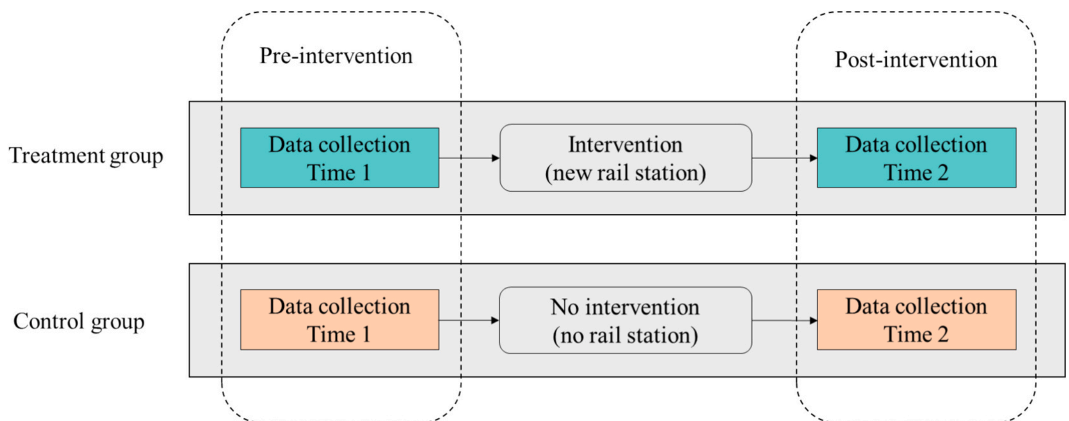

Fig. 2. The framework of a natural experiment study design.

urban population of 13.6 million (Wuhan Statistics Bureau, 2022). As the ninth most populous city in China, Wuhan is currently experiencing rapid urbanization and motorization, resulting in significant challenges to its transportation infrastructure due to increased demand for urban mobility and private vehicle use. To alleviate these problems, the government is focusing on expanding the city's rail transit system. By December 2022, Wuhan's rail transit system encompasses a total of 11 operational lines, with a combined operational mileage of 460 km (Wuhan Statistics Bureau, 2022).

The transit intervention in this study was Wuhan Metro Line 8 (Phase 2) which commenced operations in January 2021. This new metro line consists of 12 stations, running from the city center to the southern urban suburb (Fig. 3). As a major passenger flow channel in the central area of the city, the line provides a more convenient commuting option for residents along the line, as well as alleviating traffic congestion in the north-south direction of the central city. In order to ensure the exclusion of interference from surrounding lines while maintaining the representativeness of rail transit stations in the area, this study focuses on four stations of the line.

In this study, the distance between residents' residences and urban rail transit stations is used to define the treatment and control groups. Many scholars have conducted research and exploration on the threshold of the catchment area of urban rail transit stations, among which the threshold of 800 m has been widely adopted as an acceptable walking distance of most transit riders in China (Zacharias and Zhao, 2018). In addition, some Chinese studies have found that due to the widespread use of shared bicycles, the coverage area of urban rail transit stations should not only consider the walking impact range, but also consider the bicycle connection distance, which is usually 1600 m (Sun and Zacharias, 2017). For example, in areas where the distribution of urban rail transit lines is not very dense, the acceptable distance for residents to access stations will increase. Therefore, to better capture the influence range of urban rail transit stations, this study set up two treatment groups based on the thresholds of 800 m and 1600 m. The first treatment group (Group 1) selected neighborhoods within 800 m of the four stations, the second treatment group (Group 2) selected neighborhoods between 800 m and 1600 m from the stations, and the control group (Group 3) selected neighborhoods between 1600 m and 2400 m from the stations. The spatial distribution of the sampled neighborhoods is shown in Fig. 3.

3.2. Data and variables

A two-wave panel survey was conducted before and after the opening of Metro Line 8 (Phase 2) in November–December 2020 and 2021 respectively. We assembled a survey team of 25 local undergraduate and postgraduate students who were responsible for recruiting respondents and administering the survey. The survey team used a stratified, multistage sampling method to select respondents. First, To ensure the representativeness of the sample, we randomly selected approximately 50 % of the eligible neighborhoods from each buffer as sampling points. Ultimately, we identified 5, 3, and 2 neighborhoods in Group 1, 2, and 3, respectively, for each station. It is important to note that the number of candidate neighborhoods decreased as the distance from the metro station increased, as some neighborhoods were already influenced by other established metro stations. Next, 20–25 households within the selected neighborhoods were randomly chosen based on the building and house numbers. From each selected household, a qualified member

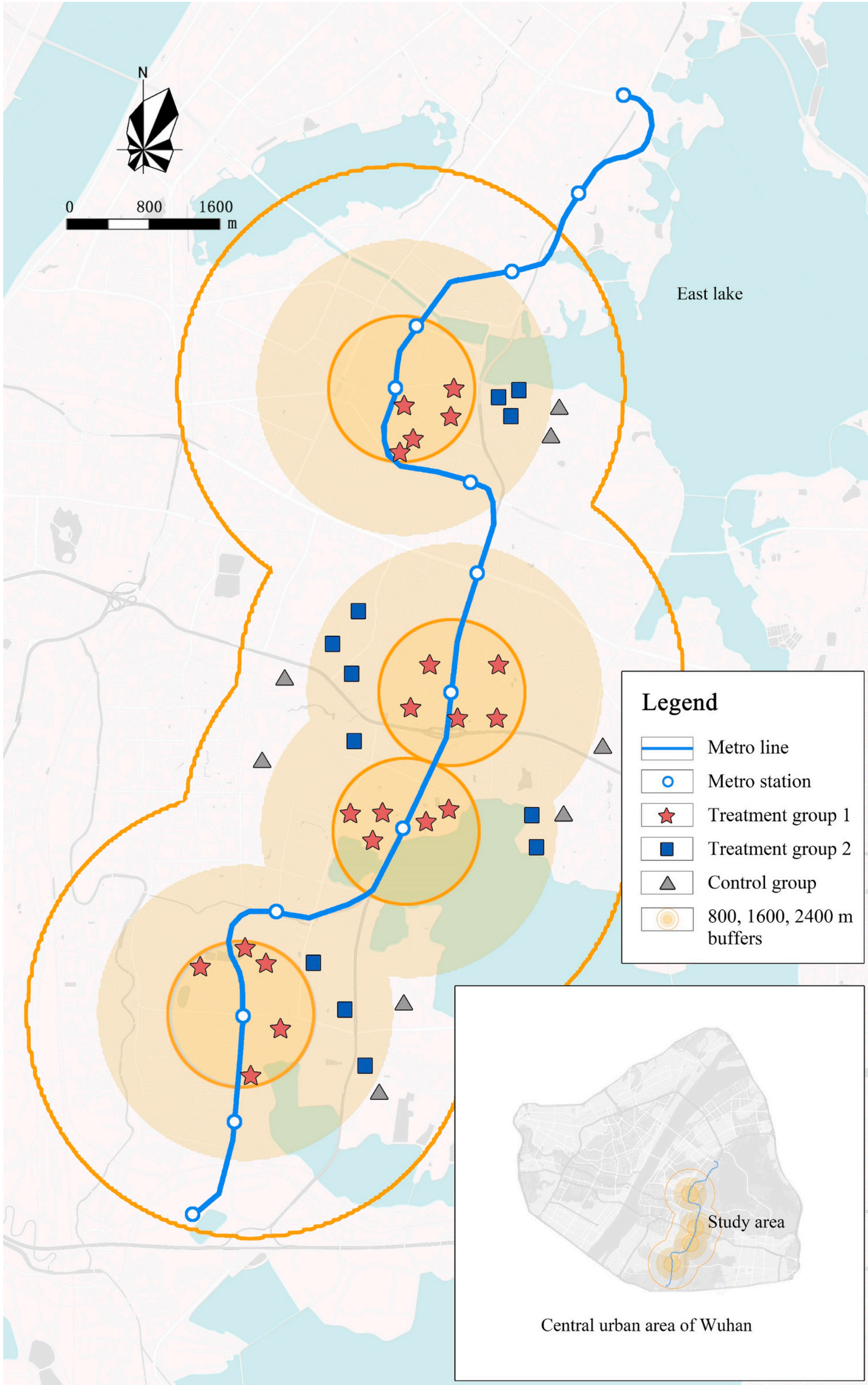

Fig. 3. Study sites for the treatment and control group.
Note: The buffer zone was calculated based on the street network distance, and was simplified for visualization.

was selected using the Kish grid method. Respondents were required to meet the following criteria: (1) be at least 19 years old, (2) have resided in the neighborhoods for at least a year to account for potential self-selection bias (Heinen et al., 2015), and (3) be capable of walking independently for a minimum of 15 min.

