Transportation Research Part D 122 (2023) 103882

Contents lists available at ScienceDirect

Transportation Research Part D

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

Effect of rail transit on travel behavior: A systematic review and meta-analysis

Jingjing Wang [a,b,1], Weibin Cheng [c,d,1], Yi Lu [b,*], Donggen Wang [e]

[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] Institute for Healthcare Artificial Intelligence Application, Guangdong Second Provincial General Hospital, Guangzhou, China
[d] School of Data Science, City University of Hong Kong, Hong Kong, China
[e] Department of Geography, Hong Kong Baptist University, Hong Kong, China

ARTICLE INFO

Keywords:
Urban rail transit
Travel behavior
Natural experiment
Systematic review
Meta-analysis

ABSTRACT

Recent research has treated new rail transit systems as 'natural experiments', and while they have generally resulted in increased ridership, the underlying reasons for this remain unclear. To address this, we undertook a systematic review and meta-analysis of natural experiment studies published between 2000 and 2023 to synthesize the effect of rail transit on various travel behavior. We selected a total of sixteen studies for review, of which eight were suitable for meta-analysis. The pooled results showed that the introduction of rail transit significantly increased the mode share of rail while significantly decreasing the mode share of bus and car, and vehicle miles traveled (VMT). Our findings suggest that rail transit has the potential to promote sustainable travel behavior. However, to establish a more reliable association, further high-quality research is needed to examine the nuanced context, extend the follow-up duration, incorporate objective measures, and appropriately define the control group.

1. Introduction

The growing reliance on cars in urban areas has led to issues such as traffic congestion and air pollution. To address these challenges, many local and regional governments worldwide have made, or plan to make, substantial investments in urban rail transit systems, with the aim of reducing society's car dependence and promoting sustainable travel behavior (Loo et al., 2010). However, urban rail transit project is one of the most costly transportation investments in a city or metropolitan area, as they require significant and irreversible investments in long-lived assets that may impose a substantial financial burden on governments (Pulido et al., 2018). Therefore, in planning and designing urban rail transit systems, it is essential to consider their effects on individuals' travel behavior, urban mobility, and sustainability (Jeihani et al., 2013). Indeed, there is an ongoing debate on the effectiveness of investments in urban rail transit, which hinges on the degree to which rail transit promotes sustainable mobility and, in particular, converts car users to rail transit users (Ingvardson and Nielsen, 2018). As some researchers have argued, rail transit infrastructure can only benefit the

* Corresponding author at: Department of Architecture and Civil Engineering, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Tong, Hong Kong, China (Y. Lu).
E-mail addresses: jwang586-c@my.cityu.edu.hk (J. Wang), chwb817@gmail.com (W. Cheng), yilu24@cityu.edu.hk (Y. Lu), dgwang@hkbu.edu.hk (D. Wang).
[1] Jingjing Wang and Weibin Cheng are co-first authors of the article.

https://doi.org/10.1016/j.trd.2023.103882
Received 11 May 2023; Received in revised form 12 August 2023; Accepted 13 August 2023
Available online 20 August 2023
1361-9209/

environment and society if the increased train traffic comes from private vehicles rather than buses or non-motorized forms of transportation such as walking and cycling (Lund et al., 2006).

In response to these arguments, numerous empirical studies have examined the effects of urban rail transit systems on travel behavior. However, most studies are based on cross-sectional data, indicating only associations among variables, and it remains unclear whether rail transit development can lead to meaningful travel behavioral changes (Chatterjee and Carey, 2018). To make causal inferences, randomized controlled trials (RCTs) would be required to track the same set of individuals before and after the implementation of the rail transit system, and thus isolate its treatment effect (Leatherdale, 2019). However, RCTs are often infeasible in travel or urban planning research because researchers cannot manipulate transportation interventions and randomly assign observations to treatment or control groups. Natural experiments, which allow the exploitation of opportunities presented by naturally occurring interventions, can serve as an alternative to RCTs (Leatherdale, 2019).

Researchers have considered the launch of new rail lines and stations as interventions to conduct natural experiments and evaluate the treatment effects of such infrastructure on individuals' travel behavior (Dai et al., 2020; Engebretsen et al., 2017). While many studies have reported an increase in rail ridership after the rail transit intervention (Deng and Zhao, 2022; Ewing and Hamidi, 2014; Spears et al., 2017; Xie, 2016), there is no consensus on the source of this increase. Although the car-to-rail modal shift has been observed in previous evidence (Huang et al., 2019; Ibraeva et al., 2021), consistent with the expectations, some studies have indicated that newly built rail transit infrastructures may not necessarily reduce car use, and the increased rail usage could result from induced travel demand or substitution for other travel modes (Cervero, 2007; Chatman, 2013; Sun et al., 2020). Furthermore, the mixed results may stem from various intervention contexts, follow-up durations, travel behavior measures, and sampling and analysis methods, making it difficult to compare them directly. Therefore, it is necessary to synthesize and quantify the mixed evidence from natural experiment studies to clarify the effectiveness of rail transit interventions.

Previous reviews have explored the effects of various built environment interventions on travel behavior, but none have specifically focused on rail transit interventions (Ewing and Cervero, 2010; Wang and Zhou, 2017).To address this gap, this study aims to conduct a systematic review of natural experiment studies that have quantitatively estimated the effects of urban rail transit systems on travel behavior. Furthermore, this study aims to synthesize all available evidence through a meta-analysis to provide overall effect estimates of rail transit systems on travel behavior. The findings of this study have the potential to provide guidelines for future research and inform policymakers on rail transit investment decisions.

2. Methods

2.1. Systematic review and meta-analysis (SR/MA)

SR/MAs are important tools for synthesizing empirical evidence and informing policy-making in various fields (Mikolajewicz and Komarova, 2019; Tawfik et al., 2019). In this study, we intend to undertake a systematic review and meta-analysis of natural experiment studies that have quantitatively estimated the effects of urban rail transit systems on travel behavior.

A systematic review involves the logical synthesis and critical appraisal of multiple studies focused on a well-defined research topic (DeLuca et al., 2008), while a meta-analysis is a statistical estimation of the results from each individual study that generates a pooled estimate of the problem under study (Gopalakrishnan and Ganeshkumar, 2013). By consolidating evidence from a large body of complex studies, a meta-analysis can provide more trustworthy treatment effect estimates than individual studies by increasing the overall sample size. However, it is crucial to acknowledge that meta-analyses may average differences among homogeneous studies and may not be representative of any individual study, particularly if studies are conducted in different urban or cultural contexts (Feinstein, 1995; Maziarz, 2022). The SR/MA used in this study follows the widely accepted guidelines of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins et al., 2019) and *Preferred Reporting Items for Systematic Reviews and* Meta-*Analyses Statement* (Moher et al., 2009).

2.2. Search and selection procedure

The present study conducted comprehensive literature research using three electronic databases, namely Web of Science, Scopus, and Transportation Research Information Service, in January 2023. In addition, the reference lists of selected articles were combed through to locate any further relevant studies. The search terms used in the query were carefully selected and included, but were not limited to, the following: rail, metro, subway, travel, transit, commute, driving, walking, natural experiment, quasi-experiment, longitudinal, retrospective and prospective.