In all, the baseline survey gathered 908 respondents, out of which 422 completed the follow-up survey, leading to a retention rate of 46.5 %. Based on previous natural experiments conducted in this field and a power calculation using G*Power 3.1.9.7, we determined that a sample size of 85 for the control group would yield a desired power value of 0.90, assuming an effect size of 0.20 and an alpha error probability of 0.05 (Wang et al., 2023a; Freedman and Schatzkin, 1992). However, it is important to note that the effect size used in the calculation was determined based on the pooled effect size from a meta-analysis (Wang et al., 2023a). Therefore, caution should be exercised when interpreting results that yield effect sizes smaller than this specified value in our study analysis. Table 1 shows the socio-demographic characteristics and objective neighborhood attributes of the respondents who completed the two-wave questionnaire survey. The overall composition of the respondents showed a slightly higher representation of females compared to males, with an average age of 44–48 years, a bachelor's degree, a monthly household income of RMB 15,001-20,000, three family members, and a car. Table 1 indicates that there were no significant differences in socio-demographic characteristics among the three groups. However, noticeable disparities were observed in some objective neighborhood characteristics. It is essential to consider these variations in the upcoming analyses.

This study employs variables from five categories: the dependent variable - travel behavior; and the control variables which might influence the respondents' travel behavior, including objective and perceived neighborhood characteristics, travel attitudes, and socio-demographic factors. Travel behavior was assessed by asking respondents to detail their frequency and duration of trips in six modes of transit: rail, bus, car, e-bike, cycling, and walking, encompassing both work and non-work trips. Specifically, respondents were requested to provide the number of days in the last month during which each mode of travel was used for both work and non-work related trips. Responses were collected using a six-point scale, indicating frequency with the following options: (1) “never,” (2) “less than once a month,” (3) “one to three times a month,” (4) “once a week,” (5) “two to three times a week,” and (6) “four to five times a week.” To facilitate statistical analysis, these response categories were transformed into numerical values that represent the estimated average number of times per month, specifically ‘0,’ ‘0.5,’ ‘2,’ ‘4.3,’ ‘10.75,’ and ‘19.35’ times, respectively. Trip duration was measured by the time respondents spent in each mode of travel on a typical workday and a day off. This data was gathered through trip diaries maintained by the respondents, which recorded all their travel activities (movements from one location to another) over a 24-h period.

The objective neighborhood characteristics encompassed neighborhood population density, land use mix, number of street intersections, number of bus stops, and number of commercial facilities. The number of street intersections, bus stops, and commercial facilities were counted within an 800-m radius from the main entrance of neighborhoods. This distance is generally regarded as walkable for the majority of individuals and is commonly employed in urban planning to assess the accessibility of an area (Zacharias and Zhao, 2018).

The description of perceived neighborhood characteristics was adapted from Huang et al. (2019), who examined the impact of rail transit on car use in Xi'an, China. Following the “D-variables” framework for the built environment (Ewing and Cervero, 2001), this study selected seven indicators related to three dimensions for the built environment: transit, accessibility, and walkability (Table 2). Respondents were asked to assess the extent to which the built environment attributes of their neighborhoods were accurate using a five-point scale. The scale ranged from (1) “not at all true” to (5) “entirely true.”

To gauge individuals' attitudes towards travel, respondents were prompted to indicate the extent to which they agreed with a statement about each mode of travel. This was done using a five-point scale that ranged from (1) “strongly disagree” to (5) “strongly agree”. The statement “I prefer to travel by rail/bus/car/e-bike/bike/foot” was developed by Stark and Hössinger (2018), who explored the relationship between travel attitudes and mode choice.

The survey also collected respondents' socio-demographic details, which include age, gender, level of education, household income, household size, and car ownership.

Table 1
Descriptive statistics of respondents ($N = 422$).

	Group 1 ($N = 236$) n (%)/Mean (SD)	Group 2 ($N = 101$) n (%)/Mean (SD)	Group 3 ($N = 85$) n (%)/Mean (SD)	*P*-value
Socio-demographics				
Male[a]	108 (45.8)	47 (46.5)	43 (50.6)	0.744
Age	46.962 (14.837)	48.436 (14.918)	43.506 (14.393)	0.064
Education level	3.814 (1.414)	3.871 (1.310)	4.200 (1.263)	0.080
Household income	5.720 (1.682)	5.713 (1.972)	6.176 (1.935)	0.117
Household size	3.263 (1.298)	3.000 (1.342)	3.235 (1.352)	0.238
Car ownership	0.856 (0.680)	0.723 (0.677)	0.729 (0.640)	0.145
Built environment				
Neighborhood population density	2.623 (0.845)	1.963 (1.159)	2.143 (1.315)	0.003*
Land use mix	1.475 (0.212)	1.451 (0.241)	1.569 (0.002)	0.060
Number of street intersections	17.608 (7.170)	19.425 (6.607)	16.429 (3.332)	0.057
Number of bus stops	24.076 (4.136)	22.082 (2.019)	22.265 (2.380)	0.137
Number of commercial facilities	625.677 (181.780)	564.149 (59.101)	593.429 (43.316)	0.034*

Note:
(a) Group 1: the first treatment group (respondents who lived within a distance of 800 m from the stations). Group 2: the second treatment group (respondents who lived within a distance of 800 to 1600 m from the stations). Group 3: the control group (respondents who lived within a distance of 1600 to 2400 mfrom the stations).
(b) The educational level was assessed using a six-point scale, ranging from (1) elementary school or lower, (2) junior high school, (3) high school (including vocational school), (4) junior college, (5) bachelor's degree, to (6) master's degree or higher.
(c) Monthly household income was assessed using an eleven-point scale, ranging from (1) ¥1500 or less, (2) ¥1501-3000, (3) ¥3001-5000, (4) ¥5001-8000, (5) ¥8001-10,000, (6) ¥10,001-15,000, (7) ¥15,001-20,000, (8) ¥20,001-25,000, (9) ¥25,001-30,000, (10) ¥30,001-40,000, to (11) ¥40,000 or more.
(d) Car ownership was assessed at the household level.
(e) Neighborhood population density was measured in units of 10,000 persons per square km.
(f) The count of street intersections, bus stops, and commercial facilities was within an 800 m buffer from neighborhoods' main entrances.
(g) ANOVA was used to assess significant differences among the three groups ([a] Chi-square test), *: $p < 0.05$.

3.3. Statistical analysis

This study used two cross-sectional analysis methods using follow-up dataset only. The aim was to investigate whether varying cross-sectional analyses could produce results consistent with those obtained from natural experiments.

3.3.1. Cross-sectional analysis 1

In Cross-sectional analysis 1, a multilevel regression model was employed to estimate the association between urban rail transit and

Table 2
Measurements of perceived neighborhood characteristics.

Built environment attributes	Statements
Transit	Easy access to transit station
Accessibility	Easy access to parks and open spaces Easy access to shopping areas Easy access to downtown Easy access to workplace
Walkability	Connected sidewalks within the neighborhood Connected bike routes around the neighborhood

travel behavior. Multilevel regression is a statistical method used to analyze data with a nested structure (Bickel, 2007). In this study, for example, residents residing in the same neighborhood often exhibit more similarities in travel habits compared to individuals randomly selected from the overall population. Traditional multivariate regression techniques treat the units of analysis as independent observations and overlook unobserved characteristics within the hierarchical structure. In contrast, multilevel regression enables modeling the effects of predictors at each level and takes into account the association and dependence between observations within the same level, such as the clustering effect of travel behaviors in neighborhoods within this study. To enhance the credibility of the model, this analysis incorporated additional variables (such as socio-demographics, objective and perceived neighborhood characteristics and travel attitudes) into the multilevel regression model as control variables.

Individual respondents (level 1) were modeled to be clustered within neighborhoods (level 2) with random intercepts. Neighborhoods are geocoded areas, created by the transportation department of Wuhan for traffic analysis. This study adopted the following equations in the multilevel regression model:

$$\text{Level 1}: Y_{ij} = \alpha_{0j} + \sum_{n=1}^{N} \alpha_{nj} X_{nij} + r_{ij}$$

$$\text{Level 2}: \alpha_{0j} = \gamma_{00} + \mu_{0j}$$

$$\alpha_{1j} = \gamma_{10} + \mu_{1j}$$

……

$$\alpha_{nj} = \gamma_{n0} + \mu_{nj}$$

Where Y_{ij} is the travel behavior outcome of respondent i in neighborhood j, which can be the trip frequency and the trip duration of a certain travel mode of an individual, or the total trip frequency and duration of an individual. X_{nij} is the explanatory variable of respondent i in neighborhood j at level 1, including socio-demographics, objective and perceived neighborhood characteristics and travel attitudes. α_{0j}, α_{1j}, …,α_{nj} is the random intercept at level 1; γ is the regression coefficient to be estimated; r_{ij} is the random effect at level 1; μ is the random error term at level 2, representing unknown factors at level 2 that affect the travel behavior.

3.3.2. Cross-sectional analysis 2

In Cross-sectional analysis 2, we employed PSM method in combination with paired *t*-tests to assess the impact of the rail line. PSM can be used to reduce bias in the cross-sectional analysis by creating comparable treatment and control groups according to the propensity score (Austin, 2008). The propensity score is a numerical value that indicates the probability or likelihood of an individual being allocated to a treatment group. It is derived from observed characteristics or covariates such as age, gender and education. Once the propensity scores are estimated, the next step is to match individuals in the treatment group with those in the control group who possess similar propensity scores. Once the matching is completed, the effect of the rail transit can be estimated through comparing the outcomes between the matched pairs. It allows for a more valid comparison of the treatment effect by reducing the influence of potential confounding variables.

In this analysis, individuals in the treatment group 1 (Group 1) and the control group (Group 3), as well as individuals in the treatment group 2 (Group 2) and the control group (Group 3), were paired based on their socio-demographic and objective neighborhood characteristics. Two group pairs (Group pair 1–3 and Group pair 2–3), consisting of matched individuals who have statistically similar socio-demographic and objective neighborhood characteristics, were obtained for analysis. Table 3 shows the socio-demographic and objective neighborhood characteristics of the respondents after matching. The results of paired *t*-tests (Chi-square tests) showed that there was no significant difference in socio-demographic and objective neighborhood characteristics between the matched samples.

Then, paired *t*-tests were utilized to determine whether there were any significant differences in the travel behavioral changes between the matched group pairs. The purpose of this analysis was to attribute the

Table 3
Descriptive statistics of respondents after matching.

	Group pair 1–3			Group pair 2–3		
	Group 1 (N = 71) n(%)/Mean (SD)	Group 3 (*N* = 71) n(%)/Mean (SD)	*P*-value	Group 2 (N = 63) n(%)/Mean (SD)	Group 3 (*N* = 63) n(%)/Mean (SD)	P-value
Socio-demographics						
Male[a]	34 (47.9)	34 (47.9)	0.725	29 (46.0)	29 (46.0)	0.729
Age	45.770 (15.386)	45.506 (15.393)	0.412	45.928 (16.184)	45.500 (13.340)	0.175
Education level	4.085 (1.344)	4.440 (1.329)	0.518	4.200 (1.263)	4.183 (1.360)	0.662
Household income	5.607 (1.713)	5.858 (1.978)	0.291	6.176 (1.935)	6.467 (1.945)	0.143
Household size	3.289 (1.344)	3.235 (1.352)	0.907	3.176 (1.353)	3.367 (1.390)	0.324
Car ownership	0.785 (0.639)	0.729 (0.640)	0.636	0.751 (0.610)	0.783 (0.661)	0.624
Built environment						
Neighborhood population density	2.510 (1.298)	2.211 (1.206)	0.112	2.004 (1.206)	2.019 (1.205)	0.769
Land use mix	1.498 (0.189)	1.482 (0.213)	0.892	1.481 (0.213)	1.521 (0.212)	0.326
Number of street intersections	17.329 (6.490)	18.623 (6.058)	0.242	18.667 (6.068)	15.578 (6.046)	0.138
Number of bus stops	23.647 (2.872)	22.131 (2.123)	0.482	22.118 (2.124)	22.144 (2.122)	0.731
Number of commercial facilities	608.043 (60.792)	591.989 (58.816)	0.124	571.667 (56.825)	582.322 (57.806)	0.178

Note: Paired t-tests were applied to assess whether there were any significant differences in socio-demographic characteristics between the group pairs ([a] Chi-square test). *: $p < 0.05$.

observed differences specifically to the transit intervention treatment.

3.3.3. Natural experiment analysis

In the natural experiment analysis, we used a difference-in-differences (DiD) model to evaluate the effect of the transit intervention on travel behavior by comparing the before-and-after changes between the treatment and the control groups. The assumption of parallel trends is a fundamental component of the Difference-in-Differences (DiD) model (Craig et al., 2017). It posits that, in the absence of the treatment, the trends in the outcomes would have followed a similar pattern for both the treatment and control groups. By relying on this assumption, the DiD model enables the identification of the treatment effect by attributing any observed disparities specifically to the treatment.

The DiD model is implemented through regression analysis, where the treatment group, time indicators, and their interaction term are included as independent variables. The coefficient of the interaction term between the treatment variable and the time variable represents the estimated effect of the transit intervention. This study adopted the following equation in the DiD model:

$$Y_{it} = \beta_0 + \beta_1 T_i + \beta_2 P_t + \beta_3 T_i P_t + \varepsilon_{it}$$

Where Y_{it} is the travel behavior outcome of respondent i at time period t, T_i is the group dummy, which is equal to 1 if i belongs to the treatment group and 0 otherwise, and P_t is the time dummy, representing the time period (T_1 (0) and T_2 (1)). β_3 is the DiD estimator that indicates the treatment effect.

To prove the robustness of the conclusions, we further performed sensitivity analyses with different buffer ranges, changing the boundaries of the first treatment plots from 800 m to 700 m and 900 m, and the second treatment plots from 1600 m to 1500 m and 1700 m, respectively.

4. Results

4.1. Cross-sectional analysis 1

Tables 4 & 5 present the multilevel regression model results for the association between rail transit and individuals' travel behavior (trip frequency and duration) after controlling for socio-demographic, objective and perceived neighborhood, and travel attitudinal characteristics.

4.1.1. Trip frequency

For rail trips, *p*-values were significant for non-work trips in both group pair 1–3 (coefficient = 5.569, p < 0.001) and group pair 2–3 (coefficient = 4.063, p < 0.001). For car trips, the significant associations between urban rail transit and trip frequency were found in group pair 1–3 for both work (coefficient = −3.541, p = 0.005) and non-work trips (coefficient = −1.834, p = 0.030). For cycling trips, p-values were significant in group pair 1–3 for both work (coefficient = −1.924, p = 0.038) and non-work trips (coefficient = −1.693, p = 0.045), and in group pair 2–3 for non-work trips (coefficient = −2.036, p = 0.030). For walking trips, only pair 1–3 showed significant associations for non-work trips (coefficient = −2.729, p = 0.021). For bus, e-bike and total trips, there was no significant association between urban rail transit and trip frequency.

4.1.2. Trip duration

For rail trips, p-values were significant only for work trips in group pair 2–3 (coefficient = 11.587, p = 0.042). For car trips, only group pair 1–3 showed significant associations between urban rail transit and trip duration for work trips (coefficient = −10.142, p = 0.009). For walking trips, the *p*-value was significant only in group pair 1–3 for non-work trips (coefficient = −7.511, p = 0.005). The analysis did not find a statistically significant relationship between urban rail transit and trip frequency for bus, e-bike, cycling and total trips.

Table 4
Results of multilevel regression analysis for urban rail transit and trip frequency.

Travel mode	Trip purpose	Group pair	Coefficient [95 % CI]	*P*-value
Rail	Work	1–3	2.490 [−0.151, 5.130]	0.066
		2–3	1.292 [−1.264, 3.847]	0.480
	Non-work	1–3	5.569 [3.757, 7.380]	<0.001***
		2–3	4.063 [1.822, 6.304]	<0.001***
Bus	Work	1–3	−0.372 [−2.817, 2.073]	0.146
		2–3	−1.171 [−3.892, 1.550]	0.596
	Non-work	1–3	−0.578 [−2.502, 1.346]	0.833
		2–3	−1.198 [−3.135, 0.739]	0.336
Car	Work	1–3	−3.541 [−5.834–1.247]	0.005**
		2–3	−2.105 [−4.414, 0.204]	0.071
	Non-work	1–3	−1.834 [−3.381, −0.286]	0.030*
		2–3	0.117 [−1.565, 1.799]	0.337
E-bike	Work	1–3	−0.543 [−2.566, 1.481]	0.896
		2–3	0.647 [−1.420, 2.715]	0.806
	Non-work	1–3	−0.554 [−1.859, 0.751]	0.606
		2–3	0.077 [−1.588, 1.741]	0.392
Bicycle	Work	1–3	−1.924 [−3.609, −0.239]	0.038*
		2–3	−1.277 [−3.219, 0.665]	0.294
	Non-work	1–3	−1.693 [−3.221, −0.165]	0.045*
		2–3	−2.036 [−3.744, −0.328]	0.030*
Walk	Work	1–3	−1.385 [−4.000, 1.230]	0.446
		2–3	−0.894 [−3.785, 1.996]	0.813
	Non-work	1–3	−2.729 [−4.903, −0.555]	0.021*
		2–3	0.473 [−1.479, 2.426]	0.950
Total	Work	1–3	−5.275 [−10.849,0.298]	0.065
		2–3	−3.508 [−9.599, 2.583]	0.386
	Non-work	1–3	−1.819 [−7.070, 3.431]	0.744
		2–3	1.995 [−3.420, 7.411]	0.702

Note: *: p < 0.05, **: p < 0.01, ***: p < 0.001. 95 % CI: 95 % coefficient interval.