The eligibility criteria for selecting studies are as follows: (1) publication after the year 2000; (2) use of quasi-experiment or natural experiment study design, such as before-and-after and experimental-control study design with an urban rail transit intervention; (3) inclusion of a measure of travel behavior change as an outcome, such as travel mode share, trip frequency, trip duration, trip distance, and vehicle miles traveled (VMT) / vehicle kilometers traveled (VKT); and (4) written in English. Various types of urban rail transit, including light rail and heavy rail, were included to obtain comprehensive evidence.

To ensure that a meta-analysis could be conducted, two additional inclusion criteria were required for studies: (1) reporting pre- and post-travel behavioral changes with corresponding mean and standard deviations (SDs), or other values that can be used to calculate SDs, such as 95 % confidence intervals (CIs), standard errors (SEs), *p*-values, and *t*-values (Higgins et al., 2019), and (2) having identical travel behavior outcomes that can be pooled with the outcomes of other studies for effect size calculation (with a minimum of three instances of each travel behavior outcome).

Following the initial search and deduplication, two researchers (JW and XW) separately reviewed the titles and abstracts of all records, and removed any ineligible studies. Articles selected by one or both researchers were undergone to full-text assessment, which was also performed independently by both researchers. Any disagreements between researchers regarding eligibility for full-text assessment were settled through consultation with a third researcher (YL).

Fig. 1 shows the search process. In total, 4,365 studies were identified, among which 123 studies underwent full-text assessment. Studies were excluded if they were deemed unrelated to urban rail transit or travel behavior outcomes, or if they did not conduct a longitudinal study. During the full-text assessment, 107 of 123 articles were excluded. Of those excluded, 23 studies did not focus on travel behavior measures, 49 studies lacked rigor in research design by not including control groups, and 35 studies did not utilize statistical models to estimate the treatment effect of urban rail transit. Ultimately, the systematic review comprised 16 studies, with 8 of them being qualified for the meta-analysis.

2.3. Quality assessment

We use the Quality Assessment Tool for Quantitative Studies (Thomas et al., 2004) to evaluate the quality of the eight studies eligible for the meta-analysis. The tool is a universal assessment instrument created to evaluate various intervention studies, including RCTs and natural experiments. It has been considered suitable for systematic reviews that assess the efficacy of public health interventions (Armijo-Olivo et al., 2012). The tool comprised six domains: (1) selection bias (representative of participants), (2) study design (random allocation of treatment and control group), (3) confounders (control of confounders during the study design or analysis), (4) blinding (assessors and participants' awareness of the intervention status and research questions), (5) data collection methods (reliability and validity of outcome measures) and (6) withdrawals and dropouts (the extent to which participants withdraw or drop out of a study before completing the study). Each domain was assessed as weak, moderate, or strong according to the standardized guide and dictionary (Thomas et al., 2004). The study's overall rating was based on these six domains. Studies with two or more weak ratings were characterized as weak. Studies with less than four strong ratings and one weak rating were classified as moderate, while those with four or more strong ratings and no weak ratings were evaluated as strong. Two researchers (JW and WC) conducted the assessment. If there were any differences of opinion between the two researchers, they were resolved through consultation with a third researcher (YL).

2.4. Statistical analysis

Given the variation in the methods utilized to estimate the effects of the intervention, both in magnitude and direction, we applied the random-effects meta-analysis model tool in Review Manager 5.3 to determine the overall intervention effects (Higgins et al., 2019). A two-stage process employed calculates the overall effect size (Fig. 2): (1) The effective size of individual studies is computed as the

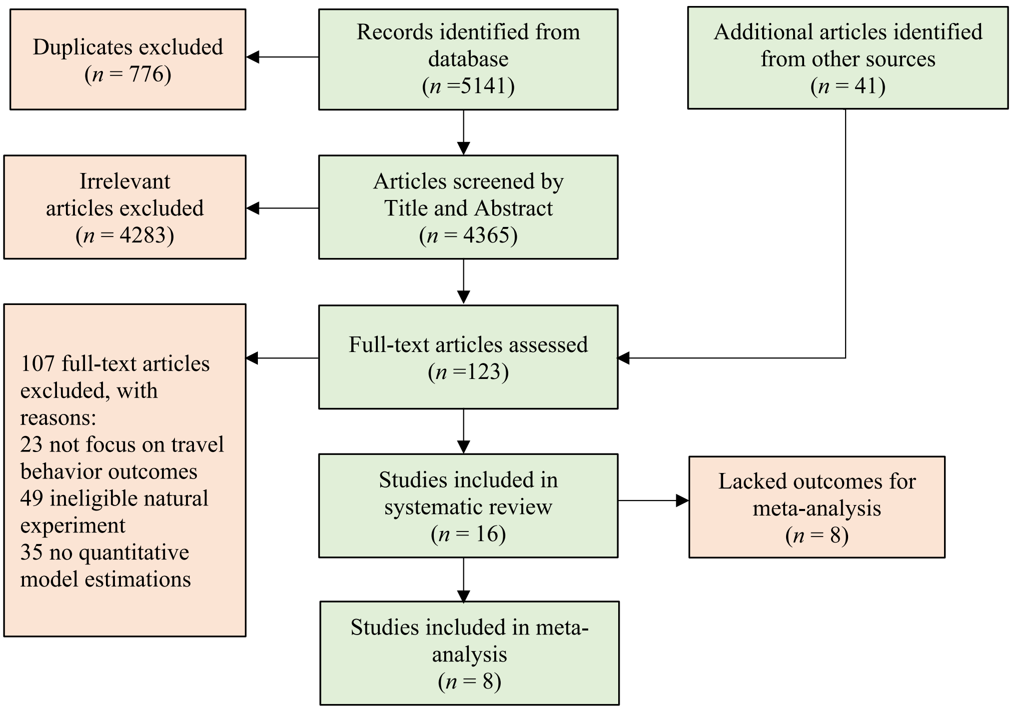

Fig. 1. Flowchart for the systematic review and meta-analysis search process.

mean difference, which is the difference between the treatment and control groups for the change of an outcome before and after the intervention. and (2) the pooled effect size across all studies is computed after weighting the effect sizes for precision, where the precision is proportional to the standard error of the mean (SEM). When the standard deviation (SD) is smaller and the sample size is larger, the SEM is reduced. A smaller SEM is indicative of a higher level of precision in estimating the mean (Andrade, 2020). The standard error of the pooled effect size is applied to calculate a confidence interval (CI) and a p-value indicating the precision of the pool estimate and the strength of the evidence respectively. Heterogeneity is assessed by the I^2 statistic, which represents the proportion of total variation in effect sizes that is due to heterogeneity, with values above 50 % indicating substantial heterogeneity (a detailed explanation of the calculation method can be found in Higgins et al. (2003)). Forest plots (Fig. 2) are a graphical representation of the results of a meta-analysis, providing a visual summary of the effect sizes and confidence intervals for each study, as well as an overall summary estimate of the effect size and its confidence interval.