4.2. Cross-sectional analysis 2

Tables 6 & 7 present the paired *t*-test results of the before-after changes in trip frequency between the matched samples of the treatment and control groups.

4.2.1. Trip frequency

For rail trips, a significant inter-group difference was found between Group 1 and Group 3 for both work (difference = 3.192 times, p = 0.006) and non-work trips (difference = 5.520 times, p = 0.008). Car trips showed a significant inter-group difference between Group 1 and Group 3 only for non-work trips (difference = −3.312 times, p = 0.012). The analysis revealed that there were no statistically significant differences between the treatment and control groups in terms of bus trips, e-bike trips, cycling trips, walking trips, and total trips.

4.2.2. Trip duration

As with trip frequency, a significant inter-group difference in rail use was found for both work and non-work trips between Group 1 and Group 3 (work trips, difference = 5.719 mins, p = 0.006; and non-work trips, difference = 9.507 mins, p = 0.031). For walking trips, a significant inter-group difference was found only for non-work trips between Group 1 and Group 3 (difference = −9.014 mins, p = 0.007). The analysis found that there were no statistically significant differences between the treatment and control groups in terms of bus trips, car trips, e-bike trips, cycling trips, and total trips.

Table 5
Results of multilevel regression analysis for urban rail transit and trip duration.

Travel mode	Trip purpose	Group pair	Coefficient [95 % CI]	P-value
Rail	Work	1–3	3.149 [−1.562, 7.860]	0.066
		2–3	11.587 [0.100, 23.075]	0.042*
	Non-work	1–3	5.299 [−3.064, 13.663]	0.320
		2–3	−0.366 [−8.122, 7.391]	0.973
Bus	Work	1–3	3.995 [−7.651, 15.642]	0.749
		2–3	2.156 [−12.507, 16.818]	0.958
	Non-work	1–3	1.350 [−8.281, 10.982]	0.975
		2–3	4.369 [−6.721, 15.460]	0.651
Car	Work	1–3	−10.142 [−18.614, −1.671]	0.009**
		2–3	−8.776 [−19.605, 2.054]	0.167
	Non-work	1–3	−6.485 [−15.227, 2.257]	0.218
		2–3	−2.155 [−14.340, 10.030]	0.961
E-bike	Work	1–3	−0.892 [−6.893, 5.109]	0.955
		2–3	−2.169 [−6.655, 2.318]	0.512
	Non-work	1–3	1.949 [−0.778, 4.677]	0.242
		2–3	1.010 [−2.059, 4.078]	0.795
Bicycle	Work	1–3	−1.881 [−4.740, 0.979]	0.254
		2–3	−1.656 [−6.101, 2.790]	0.695
	Non-work	1–3	0.156 [−1.925, 2.238]	0.925
		2–3	0.481 [−0.865, 1.827]	0.716
Walk	Work	1–3	−0.926 [−7.393, 5.542]	0.967
		2–3	−1.646 [−7.682, 4.391]	0.887
	Non-work	1–3	−7.511 [−12.448, −2.573]	0.005**
		2–3	−0.737 [−7.188, 5.714]	0.963
Total	Work	1–3	−10.697 [−13.385, 3.991]	0.147
		2–3	−0.503 [−15.653, 14.648]	0.922
	Non-work	1–3	−5.241 [−17.073, 6.592]	0.576
		2–3	3.083 [−11.032, 17.198]	0.923

Note: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$. 95 % CI: 95 % coefficient interval.

Table 6
Results of paired *t*-test analysis for before-after changes in trip frequency.

Travel mode	Trip purpose	Group pair	Difference in before-after changes (SD)	P-value
Rail	Work	1–3	3.192 (−8.510)	0.006**
		2–3	1.517 (9.861)	0.263
	Non-work	1–3	5.520 (8.164)	0.008**
		2–3	1.360 (8.790)	0.243
Bus	Work	1–3	−3.781 (−12.991)	0.132
		2–3	−1.048 (10.951)	0.485
	Non-work	1–3	−3.252 (12.797)	0.066
		2–3	−2.248 (10.709)	0.115
Car	Work	1–3	−2.008 (−8.539)	0.081
		2–3	−0.699 (9.437)	0.588
	Non-work	1–3	−3.312 (8.341)	0.012*
		2–3	−1.004 (10.836)	0.483
E-bike	Work	1–3	0.156 (−7.732)	0.879
		2–3	0.187 (7.018)	0.845
	Non-work	1–3	−0.373 (7.327)	0.669
		2–3	−2.428 (8.908)	0.142
Bicycle	Work	1–3	0.616 (−7.436)	0.534
		2–3	−1.570 (7.707)	0.140
	Non-work	1–3	1.048 (8.008)	0.274
		2–3	−3.086 (9.187)	0.113
Walk	Work	1–3	−0.683 (−11.917)	0.667
		2–3	1.757 (9.839)	0.195
	Non-work	1–3	−0.958(10.939)	0.463
		2–3	1.980 (8.834)	0.193
Total	Work	1–3	−2.508 (−27.083)	0.487
		2–3	0.144 (22.106)	0.962
	Non-work	1–3	−4.147 (25.463)	0.174
		2–3	−5.426 (23.642)	0.186

Note: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

Table 7
Results of paired t-test analysis for before-after changes in trip duration.

Travel mode	Trip purpose	Group pair	Difference in before-after changes (SD)	P-value
Rail	Work	1–3	5.719 (43.600)	0.006**
		2–3	4.694 (44.924)	0.468
	Non-work	1–3	9.507 (36.431)	0.031*
		2–3	3.333 (30.161)	0.384
Bus	Work	1–3	−6.351 (50.031)	0.342
		2–3	3.959 (57.206)	0.630
	Non-work	1–3	−9.380 (54.835)	0.154
		2–3	1.651 (54.885)	0.812
Car	Work	1–3	−1.053 (3.244)	0.747
		2–3	−2.245 (30.191)	0.605
	Non-work	1–3	−1.549 (32.881)	0.693
		2–3	−1.667 (63.016)	0.834
E-bike	Work	1–3	2.281 (24.490)	0.483
		2–3	1.082 (26.783)	0.779
	Non-work	1–3	1.479 (19.864)	0.532
		2–3	−0.317 (17.692)	0.887
Bicycle	Work	1–3	1.491 (12.391)	0.367
		2–3	−1.020 (5.101)	0.168
	Non-work	1–3	2.930 (15.318)	0.112
		2–3	−0.317 (15.342)	0.870
Walk	Work	1–3	−4.000 (22.921)	0.193
		2–3	−5.163 (27.4599)	0.194
	Non-work	1–3	−9.014 (36.692)	0.007**
		2–3	−4.619 (39.002)	0.351
Total	Work	1–3	−10.351 (48.287)	0.111
		2–3	−8.082 (52.600)	0.288
	Non-work	1–3	−6.028 (64.294)	0.432
		2–3	−1.937 (76.836)	0.842

Note: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

Table 8
Results of difference-in-differences analysis for trip frequency.

Travel mode	Trip purpose	Group pair	DiD estimator [95 % CI]	P-value
Rail	Work	1–3	3.241 [1.57, 4.91]	0.048*
		2–3	1.268 [−0.54, 3.08]	0.452
	Non-work	1–3	5.937 [4.74, 7.14]	<0.001***
		2–3	4.081 [2.62, 5.54]	0.001**
Bus	Work	1–3	−0.213 [−2.34, 1.92]	0.913
		2–3	−0.588 [−2.76, 1.58]	0.772
	Non-work	1–3	−0.303 [−2.45, 1.84]	0.839
		2–3	−0.895 [−3.10, 1.31]	0.556
Car	Work	1–3	−3.634 [−5.03, −2.24]	0.071
		2–3	−0.622 [−2.10, 0.85]	0.764
	Non-work	1–3	−2.730 [−4.31, −1.15]	0.062
		2–3	0.169 [−1.59, 1.93]	0.912
E-bike	Work	1–3	−0.342 [−1.98, 1.29]	0.855
		2–3	0.781 [−0.93, 2.49]	0.686
	Non-work	1–3	−0.571 [−2.44, 1.30]	0.680
		2–3	0.762 [−1.18, 2.71]	0.604
Bicycle	Work	1–3	−0.463 [−2.41, 1.48]	0.781
		2–3	0.684 [−1.33, 2.70]	0.692
	Non-work	1–3	0.799 [−1.24, 2.84]	0.612
		2–3	0.813 [−1.30, 2.93]	0.611
Walk	Work	1–3	−4.632 [−6.26, −3.01]	0.019*
		2–3	−3.340 [−5.19, −1.49]	0.098
	Non-work	1–3	−3.353 [−4.86, −1.85]	0.027*
		2–3	−1.559 [−2.89, −0.23]	0.298
Total	Work	1–3	−6.043 [−14.92, 2.83]	0.221
		2–3	−1.817 [−6.89, 3.26]	0.717
	Non-work	1–3	−0.220 [−5.70, 5.26]	0.960
		2–3	3.371 [−2.45, 9.20]	0.454

Note: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$. 95 % CI: 95 % coefficient interval.

4.3. Natural experiment analysis

Tables 8 & 9 present the DiD analysis of inter-group comparisons of trip frequency and duration after controlling for socio-demographic characteristics and objective neighborhood attributes.

4.3.1. Trip frequency

For rail use, the significant inter-group difference was found for non-work trips between two group pairs (group pair 1–3, difference = 5.937 times, $p < 0.001$; and group pair 2–3, difference = 4.081 times, $p = 0.001$). Meanwhile, the analysis indicated a significant inter-group difference for work trips, specifically between Group 1 and Group 3 (difference = 3.241 times, $p = 0.048$). For walking trips, the comparisons between Group 1 and Group 3 presented a significant difference for both work (difference = −4.632 times, $p = 0.019$) and non-work trips (difference = −3.353 times, $p = 0.027$). The analysis revealed that there were no statistically significant between-group differences for bus trips, car trips, cycling trips, e-bike trips, and total trips.

4.3.2. Trip duration

For rail and car trips, Group 1 showed a significant difference from Group 3 for work trips. For walking trips, Group 1 showed a significant difference from Group 3 for non-work trips. However, the analysis indicated that there was no statistically significant difference between Group 2 and Group 3. The difference and p values of these trips were: rail trips for work purpose, 6.230 min, p = 0.038; car trips for work purpose, −10.499 min, $p = 0.026$; walking trips for non-work purpose, −13.440 min, $p = 0.006$. The analysis revealed no statistically significant inter-group differences in trip duration for bus trips, cycling trips, e-bike trips, and total trips.

Sensitivity analyses with different buffer ranges showed consistency with the main results (Appendix A).

Table 9
Results of difference-in-differences analysis for trip duration.

Travel mode	Trip purpose	Group pair	DiD estimator [95 % CI]	P-value
Rail	Work	1–3	6.230 [1.65, 10.81]	0.038*
		2–3	6.257 [1.15, 11.36]	0.250
	Non-work	1–3	8.712 [1.94, 15.48]	0.068
		2–3	−0.017 [−7.15, 7.11]	0.997
Bus	Work	1–3	3.508 [−5.33, 12.35]	0.666
		2–3	5.858 [−3.21, 14.92]	0.950
	Non-work	1–3	−5.446 [−15.23, 4.34]	0.337
		2–3	−0.403 [−11.72, 10.92]	0.108
Car	Work	1–3	−10.499 [−18.42, −2.58]	0.026*
		2–3	−10.196 [−15.86, −4.53]	0.355
	Non-work	1–3	−11.955 [−19.63, −4.28]	0.079
		2–3	−4.783 [−13.81, 4.24]	0.524
E-bike	Work	1–3	5.027 [−2.26, 12.26]	0.345
		2–3	4.588 [−2.05, 11.22]	0.384
	Non-work	1–3	0.830 [−2.25, 3.91]	0.692
		2–3	−0.603 [−4.22, 3.01]	0.789
Bicycle	Work	1–3	−0.516 [−2.30, 1.27]	0.778
		2–3	−0.756 [−2.89, 1.38]	0.756
	Non-work	1–3	2.739 [−0.36, 5.84]	0.182
		2–3	1.268 [−2.14, 4.67]	0.408
Walk	Work	1–3	−7.840 [−13.24, −2.44]	0.099
		2–3	−5.374 [−9.61, −1.14]	0.269
	Non-work	1–3	−13.440 [−21.32, −5.56]	0.006**
		2–3	−6.341 [−14.80, 2.11]	0.217
Total	Work	1–3	−4.090 [−15.05, 6.87]	0.687
		2–3	0.377 [−9.85, 10.60]	0.972
	Non-work	1–3	−18.560 [−31.73, −5.39]	0.122
		2–3	−12.073 [−26.96, 2.81]	0.176

Note: *: p < 0.05, **: p < 0.01, ***: p < 0.001. 95 % CI: 95 % coefficient interval.

4.4. The comparison between cross-sectional analyses and natural experiment

Table 10 shows the comparison between the results of the two cross-sectional analyses and the natural experiment. For the comparison between Cross-sectional analysis 1 and the natural experiment, of the 28 results analyzed, there were 15 different results and 13 identical ones. The main differentiation was that Cross-sectional analysis 1 revealed significant evidence for car and cycling trips, whereas the natural experiment did not detect any effect. Additionally, Cross-sectional analysis 2 indicated insignificant evidence for walking trips, while the natural experiment yielded significant evidence.

To enhance the models' robustness, we employed a more stringent p-value threshold (0.01) in the cross-sectional analyses and then compared the adjusted results to those of the natural experiment. Table 11 shows the comparison between the results of the two cross-sectional analyses (adjusted for P-value) and the natural experiment. Using a stricter significance level, the results of the cross-sectional analyses were closer to the natural experiment. Of the 28 results analyzed, Cross-sectional analyses 1 and 2 differed from the natural experiment in 5 and 4 results, respectively.

Table 12 presents the effect sizes of significant results from both the cross-sectional analysis (adjusted for P-value) and the natural experiment, allowing for a further comparison of the magnitudes of effect between the two methods. The results showed that the natural experiment analysis yielded slightly larger effect sizes compared to the cross-sectional analyses.

5. Discussion

5.1. Consistency between cross-sectional analyses and natural experiment

The results for the cross-sectional analyses and natural experiment came to a similar conclusion, showing that the new rail line resulted in an increase in the use of rail transit and a decrease in driving and walking, while there was no significant effect on the use of other modes and total trips. And the impact distance threshold was 800 m for most travel modes.

Specifically, for rail trips, the impact of the new metro system on the frequency of non-work rail trips extended up to a distance of 1600 m. This implies that individuals are willing to travel longer distances to access rail services for non-commuting purposes. In China, cycling has been recognized as a crucial mode of transportation for covering the first and last miles to and from rail stations, particularly in suburban areas with limited public transportation options (Deng and Zhao, 2022; Sun et al., 2020). Thus, there is a growing recognition of the need to consider cycling distances when defining the catchment areas of rail stations, aligning with broader urban planning objectives of transit-oriented development.

For car trips, the rail transit system only exhibited an impact on car trips with work purposes, while it had no influence on non-work trips. These findings suggest that rail transit appears to have a greater impact on reducing car travel for work purposes. This could be attributed to factors such as rush hour congestion, which may prompt individuals to consider alternative commuting modes (Ben-Elia and Ettema, 2011; De Vos et al., 2012). To reduce non-work car trips and encourage a shift towards public transportation, it is crucial to implement additional public transit infrastructure optimizations and policy management measures, such as parking management. These initiatives aim to facilitate the transition from car usage to public transit by improving accessibility and convenience.

Similar findings were observed for walking trips, indicating that non-work walking behavior was significantly influenced. These findings present a discrepancy compared to previous studies that have indicated that proximity to new metro stations can promote active travel, particularly walking (Huang et al., 2017). One possible explanation could be

Table 10
The comparison between the results of the cross-sectional analyses and the natural experiment.