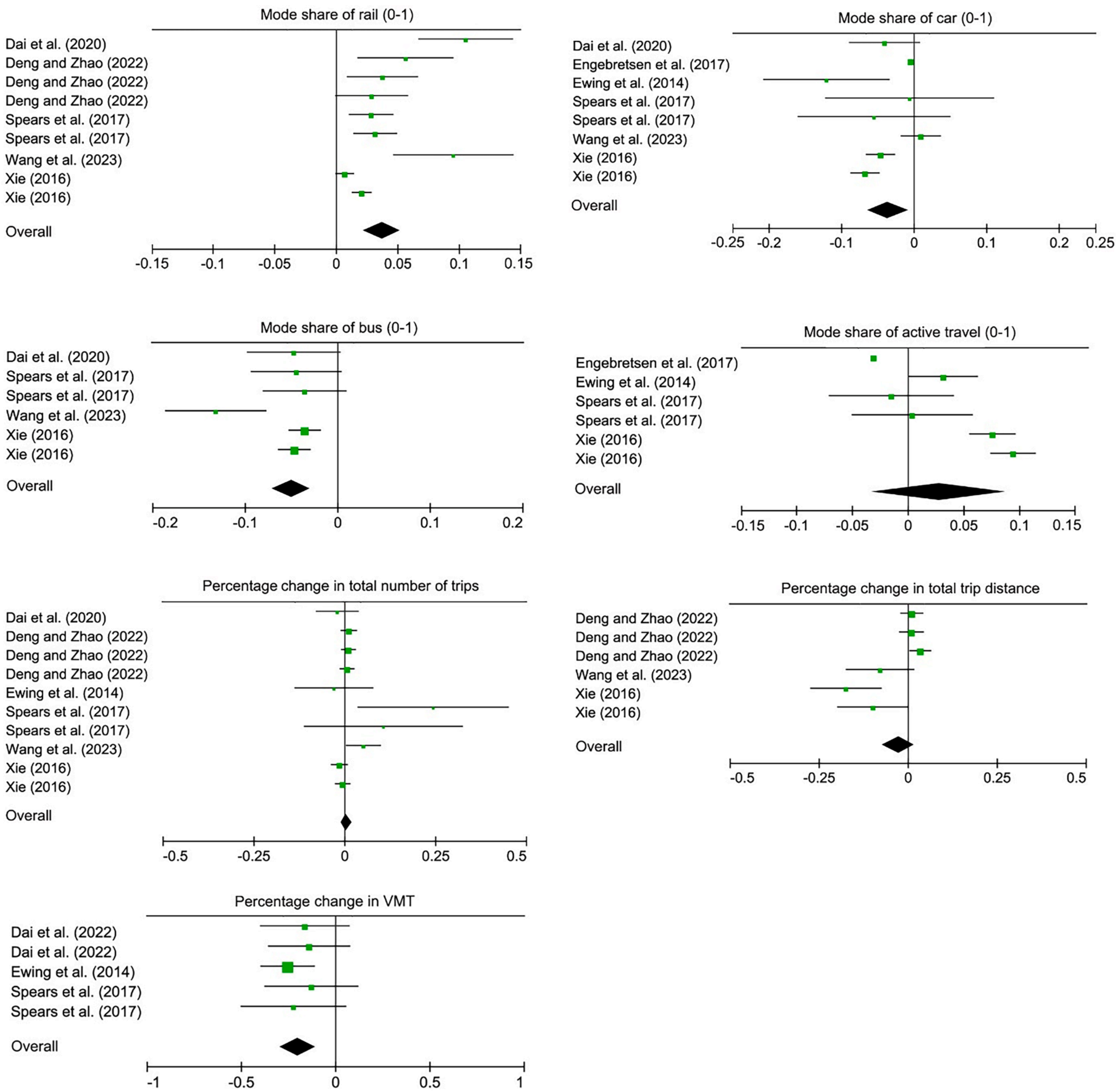

Fig. 2. Forest plots for seven types of travel behavior outcomes following urban rail transit interventions. Note: a. The square shape represents the effect size for each study. The center of each shape represents the point estimate (i.e., mean difference), and the horizontal lines extending from each shape represent the confidence interval for that estimate. The diamond shape at the bottom of the plot represents the pooled effect size across all studies. The center of the diamond represents the point estimate of the pooled effect size, and the width of the diamond represents the confidence interval. If a diamond does not cross the vertical line in the middle (i.e., a mean difference of zero), this suggests that the pooled estimate is statistically significant. b. Note: In order to make a standardized comparison across the studies, we calculated the effects in terms of percentage change for the total number of trips, the total trip distance and VMT variables.

3. Results

3.1. Systematic review

There were 16 studies eligible for inclusion in the systematic review (Appendix Table 1). Most of the studies were conducted in China (n = 6) and the United States (n = 5), with the remaining five studies conducted in Singapore (n = 2), Norway (n = 1), the United Kingdom (n = 1), and Portugal (n = 1). Two types of rail transit intervention were examined: light rail (n = 7) and heavy rail (n = 9). The studies employed three types of research design, namely prospective longitudinal study (n = 4), in which data were collected from a targeted group at baseline and followed over a period of time; retrospective longitudinal study (n = 2), in which data were collected from past records or surveys of a targeted group; and repeated cross-sectional study (n = 10), in which data were collected from different groups of individuals at multiple points of time. Most studies used travel surveys as their data source (n = 11), while others used official transport data (n = 3), mobile phone data (n = 1), and passenger counts (n = 1). The follow-up periods varied from 1 to 21 years, with the longitudinal studies (1 to 2.5 years) generally shorter follow-up durations than the repeated cross-sectional studies (mostly ≥ 4 years). The data analysis units were mainly individual person (n = 10), while household, community, city and other units were also used. The sample sizes at the personal or household level varied greatly, ranging from 285 to 21,859 at baseline and from 173 to 23,960 at follow-up, with repeated cross-sectional studies usually involving larger sample sizes (mostly > 1,000 participants) than longitudinal studies (mostly < 1,000 participants). For the prospective longitudinal studies, the retention rates in the follow-up surveys were generally above 50 %, except for one large-sample study (32.6 %) (Sun et al., 2020).

The definitions of the terms '*treatment*' and '*control*' groups were primarily based on the level of convenience in accessing the rail service. In most studies, a single distance threshold to a new rail line or rail station was used to differentiate the treatment and control groups (e.g., 500 m, 1000 m). Participants living within the threshold were regarded as the treatment group, while those living beyond the threshold were classified as the control group. Some studies set control corridors that shared similar attributes with rail transit corridors but lacked a rail transit service (Ewing and Cervero, 2010; Senior, 2009). The DiD method was most commonly used to estimate the intervention effects (n = 9), while some other statistical methods, such as the ordered response model, structural equations model, and propensity score matching were also employed.

The travel behavior metrics analyzed in the reviewed studies included mode share, trip frequency/ number of each travel mode (such as rail, car, bus, and active travel), VMT/VKT, total number of trips and total trip distance. Some studies also examined ridership and travel time (Liu and Li, 2020; Sun et al., 2020; Werner et al., 2016). The association between each travel behavior metric and the rail transit intervention was coded positive, negative, or non-significant, based on the findings of each study (Table 1).

Overall, the establishment of urban rail transit was found to have a positive association with increased usage of rail travel. Furthermore, most studies reported a negative association between urban rail transit and the usage of bus and car travel, while no significant associations were found between urban rail transit and the total number of trips. The results for active travel were mixed,

Table 1
Associations between rail transit and travel behavior outcomes of the studies in the systematic review.