Travel mode	Trip purpose	Group pair	Trip frequency			Trip duration		
			Cross-sectional analysis1	Cross-sectional analysis2	Natural experiment	Cross-sectional analysis1	Cross-sectional analysis2	Natural experiment
Rail	Work	1–3	**0**	+	+	**0**	+	+
		2–3	0	0	0	+	0	0
	Non-work	1–3	+	+	+	0	+	0
		2–3	+	**0**	+	0	0	0
Bus	Work	1–3	0	0	0	0	0	0
		2–3	0	0	0	0	0	0
	Non-work	1–3	0	0	0	0	0	0
		2–3	0	0	0	0	0	0
Car	Work	1–3	−	0	0	−	**0**	−
		2–3	0	0	0	0	0	0
	Non-work	1–3	−	−	0	0	0	0
		2–3	0	0	0	0	0	0
E-bike	Work	1–3	0	0	0	0	0	0
		2–3	0	0	0	0	0	0
	Non-work	1–3	0	0	0	0	0	0
		2–3	0	0	0	0	0	0
Bicycle	Work	1–3	−	0	0	0	0	0
		2–3	0	0	0	0	0	0
	Non-work	1–3	−	0	0	0	0	0
		2–3	−	0	0	0	0	0
Walk	Work	1–3	**0**	**0**	−	0	0	0
		2–3	0	0	0	0	0	0
	Non-work	1–3	−	**0**	−	−	−	−
		2–3	0	0	0	0	0	0
Total	Work	1–3	0	0	0	0	0	0
		2–3	0	0	0	0	0	0
	Non-work	1–3	0	0	0	0	0	0
		2–3	0	0	0	0	0	0

Note: Significant evidence was marked (+) or (−) to indicate the direction of the association, and (0) to indicate insignificant results. The different outcomes of the cross-sectional analyses and the natural experiment were bolded for legibility.

Table 11
The comparison between the results of the cross-sectional analyses (adjusted for *P*-value) and the natural experiment.

Travel mode	Trip purpose	Group pair	Trip frequency			Trip duration		
			Cross-sectional analysis1	Cross-sectional analysis2	Natural experiment	Cross-sectional analysis1	Cross-sectional analysis2	Natural experiment
Rail	Work	1–3	**0**	+	+	**0**	+	+
		2–3	0	0	0	0	0	0
	Non-work	1–3	+	+	+	0	0	0
		2–3	+	**0**	+	0	0	0
Bus	Work	1–3	0	0	0	0	0	0
		2–3	0	0	0	0	0	0
	Non-work	1–3	0	0	0	0	0	0
		2–3	0	0	0	0	0	0
Car	Work	1–3	−	0	0	−	**0**	−
		2–3	0	0	0	0	0	0
	Non-work	1–3	0	0	0	0	0	0
		2–3	0	0	0	0	0	0
E-bike	Work	1–3	0	0	0	0	0	0
		2–3	0	0	0	0	0	0
	Non-work	1–3	0	0	0	0	0	0
		2–3	0	0	0	0	0	0
Bicycle	Work	1–3	0	0	0	0	0	0
		2–3	0	0	0	0	0	0
	Non-work	1–3	0	0	0	0	0	0
		2–3	0	0	0	0	0	0
Walk	Work	1–3	**0**	**0**	−	0	0	0
		2–3	0	0	0	0	0	0
	Non-work	1–3	**0**	**0**	−	−	−	−
		2–3	0	0	0	0	0	0
Total	Work	1–3	0	0	0	0	0	0
		2–3	0	0	0	0	0	0
	Non-work	1–3	0	0	0	0	0	0
		2–3	0	0	0	0	0	0

Note: Strong evidence was marked (+) or (−) to indicate the direction of the association, and (0) to indicate insignificant results. The different outcomes of the cross-sectional analyses and the natural experiment were bolded for legibility.

Table 12
The comparison of the effect sizes of the cross-sectional analyses (adjusted for P-value) and the natural experiment.

Travel mode	Trip purpose	Group pair	Trip frequency			Trip duration		
			Cross-sectional analysis1	Cross-sectional analysis2	Natural experiment	Cross-sectional analysis1	Cross-sectional analysis2	Natural experiment
Rail	Work	1–3	/	3.192	3.241	/	5.719	6.230
	Non-work	1–3	5.569	5.520	5.973	/	/	/
		2–3	4.063	/	4.081	/	/	/
Car	Work	1–3	/	/	/	−10.142	/	−10.499
		2–3	/	/	/	/	/	/
	Non-work	1–3	/	/	/	/	/	/
		2–3	/	/	/	/	/	/
Walk	Work	1–3	/	/	/	/	/	/
		2–3	/	/	/	/	/	/
	Non-work	1–3	−2.729	/	−3.353	−7.511	−9.014	−13.440
		2–3	/	/	/	/	/	/

Note: /: Results that were not significant or lacked significant results for comparison were not included in the analysis.

that the introduction of new rail transit options improves the convenience of daily travel from residential areas to various destinations, such as restaurants or shopping malls. Furthermore, the pedestrian walking environment around rail stations in Wuhan suffers from low quality, primarily due to inadequate sidewalk infrastructure and challenging climatic conditions. As a result, some residents may opt for rail transit due to its comfort and speed advantages over walking. Therefore, it is essential to integrate comprehensive urban planning interventions around rail stations (such as pedestrian infrastructure) in conjunction with urban rail investments to achieve a synergistic effect in enhancing sustainable travel behavior.

5.2. Discrepancy between cross-sectional analyses and natural experiment

Comparing the results of the cross-sectional analyses and the natural experiment indicated that improving the model for the cross-sectional analyses (controlling for additional variables or combining it with the PSM method), while setting a stricter p-value (0.01), can make the cross-sectional analysis produce similar results to those of the natural experiment. This suggests that a well-designed model combined with a stricter significance level can help to enhance the consistency and reliability of the cross-sectional analysis.

Upon further comparison of effect sizes, it was found that the natural experiment analysis produced larger effect sizes in comparison to the cross-sectional analyses, indicating a more pronounced impact of the new rail line on travel behavior. One potential explanation is the causal nature of the study design, which helps isolate the effects of the rail line from other confounding factors, thus providing more robust evidence of the intervention's impact (Leatherdale, 2019). Moreover, the natural experiment design may capture unmeasured or latent factors that cross-sectional analyses might miss (Streeter et al., 2017). For instance, the anticipation and excitement surrounding the upcoming rail line, as well as the subsequent cultural and behavioral shifts within the community, could contribute to the observed larger effect sizes. These contextual factors, which are difficult to capture in cross-sectional studies, may have a significant influence on travel behavior.

To minimize discrepancies in effect results between cross-sectional analyses and natural experiments, optimizing cross-sectional studies can be achieved through various strategies. These include enhancing control over confounding factors, utilizing advanced statistical techniques like propensity score matching (PSM) or instrumental variable approaches, and conducting sensitivity analyses and robustness checks. By implementing these methodological enhancements, cross-sectional studies can yield more reliable and robust results, serving as a valuable complement to natural experiments when research resources are limited.

5.3. Limitations and future directions

There are several limitations to this study. Firstly, the reliance on self-reported data collected through travel surveys may introduce recall bias and other measurement biases. Future studies can enhance the assessment of individuals' travel behavior by combining subjective measurement methods, such as traditional surveys, with objective ones, such as GPS trajectory data. This approach allows for a more comprehensive and accurate tracking of travel behavior by capturing both self-reported information and real-time, location-based data. Secondly, the subjective neighborhood characteristics and travel attitudes were measured by a single item, which may limit the reliability of this variable. Future studies could consider employing multiple items and conducting Confirmatory Factor Analysis (CFA) to measure different aspects of subjective neighborhood characteristics and travel attitudes. Thirdly, the study's limited scope to a single city indeed restricts the findings' generalizability. In particular, more comparative empirical studies are required to further validate the feasibility and accuracy of cross-sectional studies, taking into account specific contextual factors.

6. Conclusion

This study used both cross-sectional and natural experiment analyses to evaluate the treatment effects of urban rail transit on individuals' travel behavior based on the same dataset. The findings from both the cross-sectional analyses and the natural experiment yielded a consistent outcome, indicating that the introduction of the new rail infrastructure resulted in an increase in rail transit usage and a decrease in both driving and walking. The comparison between the cross-sectional analyses and the natural experiment revealed that refining the model used in the cross-sectional analyses, combined with a more stringent p-value threshold, can lead to cross-sectional analysis results that closely align with those obtained from the natural experiment. These findings suggest that well-designed cross-sectional studies can offer reliable insights and serve as a cost-effective alternative to resource-intensive natural experiments.