Travel behavior variables	Positive	Negative	No significance
Total number/frequency of trips	2	0	3
Total trip distance	1	0	2
Total travel time of trips	0	0	1
Pooled of any total travel behavior index	3	0	6
Mode share of car	0	2	1
Number/frequency of car trips	0	2	1
Travel time of car trips	0	0	1
VMT/VKT	0	4	0
AADT	0	1	0
Pooled of any car travel index	0	9	3
Mode share of rail	3	0	0
Number/frequency of rail trips	4	0	0
Distance of rail trips	1	0	0
Travel time of rail trips	2	0	0
Pooled of any rail travel index	10	0	0
Mode share of bus	0	1	2
Number/frequency of bus trips	0	1	1
Travel time of bus trips	0	1	0
Bus ridership	0	1	0
Pooled of any bus travel index	0	4	3
Mode share of active travel	1	0	0
Number of walking trips	0	0	1
Number of cycling trips	0	0	1
Travel time of walking trips	0	1	0
Travel time of cycling trips	0	1	0
Pooled of any active travel index	1	2	2

Note: Public transit = rail and bus; active travel = walking and cycling; VMT = vehicle miles traveled; VKT = vehicle kilometers traveled; AADT = vehicular traffic within road segments for both directions on any given day during a year. Travel behavior variables were pooled by travel mode.

with some studies reporting a positive association and others reporting no significant association.

3.2. Research quality

Of the eight studies in the meta-analysis, six were rated as moderate (Dai et al., 2020; Dai et al., 2022; Deng and Zhao, 2022; Spears et al., 2017; Wang et al., 2023; Xie, 2016), while the remaining two were rated as weak (Engebretsen et al., 2017; Ewing and Hamidi, 2014) (Table 2). All the studies were rated as moderate in the dimensions of selection bias and study design, as they were observational studies with a non-random selection of participants. Some studies had issues in cofounders, as they did not control for relevant confounders (Deng and Zhao, 2022; Engebretsen et al., 2017; Ewing and Hamidi, 2014; Spears et al., 2017). One study was rated as strong in the dimension of data collection methods, as it used mobile phone data that provide objective measurements (Deng and Zhao, 2022). In six repeated cross-sectional studies (Dai et al., 2020; Dai et al., 2022; Engebretsen et al., 2017; Ewing and Hamidi, 2014; Wang et al., 2023; Xie, 2016), the dimension of withdrawals and dropouts was rated as weak, as the dropout was high.

3.3. Meta-Analysis

The meta-analysis included eight studies that reported travel behavior outcomes comparable to those of other studies (Dai et al., 2020; Dai et al., 2022; Deng and Zhao, 2022; Engebretsen et al., 2017; Ewing and Hamidi, 2014; Spears et al., 2017; Wang et al., 2023; Xie, 2016). Some studies provided multiple comparison pairs due to having multiple control groups (Dai et al., 2022; Deng and Zhao, 2022) or follow-up periods (Spears et al., 2017; Xie, 2016). We identified 50 comparison pairs from the studies and converted the outcomes into seven categories: mode shares of rail, bus, car, and active travel, total number of trips and total trip distance, and VMT, which were then pooled.

Fig. 2 and Table 3 present the overall effect of introducing new rail transportation on various travel behavior outcomes. The results indicate a significant positive relationship between rail transit and the mode share of rail (0.04, 95 % CI = 0.02, 0.05, $P < 0.001$), but there was an indication of heterogeneity ($I^2 = 83$ %, $P < 0.001$). However, significant negative associations were found between rail transit and the mode shares of bus and car, as well as VMT, (for bus, −0.05, 95 % CI = −0.07, −0.03, $P < 0.001$; for car, −0.04, 95 % CI = −0.06, −0.01, $P = 0.01$; for VMT, −0.20, 95 % CI = −0.29, −0.11, $P < 0.001$), with no heterogeneity for bus and VMT, but with heterogeneity for car (for bus, $I^2 = 54$ %, $P = 0.05$; for car, $I^2 = 89$ %, $P < 0.001$; for VMT, $I^2 = 0$ %, $P = 0.87$). The pooled effects on the mode share of active travel (0.03, 95 % CI = −0.03, 0.09, $P = 0.37$), the total number of trips (0.00, 95 % CI = −0.01, 0.01, $P = 0.72$) and the total trip distance (−0.05, 95 % CI = −0.11, 0.02, $P = 0.14$) were insignificant.

4. Discussion

4.1. Major findings

This study presents the first SR/MA based on natural experiment studies investigating the effects of newly introduced rail transit infrastructures on travel behavior. The pooled effect was estimated separately for various travel mode outcomes, revealing a pattern of modal shift in response to rail transit infrastructure. The introduction of new rail transit systems significantly increased the mode share of rail ridership while significantly decreasing the mode share of bus, car and vehicle miles traveled (VMT). However, it has no significant effect on the mode share of active travel, the total number of trips and the total trip distance. The findings of the SR/MA suggest that the transition from bus and car might be the reason for the increase in rail ridership in response to the intervention.

Promoting public transit use as a substitute for car use is a crucial expected outcome of urban rail investments. Prior studies have established that individuals residing close to rail stations tend to use their personal vehicles less frequently, perhaps due to the availability of a convenient transit option, particularly in areas characterized by high levels of traffic congestion (Ewing and Hamidi, 2014; Huang et al., 2019). For individuals located in or around urban centers, using rail transit is generally more appealing than driving, thereby increasing their likelihood of switching to transit travel after its introduction (Paulley et al., 2006). Nevertheless, travel demand and personal preferences are likely to differ among residents. Discretionary riders who have greater flexibility in choosing their travel times may prefer the convenience and flexibility of driving (Jeihani et al., 2013). Similarly, higher-income males

Table 2
Quality rating of the studies in the meta-analysis.

Study	Selection bias	Study design	Cofounders	Blinding	Data collection methods	Withdrawals and dropouts	Overall rating
Ewing and Hamidi (2014)	Moderate	Moderate	Weak	Moderate	Moderate	Weak	Weak
Xie (2016)	Moderate	Moderate	Moderate	Moderate	Moderate	Weak	Moderate
Spears et al. (2017)	Moderate	Moderate	Weak	Moderate	Moderate	Moderate	Moderate
Engebretsen et al. (2017)	Moderate	Moderate	Weak	Moderate	Moderate	Weak	Weak
Dai et al. (2020)	Moderate	Moderate	Moderate	Moderate	Moderate	Weak	Moderate
Deng and Zhao (2022)	Moderate	Moderate	Weak	Moderate	Strong	Strong	Moderate
Dai et al. (2022)	Moderate	Moderate	Moderate	Moderate	Moderate	Weak	Moderate
Wang et al. (2023)	Moderate	Moderate	Moderate	Moderate	Moderate	Weak	Moderate

Table 3
Meta-analysis results.