Funding

The work described in this paper was supported by the Research Grants Council of the Hong Kong SAR (Project No. CityU11207520), the Humanities and Social Sciences Fund of Ministry of Education of China (Project No. 24YJC840031), the Postdoctoral Fellowship Program of CPSF (Project No. GZC20241261), the Shanghai Municipal Natural Science Foundation (Project No. 23ZR1465100), and the Fundamental Research Funds for the Central Universities of China.

CRediT authorship contribution statement

Jingjing Wang: Writing – original draft, Data curation. **Yi Lu:** Writing – review & editing, Project administration, Conceptualization. **Mi Diao:** Methodology. **Ye Liu:** Supervision.

Data availability

Data will be made available on request.

Appendix A. Appendix

Table A1
Sensitivity analysis: results of difference-in-differences analysis for trip frequency.

Travel mode	Trip purpose	Group pair	Sensitivity analysis 1		Sensitivity analysis 2	
			DiD estimator [95 % CI]	*P*-value	DiD estimator [95 % CI]	P-value
Rail	Work	1–3	0.880 [0.05, 1.72]	0.039*	1.048 [0.24, 1.86]	0.011*
		2–3	0.758 [−0.51, 1.53]	0.153	0.033 [−0.96, 1.03]	0.948
	Non-work	1–3	0.865 [0.20, 1.53]	0.011*	0.924 [0.27, 1.57]	0.005**
		2–3	0.758 [0.01, 1.51]	0.043*	0.473 [0.03, 0.92]	0.048*
Bus	Work	1–3	0.205 [−0.69, 1.11]	0.654	−0.278 [−1.23, 0.68]	0.568
		2–3	0.058 [−0.76, 0.88]	0.889	−0.381 [−1.51, 0.75]	0.506
	Non-work	1–3	−0.268 [−0.96, 0.43]	0.448	−0.357 [−1.07, 0.35]	0.325
		2–3	0.058 [−0.76, 0.88]	0.889	−0.334 [−1.17, 0.50]	0.432
Car	Work	1–3	−0.989 [−1.89, −0.09]	0.032*	−0.606 [−1.54, −0.03]	0.075
		2–3	−1.300 [−2.13, 1.47]	0.102	−0.055 [−1.19, 1.08]	0.923
	Non-work	1–3	−0.454 [−0.98, 0.07]	0.078	−0.535 [−1.09, 0.02]	0.067
		2–3	−1.300 [−3.07, 0.47]	0.102	0.360 [−0.54, 1.26]	0.430
E-bike	Work	1–3	−0.861 [−2.19, 0.32]	0.155	−0.125 [−1.06, 0.81]	0.793
		2–3	−1.024 [−2.26, 0.21]	0.114	−0.208 [−1.30, 0.89]	0.709
	Non-work	1–3	−0.214 [−0.92, 0.49]	0.551	−0.245 [−0.96, 0.47]	0.504
		2–3	−0.608 [−4.90, 3.68]	0.780	−0.598 [−0.94, 0.26]	0.172
Bicycle	Work	1–3	−0.174 [−2.58, 2.24]	0.887	−0.464 [−2.83, 1.90]	0.700
		2–3	−1.854 [−4.11, 0.41]	0.107	−1.556 [−4.89, 1.78]	0.359
	Non-work	1–3	1.055 [−1.18, 3.29]	0.353	1.091 [−1.44, 3.62]	0.397
		2–3	−1.853 [−4.11, 0.41]	0.107	−0.129 [−3.34, 3.09]	0.937
Walk	Work	1–3	−1.341 [−2.16, −0.53]	0.001**	−0.335 [−0.64, −0.03]	0.049*
		2–3	−1.402 [−2.19, −0.61]	0.001**	−0.224 [−1.28, 0.83]	0.677
	Non-work	1–3	−1.083 [−2.20, 0.03]	0.065	−0.308 [−0.67, 0.05]	0.076
		2–3	−1.402 [−3.41, 0.61]	0.101	−0.325 [−0.97, 0.32]	0.324
Total	Work	1–3	−6.562 [−13.41, 1.29]	0.160	−3.334 [−10.56, 3.89]	0.365
		2–3	−10.637 [−15.07, 1.95]	0.102	−5.040 [−14.95, 4.91]	0.319
	Non-work	1–3	0.767 [−6.19, 7.72]	0.829	−2.201 [−9.63, 5.23]	0.561
		2–3	−10.637 [−23.22, 1.95]	0.102	−2.361 [−11.55, 6.83]	0.613

Note:
(a) Sensitivity analysis 1: Group 1: the first treatment group (respondents who lived within a distance of 700 m from the stations). Group 2: the second treatment group (respondents who lived within a distance of 700 to 1500 m from the stations). Group 3: the control group (respondents who lived within a distance of 1500 to 2400 mfrom the stations).
(b) Sensitivity analysis 2: Group 1: the first treatment group (respondents who lived within a distance of 900 m from the stations). Group 2: the second treatment group (respondents who lived within a distance of 900 to 1700 m from the stations). Group 3: the control group (respondents who lived within a distance of 1700 to 2400 mfrom the stations).
(c) *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.
(d) 95 % CI: 95 % coefficient interval.

Table A2
Sensitivity analysis: results of difference-in-differences analysis for trip duration.

Travel mode	Trip purpose	Group pair	Sensitivity analysis 1		Sensitivity analysis 2	
			DiD estimator [95 % CI]	P-value	DiD estimator [95 % CI]	P-value
Rail	Work	1–3	0.530 [0.09, 0.97]	0.048*	4.207 [0.60, 7.81]	0.048*
		2–3	0.876 [−7.82, 10.07]	0.585	0.853 [−13.33, 15.04]	0.906
	Non-work	1–3	7.166 [−0.71, 15.04]	0.054	7.086 [−0.98, 15.16]	0.086
		2–3	7.994 [−2.73, 18.72]	0.172	−0.081 [−10.83, 10.67]	0.525
Bus	Work	1–3	−1.310 [−21.22, 18.60]	0.897	−2.024 [−19.60, 15.55]	0.821
		2–3	−0.043 [−17.56, 17.48]	0.996	−3.383 [−25.09, 18.33]	0.759
	Non-work	1–3	−7.892 [−20.50, 4.72]	0.219	−4.312 [−14.27, 5.65]	0.396
		2–3	−3.014 [−9.04, 15.07]	0.624	5.041 [−8.39, 18.48]	0.461
Car	Work	1–3	−5.646 [−12.01, 0.72]	0.056	−8.498 [−16.43, −0.57]	0.046*
		2–3	−4.116 [−22.91, 14.68]	0.667	−13.397 [−38.57, 11.77]	0.295
	Non-work	1–3	−2.651 [−5.44, 0.14]	0.072	−2.765 [−5.84, 0.31]	0.081
		2–3	−2.084 [−15.36, 11.19]	0.758	−1.010 [−11.79, 9.77]	0.558
E-bike	Work	1–3	2.347 [−5.20, 9.90]	0.541	3.745 [−5.16, 12.65]	0.409

(*continued on next page*)

Table A2 (*continued*)

Travel mode	Trip purpose	Group pair	Sensitivity analysis 1		Sensitivity analysis 2	
			DiD estimator [95 % CI]	P-value	DiD estimator [95 % CI]	P-value
		2–3	2.014 [−5.50, 9.53]	0.599	2.558 [−7.39, 12.51]	0.613
	Non-work	1–3	−1.474 [−6.20, 3.25]	0.540	−0.914 [−4.61, 2.78]	0.628
		2–3	−0.609 [−4.90, 3.68]	0.780	−0.587 [−5.33, 4.16]	0.808
Bicycle	Work	1–3	1.640 [−3.58, 6.86]	0.537	1.068 [−2.89, 5.03]	0.596
		2–3	0.038 [−4.56, 4.64]	0.987	−0.444 [−6.06, 5.17]	0.876
	Non-work	1–3	1.600 [−0.66, 3.86]	0.164	1.528 [−1.08, 4.14]	0.251
		2–3	−0.288 [−2.84, 2.27]	0.825	−0.548 [−4.43, 3.33]	0.781
Walk	Work	1–3	−3.386 [−7.16, 0.39]	0.098	−7.389 [−16.87, 2.09]	0.126
		2–3	−6.208 [−14.58, 2.17]	0.146	−8.420 [−19.37, 2.53]	0.131
	Non-work	1–3	−8.166 [−17.65, 1.32]	0.091	−10.329 [−19.87, −0.78]	0.034*
		2–3	−9.995 [−23.58, 3.59]	0.120	−7.547 [−19.21, 4.11]	0.204
Total	Work	1–3	−5.826 [−27.43, 15.78]	0.596	−8.715 [−28.14, 10.71]	0.378
		2–3	−12.192 [−31.73, 7.35]	0.221	−21.613 [−46.66, 3.44]	0.100
	Non-work	1–3	−11.418 [−28.35, 5.51]	0.186	−7.838 [−23.97, 8.29]	0.340
		2–3	−1.967 [−17.47, 13.54]	0.803	−2.066 [−21.24, 17.11]	0.832

Note:
(a) Sensitivity analysis 1: Group 1: the first treatment group (respondents who lived within a distance of 700 m from the stations). Group 2: the second treatment group (respondents who lived within a distance of 700 to 1500 m from the stations). Group 3: the control group (respondents who lived within a distance of 1500 to 2400 mfrom the stations).
(b) Sensitivity analysis 2: Group 1: the first treatment group (respondents who lived within a distance of 900 m from the stations). Group 2: the second treatment group (respondents who lived within a distance of 900 to 1700 m from the stations). Group 3: the control group (respondents who lived within a distance of 1700 to 2400 mfrom the stations).
(c) *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.
(d) 95 % CI: 95 % coefficient interval.