Mode share of rail (0–1)

Study	Treatment Before-after MD	SD	Total	Control Before-after MD	SD	Total	Weight	Intergroup difference [95 % CI]
Dai et al. (2020)	0.146	0.339	561	0.041	0.317	561	7.8 %	0.10 [0.07, 0.14]
Deng and Zhao (2022)	0.075	0.595	1801	0.019	0.578	1676	7.7 %	0.06 [0.02, 0.10]
Deng and Zhao (2022)	0.075	0.595	1801	0.037	0.183	1682	10.2 %	0.04 [0.01, 0.07]
Deng and Zhao (2022)	0.075	0.595	1801	0.046	0.289	2640	9.9 %	0.03 [−0.00, 0.06]
Spears et al. (2017)	0.033	0.070	138	0.005	0.068	92	13.3 %	0.03 [0.01, 0.05]
Spears et al. (2017)	0.030	0.064	150	−0.002	0.073	97	13.4 %	0.03 [0.01, 0.05]
Wang et al. (2023)	0.271	0.418	524	0.176	0.388	524	5.9 %	0.10 [0.05, 0.14]
Xie (2016)	0.004	0.147	3342	−0.002	0.190	4205	16.0 %	0.01 [−0.00, 0.01]
Xie (2016)	0.020	0.166	3342	−0.001	0.189	4205	15.9 %	0.02 [0.01, 0.03]
Total [95 % CI]			13,460			15,682	100 %	0.04 [0.02, 0.05]
Heterogeneity: $I^2 = 83$ % ($P < 0.001$)								
Overall effect: $P < 0.001$								

Mode share of bus (0–1)

Study	Treatment Before-after MD	SD	Total	Control Before-after MD	SD	Total	Weight	Intergroup difference [95 % CI]
Dai et al. (2020)	−0.027	0.437	561	0.021	0.429	561	10.7 %	−0.05 [−0.10, 0.00]
Spears et al. (2017)	−0.017	0.192	150	0.028	0.193	97	11.2 %	−0.04 [−0.09, 0.00]
Spears et al. (2017)	−0.030	0.181	137	0.006	0.164	92	12.5 %	−0.04 [−0.08, 0.01]
Wang et al. (2023)	−0.061	0.456	524	0.071	0.448	524	9.6 %	−0.13 [−0.19, −0.08]
Xie (2016)	−0.056	0.389	3342	−0.020	0.376	4205	28.1 %	−0.04 [−0.05, −0.02]
Xie (2016)	−0.031	0.390	3342	0.016	0.384	4205	27.9 %	−0.05 [−0.06, −0.03]
Total [95 % CI]			8056			9684	100 %	−0.05 [−0.07, −0.03]
Heterogeneity: $I^2 = 54$ % ($P = 0.05$)								
Overall effect: $P < 0.001$								

Mode share of car (0–1)

Study	Treatment Before-after MD	SD	Total	Control Before-after MD	SD	Total	Weight	Intergroup difference [95 % CI]
Dai et al. (2020)	−0.070	0.407	561	−0.029	0.432	561	12.3 %	−0.04 [−0.09, 0.01]
Engebretsen et al. (2017)	−0.051	0.006	7649	−0.046	0.004	15,249	19.4 %	−0.00 [−0.01, −0.00]
(Ewing and Hamidi, 2014)	−0.141	0.608	348	−0.020	0.652	465	6.8 %	−0.12 [−0.21, −0.03]
Spears et al. (2017)	0.023	0.458	136	0.029	0.426	92	4.5 %	−0.01 [−0.12, 0.11]
Spears et al. (2017)	−0.024	0.415	149	0.031	0.408	96	5.3 %	−0.06 [−0.16, 0.05]
Wang et al. (2023)	−0.003	0.221	524	−0.012	0.236	524	16.4 %	0.01 [−0.02, 0.04]
Xie (2016)	−0.004	0.449	3342	0.043	0.440	4205	17.7 %	−0.05 [−0.07, −0.03]
Xie (2016)	−0.020	0.446	3342	0.048	0.442	4205	17.7 %	−0.07 [−0.09, −0.05]
Total [95 % CI]			16,051			25,397	100 %	−0.04 [−0.06, −0.01]
Heterogeneity: $I^2 = 89$ % ($P < 0.001$)								
Overall effect: $P = 0.01$								

Mode share of active travel (0–1)

Study	Treatment Before-after MD	SD	Total	Control Before-after MD	SD	Total	Weight	Intergroup difference [95 % CI]
Engebretsen et al. (2017)	−0.001	0.005	7649	0.030	0.003	15,249	17.7 %	−0.03 [−0.03, −0.03]
(Ewing and Hamidi, 2014)	0.055	0.221	348	0.024	0.227	465	16.9 %	0.03 [0.00, 0.06]
Spears et al. (2017)	0.000	0.221	137	0.015	0.205	91	15.3 %	−0.02 [−0.07, 0.04]
Spears et al. (2017)	0.013	0.212	149	0.009	0.213	97	15.4 %	0.00 [−0.05, 0.06]
Xie (2016)	0.056	0.458	3342	−0.020	0.458	4205	17.3 %	0.08 [0.05, 0.10]
Xie (2016)	0.031	0.449	3342	−0.063	0.449	4205	17.3 %	0.09 [0.07, 0.11]
Total [95 % CI]			14,967			24,312	100 %	0.03 [−0.03, 0.09]
Heterogeneity: $I^2 = 98$ % ($P < 0.001$)								
Overall effect: $P = 0.37$								

Percentage change in total number of trips

Study	Treatment Before-after MD	SD	Total	Control Before-after MD	SD	Total	Weight	Intergroup difference [95 % CI]
Dai et al. (2020)	−0.036	0.476	561	−0.015	0.531	561	3.9 %	−0.02 [−0.08, 0.04]
Deng and Zhao (2022)	0.004	0.345	1801	−0.007	0.331	1676	14.2 %	0.01 [−0.01, 0.03]
Deng and Zhao (2022)	0.004	0.345	1801	−0.006	0.272	1682	15.3 %	0.01 [−0.01, 0.03]
Deng and Zhao (2022)	0.004	0.345	1801	−0.002	0.331	2640	15.5 %	0.01 [−0.01, 0.03]
(Ewing and Hamidi, 2014)	0.135	0.752	348	0.165	0.815	465	1.3 %	−0.03 [−0.14, 0.08]

(continued on next page)

Table 3 (*continued*)

Percentage change in total number of trips

Study	Treatment			Control			Weight	Intergroup difference [95 % CI]
	Before-after MD	SD	Total	Before-after MD	SD	Total		
Spears et al. (2017)	0.088	0.877	149	−0.155	0.763	96	0.4 %	0.24 [0.04, 0.45]
Spears et al. (2017)	0.043	0.889	135	−0.063	0.774	91	0.3 %	0.11 [−0.11, 0.32]
Wang et al. (2023)	−0.212	0.384	524	−0.263	0.406	524	5.5 %	0.05 [0.00, 0.10]
Xie (2016)	−0.006	0.245	3342	0.009	0.25	4205	21.5 %	−0.01 [−0.03, −0.00]
Xie (2016)	−0.025	0.236	3342	−0.019	0.209	4205	22.2 %	−0.01 [−0.02, 0.00]
Total [95 % CI]			13,804			16,145	100 %	0.00 [−0.01, 0.01]
Heterogeneity: $I^2 = 56$ % (P = 0.02)								
Overall effect: P = 0.72								

Percentage change in total trip distance

Study	Treatment			Control			Weight	Intergroup difference [95 % CI]
	Before-after MD	SD	Total	Before-after MD	SD	Total		
Deng and Zhao (2022)	0.151	0.466	1801	0.141	0.631	2640	17.8 %	0.01 [−0.02, 0.04]
Deng and Zhao (2022)	0.151	0.466	1801	0.142	0.569	1682	17.7 %	0.01 [−0.03, 0.04]
Deng and Zhao (2022)	0.151	0.466	1801	0.117	0.449	1676	17.9 %	0.03 [0.00, 0.06]
Wang et al. (2023)	0.101	0.791	524	0.18	0.800	524	13.0 %	−0.08 [−0.18, 0.02]
Xie (2016)	−0.100	0.995	3342	0.075	1.156	4205	16.8 %	−0.17 [−0.22, −0.13]
Xie (2016)	−0.095	1.006	3342	0.0040	1.148	4205	16.8 %	−0.10 [−0.15, −0.05]
Total [95 % CI]			12,611			14,932	100 %	−0.05 [−0.11, 0.02]
Heterogeneity: $I^2 = 93$ % (P < 0.001)								
Overall effect: P = 0.14								