References

Austin, P.C., 2008. A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. Stat. Med. 27 (12), 2037–2049. https://doi.org/10.1002/sim.3150.

Ben-Elia, E., Ettema, D., 2011. Changing commuters' behavior using rewards: a study of rush-hour avoidance. Transport. Res. F: Traffic Psychol. Behav. 14 (5), 354–368.

Bickel, R., 2007. Multilevel Analysis for Applied Research: It's Just Regression! Guilford Press.

Cao, X., Schoner, J., 2014. The influence of light rail transit on transit use: an exploration of station area residents along the Hiawatha line in Minneapolis. Transp. Res. A Policy Pract. 59, 134–143. https://doi.org/10.1016/j.tra.2013.11.001.

Chatman, D.G., 2013. Does TOD need the T? J. Am. Plan. Assoc. 79 (1), 17–31. https://doi.org/10.1080/01944363.2013.791008.

Cheng, L., Jin, T., Wang, K., Lee, Y., Witlox, F., 2022. Promoting the integrated use of bikeshare and metro: a focus on the nonlinearity of built environment effects. Multim. Transp. 1 (1), 100004.

Craig, P., Katikireddi, S.V., Leyland, A., Popham, F., 2017. Natural experiments: an overview of methods, approaches, and contributions to public health intervention research. Annu. Rev. Public Health 38 (1), 39–56. https://doi.org/10.1146/annurev-publhealth-031816-044327.

De Vos, J., Derudder, B., Van Acker, V., Witlox, F., 2012. Reducing car use: changing attitudes or relocating? The influence of residential dissonance on travel behavior. J. Transp. Geogr. 22, 1–9.

Deng, Y., Zhao, P., 2022. The impact of new metro on travel behavior: panel analysis using mobile phone data. Transp. Res. A Policy Pract. 162, 46–57. https://doi.org/10.1016/j.tra.2022.05.013.

Du, Q., Zhou, Y., Huang, Y., Wang, Y., Bai, L., 2022. Spatiotemporal exploration of the non-linear impacts of accessibility on metro ridership. J. Transp. Geogr. 102, 103380. https://doi.org/10.1016/j.jtrangeo.2022.103380.

Ewing, R., Cervero, R., 2001. Travel and the built environment: a synthesis. Transp. Res. Rec. 1780 (1), 87–114. https://doi.org/10.3141/1780-10.

Freedman, L.S., Schatzkin, A., 1992. Sample size for studying intermediate endpoints within intervention trials or observational studies. Am. J. Epidemiol. 136 (9), 1148–1159. https://doi.org/10.1093/oxfordjournals.aje.a116581.

Heinen, E., Panter, J., Mackett, R., Ogilvie, D., 2015. Changes in mode of travel to work: a natural experimental study of new transport infrastructure. Int. J. Behav. Nutr. Phys. Act. 12 (1), 81. https://doi.org/10.1186/s12966-015-0239-8.

Huang, R.Z., Moudon, A.V., Zhou, C., Stewart, O.T., Saelens, B.E., 2017. Light rail leads to more walking around station areas. J. Transp. Health 6, 201–208. https://doi.org/10.1016/j.jth.2017.02.002.

Huang, X., Cao, X., Yin, J., Cao, X., 2019. Can metro transit reduce driving? Evidence from Xi'an, China. Transp. Policy 81, 350–359. https://doi.org/10.1016/j.tranpol.2018.03.006.

Kärmeniemi, M., Lankila, T., Ikäheimo, T., Koivumaa-Honkanen, H., Korpelainen, R., 2018. The built environment as a determinant of physical activity: a systematic review of longitudinal studies and natural experiments. Ann. Behav. Med. 52 (3), 239–251. https://doi.org/10.1093/abm/kax043.

Leatherdale, S.T., 2019. Natural experiment methodology for research: a review of how different methods can support real-world research. Int. J. Soc. Res. Methodol. 22 (1), 19–35. https://doi.org/10.1080/13645579.2018.1488449.

Melissa, Morrison, R.S., 2009. Study design, precision, and validity in observational studies. J. Palliat. Med. 12 (1), 77–82. https://doi.org/10.1089/jpm.2008.9690.

Nasri, A., Zhang, L., 2014. The analysis of transit-oriented development (TOD) in Washington, D.C. and Baltimore metropolitan areas. Transp. Policy 32, 172–179. https://doi.org/10.1016/j.tranpol.2013.12.009.

Spector, P.E., 2019. Do not cross me: optimizing the use of cross-sectional designs. J. Bus. Psychol. 34 (2), 125–137. https://doi.org/10.1007/s10869-018-09613-8.

Stark, J., Hössinger, R., 2018. Attitudes and mode choice: measurement and evaluation of interrelation. Transp. Res. Proc. 32, 501–512. https://doi.org/10.1016/j.trpro.2018.10.038.

Streeter, A.J., Lin, N.X., Crathorne, L., Haasova, M., Hyde, C., Melzer, D., Henley, W.E., 2017. Adjusting for unmeasured confounding in nonrandomized longitudinal studies: a methodological review. J. Clin. Epidemiol. 87, 23–34.

Sun, G., Zacharias, J., 2017. Can bicycle relieve overcrowded metro? Managing short-distance travel in Beijing. Sustain. Cities Soc. 35, 323–330. https://doi.org/10.1016/j.scs.2017.08.010.

Sun, G., Zhao, J., Webster, C., Lin, H., 2020. New metro system and active travel: a natural experiment. Environ. Int. 138, 105605. https://doi.org/10.1016/j.envint.2020.105605.

Tao, S., Cheng, L., He, S., Witlox, F., 2023. Examining the non-linear effects of transit accessibility on daily trip duration: a focus on the low-income population. J. Transp. Geogr. 109, 103600. https://doi.org/10.1016/j.jtrangeo.2023.103600.

Wang, X., Cheng, Z., 2020. Cross-sectional studies: strengths, weaknesses, and recommendations. Chest 158 (1, Supplement), S65–S71. https://doi.org/10.1016/j.chest.2020.03.012.

Wang, J., Cheng, W., Lu, Y., Wang, D., 2023a. Effect of rail transit on travel behavior: a systematic review and meta-analysis. Transp. Res. Part D: Transp. Environ. 122, 103882.

Wang, J., Lu, Y., Yang, Y., Peng, J., Liu, Y., Yang, L., 2023b. Influence of a new rail transit line on travel behavior: evidence from repeated cross-sectional surveys in Hong Kong. J. Transp. Geogr. 106, 103526. https://doi.org/10.1016/j.jtrangeo.2022.103526.

Wuhan Statistics Bureau, 2022. Wuhan Statistical Yearbook (2007–2021). China Statistics Press, Beijing, China.

Xie, L., 2016. Automobile usage and urban rail transit expansion: evidence from a natural experiment in Beijing, China. Environ. Dev. Econ. 21 (5), 557–580. https://doi.org/10.1017/s1355770x16000048.

Zacharias, J., Zhao, Q., 2018. Local environmental factors in walking distance at metro stations. Publ. Transp. 10 (1), 91–106. https://doi.org/10.1007/s12469-017-0174-y.

Zhong, H., Li, W., Boarnet, M.G., 2021. A two-dimensional propensity score matching method for longitudinal quasi-experimental studies: a focus on travel behavior and the built environment. Environ. Plann. B 2399808320982305. https://doi.org/10.1177/2399808320982305.