Percentage change in VMT

Study	Treatment			Control			Weight	Intergroup difference [95 % CI]
	Before-after MD	SD	Total	Before-after MD	SD	Total		
Dai et al. (2022)	−0.397	2.771	801	−0.233	2.938	1684	15.3 %	−0.16 [−0.40, 0.07]
Dai et al. (2022)	−0.397	2.771	801	−0.256	2.875	2972	18.2 %	−0.14 [−0.36, 0.08]
Ewing and Hamidi (2014)	0.069	1.014	348	0.324	1.07	465	41.5 %	−0.26 [−0.40, −0.11]
Spears et al. (2017)	−0.078	1.032	146	0.051	0.907	93	13.9 %	−0.13 [−0.38, 0.12]
Spears et al. (2017)	−0.04	1.125	134	0.184	0.989	89	11.0 %	−0.22 [−0.50, 0.06]
Total [95 % CI]			2230			5303	100 %	−0.20 [−0.29, −0.11]
Heterogeneity: $I^2 = 0$ % (P = 0.87)								
Overall effect: P < 0.001								

Note: In order to make a standardized comparison across the studies, we calculated the effects in terms of percentage change for the total number of trips, the total trip distance and VMT variables.

are more likely to rely on personal vehicles as their primary mode of transportation, regardless of the improvement in transit service (Pan et al., 2013).

The shift from bus to rail can be explained by the fact that rail transit typically offers faster, more comfortable and convenient service compare to bus services, as noted in previous studies (Henry and Litman, 2006). Besides, the adjustment of bus service may be another reason for the shift from bus to rail in travel mode. After the implementation of rail transit, the bus transit agency may reduce the frequency or even eliminate certain bus lines that overlap with the rail transit routes, which could lead to a decrease in the overall supply of bus service. This reduction in bus service could further discourage the use of buses, leading to a decline in bus ridership. In cities with integrated public transport systems, rail transit often serve as a backbone of public transport service, with buses providing a complementary service due to their different capacities and inherent characteristics (Verma and Dhingra, 2006). For example, in Hong Kong, rail and bus account for 30 % and 27 %, respectively, of the daily trips made by urban residents during weekdays (Transport Department of Hong Kong, 2011).

The findings also indicate that the effect of rail transit interventions on active travel was insignificant. The introduction of urban rail transit systems may lead to an increase in active travel as households residing near rail stations may find it more convenient to use nearby facilities by walking or cycling (Zhao and Li, 2018). However, the time taken to travel to previous active travel destinations may decrease due to the introduction of rail transit, resulting in some active travelers shifting to using rail as their primary mode of transportation (Hong et al., 2016). Additionally, poor connectivity of sidewalks or cycling lanes, lack of tree shade, and concerns over traffic safety may discourage active travel (Sun et al., 2020). Hence, residents may choose to give up walking or cycling when other travel options are available. It should be noted that walking and cycling trips to and from rail stations may be excluded in some travel surveys, as they only captured the predominant mode of transportation within a trip chain. Therefore, the effect of rail transit interventions on active travel behavior may be underestimated.

Furthermore, it is worth mentioning that the lack of significant effects on the total trips may be explained by the fact that rail transit systems are typically implemented in areas with high demand for travel, and therefore the new rail services may just substitute or

Table A1
Characteristics of included studies for systematic review.

Author, year	Study location	Intervention type	Study design	Data source	Follow-up duration (year)	Data analysis unit	Sample size (N)		Treatment group	Control group	Travel behavior metric (associations with rail transit)	Statistical method
							Time 1	Time 2				
Senior (2009)	Greater Manchester, UK	Light rail	Longitudinal (prospective)	Survey	2.5	Individual	1002	614	Rail corridor	Heavy rail and non-rail corridors	Frequency of bus use (-) Frequency of rail use (+)	Chi-square test
(Ewing and Hamidi, 2014)	Portland, US	Light rail	Repeated cross-sectional	Survey	17	Household	634	991	Rail corridor	Highway corridor	VMT (-)	Linear regression model
Xie (2016)	Beijing, China	Heavy rail	Repeated cross-sectional	Survey	1 and 2	Individual	7547	7547	TAZs with smaller distance to stations	TAZs with unchanged distance to stations	Mode share of rail (+) Mode share of bus (0) Mode share of car (-) Mode share of active travel (+) Total number of trips (0)Total trip distance (0)	Difference-in-differences (DiD) model
Spears et al. (2017)	Los Angeles, US	Light rail	Longitudinal (prospective)	Survey	0.5 and 1.5	Household	285	173	Within 1000 m of the stations	Between 1000 m and 5000 m of the stations	VMT (-) Number of rail trips (+) Number of bus trips (0) Number of car trips (0) Number of walking trips (0) Number of cycling trips (0) Total number of trips (0)	DiD
Engebretsen et al. (2017)	Bergen, Norway	Light rail	Repeated cross-sectional	Survey	5	Individual	21,859	23,960	Within 1000 m of the stations	1000 m away from the stations	Mode share of public transit (+)	Logistic regression model
Cao and Ermagun (2016)	Minneapolis, US	Light rail	Longitudinal (retrospective)	Survey	–	Individual	597		Movers into rail corridors	Movers into non-rail urban and suburban corridors	Frequency of car use (-) Frequency of rail use (+)	Structural equations model (SEM)
Werner et al. (2016)	Salt Lake City, US	Light rail	Repeated cross-sectional	Passenger counts	1	Catchment area	–	–	Within 0.25 miles of the stations	0.25 miles away from the stations	Public transit ridership (+)	Fixed effects repeated measures regression model
Huang et al. (2019)	Xi'an, China	Heavy rail	Longitudinal (retrospective)	Survey	–	Individual	593		Movers into rail corridors	Movers into control corridors	VKT (-)	Ordered response model

(*continued on next page*)

Table A1 (*continued*)

Author, year	Study location	Intervention type		Study design	Data source	Follow-up duration (year)	Data analysis unit	Sample size (N) Time 1	Time 2	Treatment group	Control group	Travel behavior metric (associations with rail transit)	Statistical method
Dai et al. (2020)	Singapore	Heavy rail	Repeated cross-sectional		survey	4	Individual	1122	1122	Within 500 m of the stations	Between 500 m and 1000 m of the stations	Mode share of rail (+) Mode share of bus (0) Mode share of car (-) Total number of trips (0)	Two-dimensional propensity score matching (2DPSM) and DiD
Liu and Li (2020)	43 cities, China	Heavy rail	Repeated cross-sectional		Official transport data	21	City	–	–	Cities with subways	Cities without subways (with subway plan)	Bus ridership (-)	DiD
Sun et al. (2020)	Nanchang, China	Heavy rail	Longitudinal (prospective)		Survey	1	Individual	5436	1770	Within 800 m of the stations	1600 m and 5000 m away from the stations	Travel time of rail trips (+) Travel time of bus trips (-) Travel time of car trips (0) Travel time of walking trips (-) Travel time of cycling trips (-)	DiD
Ibraeva et al. (2021)	Porto, Portugal	Heavy rail	Repeated cross-sectional		Official transport data	10	Parish	–	–	Metro-served parishes	Non-metro-served parishes	Number of car trips (-)	DiD and Spatial DID (SDID) model
Tao et al. (2021)	Twin Cities, US	Light rail	Repeated cross-sectional		Official transport data	9	Road segment	–	–	One-mile buffer along the rail line	One-mile buffers along the highways	AADT (-)	DiD
Deng and Zhao (2022)	Shenzhen, China	Heavy rail	Longitudinal (prospective)		Mobile phone data	1 and 1.3	Individual	7799	7799	Between 0 and 1 km, 1–2 km and 2–3 km of the stations	3 km away from the stations	Total frequency of trips (+) Frequency of rail trips (+) Total trip distance (+) Distance of rail trips (+) Total travel time of trips (0)Travel time of rail trips (+)	DiD
Dai et al. (2022)	Singapore	Heavy rail	Repeated cross-sectional	survey	4	Individual	1444	1444	Within 500 m of the stations	Between 500 and 1000 m and 500–1500 m of the stations	VKT (-)	2DPSM and DiD	
Wang et al. (2023)	Hong Kong, China	Heavy rail	Repeated cross-sectional	Survey	9	Individual	1048	1048	Within 500 m of the stations	Between 500 and 1500 m of the stations	Mode share of rail (+) Mode share of bus (-)	2DPSM and paired *t*-test	

(*continued on next page*)

Table A1 (*continued*)

Author, year	Study location	Intervention type	Study design	Data source	Follow-up duration (year)	Data analysis unit	Sample size (N)		Treatment group	Control group	Travel behavior metric (associations with rail transit)	Statistical method
							Time 1	Time 2				
										Mode share of car (0) Total number of trips (+) Total trip distance (0)		

Note: Public transit = rail and bus; active travel = walking and cycling; TAZ = traffic analysis zone; VMT = vehicle miles traveled; VKT = vehicle kilometers traveled; AADT = vehicular traffic within road segments for both directions on any given day during a year. “+” = statistically significant positive association; “-” = statistically significant negative association; “0” = no significant association.
.

redistributed the existing travel demand (Ettema and Nieuwenhuis, 2017). Moreover, the effectiveness of rail transit in inducing travel demand may also depend on the existing transportation infrastructure, land use, and socioeconomic factors in the study area (Chatman, 2013). Therefore, the lack of significant effects on total trips should not be interpreted as a failure of rail transit interventions in promoting sustainable urban mobility, but rather as a need for more comprehensive and context-specific planning strategies to achieve the desired outcomes of transit-oriented development (TOD).

Although the intervention effects on the travel modes were pooled and quantified in our meta-analysis, the results should be treated with caution as there was some evidence of heterogeneity. When interpreting the heterogeneity, it is important to consider both the contexts of rail interventions on the results. The contexts of rail infrastructural interventions remain an important consideration when examining how the effect of an intervention varies in different settings (Campbell et al., 2007). Urban rail infrastructure is designed to suit cities' specific social and economic characteristics and their population's travel demand, which may all moderate its effects (Moore et al., 2014). Furthermore, the effectiveness of the interventions may differ with the built environment around rail stations and the connectivity to existing public transport systems, public facilities and parks of the city (Chatman, 2013; Xiao et al., 2019). Another factor that may explain the variance in the analyzed studies is the diversity of their methodologies, such as the types of travel behavior measured, measurement tools, length of follow-up duration, definition of treatment and control groups and the statistical analysis method.

4.2. Data sources

Existing research in the field of travel behavior analysis primarily relies on survey data, which includes both repeated cross-sectional data and panel data. Repeated cross-sectional datasets captured different cohorts of participants in different waves of survey and thus can only provide insight into the aggregate population-level implications of changes in travel behavior (Zhong et al., 2021). However, different cohorts of participants could lead to inaccurate estimates, as any changes in travel patterns observed could be due to changes in demographic factors over time (Cao and Cao, 2014). Nevertheless, repeated cross-sectional data are particularly useful for before-and-after studies of travel behavior because of their widespread availability and ease of access. In addition, the issue of demographic change over time in cross-sectional data can be partially mitigated by using statistical techniques such as PSM (Dai et al., 2020; Dai et al., 2022).

Panel data is gathered for a particular group or population at two different points in time – before and after the implementation of rail transit infrastructure, which can detect changes in the outcomes of the same participants and thus provide more robust evidence for inferring a causal effect (Spears et al., 2017). However, panel datasets tend to have relatively small sample sizes because tracking participants for years is a time-consuming and costly endeavor (Zhong et al., 2021). Due to the limited sample size and potential dropouts of participants during the tracking process, as well as self-reported bias, the results may remain biased.

Moreover, individual travel behavior is highly complex, with both predictable and unpredictable patterns occurring on a daily, weekly, and monthly basis (Järv et al., 2014). Conventional survey data can only capture a brief moment in time of an individual's travel habits, typically no longer than a single day or a few days. To address these limitations, researchers have started using big data, such as mobile phone data, to measure individuals' travel behavior (Deng and Zhao, 2022). Mobile phone data is gathered automatically and continuously, which permits larger population samples and the ability to record continuous travel information over an extended duration of time. However, mobile phone data cannot accurately provide sociodemographic attributes or identify transformation between different travel modes, such as transfer between bus and rail, compared to traditional survey data.

4.3. Statistical methods

In terms of statistical methods, DiD method is the most well-established and widely used in natural experiment studies. The DiD method measures the difference in outcome changes between those who have been exposed to an intervention (such as the implementation of a new infrastructure or policy) and those who have remained unexposed (Craig et al., 2017). The method relies on the assumption that, without the intervention, changes in outcomes would have occurred equally in both groups (Dimick and Ryan, 2014). Therefore, any differences between these two groups can be traced to the intervention's effect. Compared to traditional regression models, the DiD approach has the advantage of controlling for both unobserved and observed differences in the groups' fixed characteristics, making it less susceptible to bias from unmeasured confounders or measurement errors (Craig et al., 2017). Recent developments, such as the PSM method, used to match exposed and unexposed groups, can help control for unmeasured confounders and allow balanced comparisons (Austin, 2008). Therefore, the combination use of DiD and PSM may further strengthen causal inference.

In addition to DiD, the Interrupted Time Series (ITS) design with a control group, also known as controlled ITS, is another valid option for evaluating the impact of rail interventions (Lopez et al., 2019). While the DiD design measures the outcome at a single baseline time point and a single follow-up time point, it can be difficult to confirm the underlying assumption that the pre-intervention trends are parallel in both the control and intervention groups. This can lead to bias due to the unparallel trends in the two groups. The ITS design, on the other hand, measures outcomes at multiple baseline and follow-up time points and allows for extrapolation of pre-intervention trends in the two groups. This allows for the bias of DiD to be addressed by verifying whether the trends are parallel in the two groups (Lopez et al., 2019). Furthermore, ITS design allows for the identification of the post-intervention trend for the intervention group in detail, such as a time delay between the intervention and its subsequent impact, and the direction of effect trends over time, whether downward, flat, or upwards trends. Therefore, a controlled ITS design, which combines the strengths of both ITS and DiD, can further enhance the ability to infer causality (Craig et al., 2017; Lopez et al., 2019).

4.4. Limitations

A major limitation of our study is the small sample size in the meta-analysis. Despite our efforts to obtain original, unreported data from the authors of all publications included in this meta-analysis, the number of articles that met our inclusion criteria was limited. A small sample size may lead to low statistical power (Gurevitch et al., 2018). Second, there was significant heterogeneity for some travel behavior evidence, which suggested that their pooled effect sizes may not represent the common population (Ng et al., 2006). Although we discussed potential sources of heterogeneity, such as the difference in the study context, the definition of the catchment area, and the length of the follow-up period, we were unable to perform subgroup analyses to identify the sources of heterogeneity due to the limited eligible studies. Third, there remains the risk of publication bias, where studies with null or negative results are less likely to be published. This can lead to an overrepresentation of studies with positive results in the meta-analysis, potentially leading to an overestimation of the true effect size of the intervention (Thornton and Lee, 2000).

Despite such limitations, this study is timely and meaningful for several reasons. First, the natural experiments included in this review represent a diverse range of geographic locations, study populations, and study designs. This diversity increases the generalizability of our findings and suggests that our results may be applicable to a broad range of settings. Second, we conducted a rigorous assessment of study quality and excluded studies that did not meet our inclusion criteria, which helps to ensure that the studies included in our analysis were of sufficient quality. Such high standard ensures rigorous findings of the relationship between urban rail transit and travel behavior. Third, our analysis provides a comprehensive synthesis of the existing literature in this field, which may be particularly useful for policymakers who need to make informed decisions with high-quality and reliable evidence on a particular topic (Valentine et al., 2010). Finally, we believe that our study provides a valuable guide for future natural experimental study on the effect of rail transit on travel behavior.

4.5. Recommendations for future studies

The context of rail infrastructural interventions is crucial in determining the treatment effect o urban rail transit systems on travel behavior. However, most studies did not account for the complex context of each rail project studied, such as TOD implementation, the site and route selection of rail transit, and other municipal land use policies. This limitation impedes our understanding of the impacts of rail transit investments and municipal land use policies. To overcome this limitation, future studies should provide a detailed description of local contexts, which will allow their findings to be generalized to other locations. Additionally, it is important to distinguish the impacts of the built environment of station areas from the effects of rail transit itself. In other words, if there is an impact, planners should identify whether it can be attributed to transportation investments, land use policies, or both.

We offer several recommendations for improving research methods. First, we suggest that longer follow-up durations should be used, particularly for prospective longitudinal studies, where follow-up periods of 1 to 2.5 years may be insufficient for new travel habits to form. By tracking cohorts over a longer period (e.g., 5 years or more), we can obtain time series data for Controlled ITS analysis, which allows for stronger causal inference than simple pre-post DiD design and facilitates the detection of long-term travel behavioral changes.

Second, most studies relied on self-reported data from travel surveys, which are subject to potential recall bias and social desirability biases. Combining subjective (e.g., travel survey) and objective (e.g., mobile phone signaling data, wearable devices) measurement methods can accurately track changes in individuals' travel behavior (Song et al., 2020).

Third, some studies did not define control groups clearly, which made it difficult to isolate the treatment effect from temporal confounding factors. To address this issue, we suggest that future studies should define the treatment and control groups appropriately, matching built environment attributes (e.g., housing types, land use mixture, and density of public facilities) and demographic profiles closely. Moreover, almost all studies used only a single distance threshold to define treatment and control groups, making it difficult to identify the dose–response effect of rail transit interventions (Xie et al., 2021). To better understand the dose–effect, future studies should use multiple distance thresholds (e.g., 500, 1000, 1500 and 2000 m) to examine the effects of rail transit interventions.

5. Conclusions

Our study represents the first SR/MA based on natural experiment evidence to examine the effects of rail transit infrastructures on travel behavior. Using a random-effects meta-analysis model, we estimated the effects of rail transit on various travel modes and identified a pattern of modal shift. Our results indicate that rail ridership increased following rail transportation interventions, while the mode shares of bus and car declined. Further studies should make an effort to provide detailed descriptions of the local context, use longer follow-up durations, measure outcomes objectively, and define treatment and control groups appropriately.

Funding

The work described in this paper was fully supported by the Research Grants Council of the Hong Kong SAR (Project No. CityU11207520).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors wish to acknowledge data and other assistance from the following researchers: António Pais Antunes, Jason Cao, Robert Cervero, Carey Curtis, Fangzhou Dai, Yiling Deng, Mi Diao, Øystein Engebretsen, Xiaoyan Huang, Anna Ibraeva, Doina Olaru, Tao Tao, Di Wei, Xueying Wu.

Appendix A

See Table A1.

References

Andrade, C., 2020. Mean difference, standardized mean difference (SMD), and their use in meta-analysis: as simple as it gets. J. Clin. Psychiatry 81 (5), 11349.

Armijo-Olivo, S., Stiles, C.R., Hagen, N.A., Biondo, P.D., Cummings, G.G., 2012. Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. J. Eval. Clin. Pract. 18 (1), 12–18. https://doi.org/10.1111/j.1365-2753.2010.01516.x.

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.

Campbell, N.C., Murray, E., Darbyshire, J., Emery, J., Farmer, A., Griffiths, F., Guthrie, B., Lester, H., Wilson, P., Kinmonth, A.L., 2007. Designing and evaluating complex interventions to improve health care. Br. Med. J. 334 (7591), 455–459. https://doi.org/10.1136/bmj.39108.379965.be.

Cao, J., Cao, X., 2014. The Impacts of LRT, Neighbourhood Characteristics, and Self-selection on Auto Ownership: Evidence from Minneapolis-St. Paul. *Urban Studi.* 51 (10), 2068–2087. https://doi.org/10.1177/0042098013505887.

Cao, J., Ermagun, A., 2016. Influences of LRT on travel behaviour: A retrospective study on movers in Minneapolis. Urban Stud. 54 (11), 2504–2520. https://doi.org/10.1177/0042098016651569.

Cervero, R., 2007. Transit-Oriented Development's Ridership Bonus: A Product of Self-Selection and Public Policies. Environ. Plann. A: Econ. Space 39 (9), 2068–2085. https://doi.org/10.1068/a38377.

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

Chatterjee K., Carey R. N. (2018). Special Issue on the theory, design and evaluation of behaviour change interventions in transport. *Journal of Transport & Health, 10*, 7-10. doi:https://doi.org/10.1016/j.jth.2018.08.009.

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.

Dai, F., Diao, M., Sing, T.F., 2020. Effects of rail transit on individual travel mode shares: A two-dimensional propensity score matching approach. Transport. Res. Part D-Transp. Environ. 89 https://doi.org/10.1016/j.trd.2020.102601.

Dai, F., Diao, M., Sing, T.F., 2022. A two-dimensional propensity score matching approach to estimating the treatment effect of urban rail transit lines on vehicle travel. Transportation. https://doi.org/10.1007/s11116-022-10360-2.

DeLuca, J.B., Mullins, M.M., Lyles, C.M., Crepaz, N., Kay, L., Thadiparthi, S., 2008. Developing a Comprehensive Search Strategy for Evidence Based Systematic Reviews. Evid. Based Libr. Inf. Pract. 3 (1), 3–32. https://doi.org/10.18438/B8KP66.

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.

Dimick, J.B., Ryan, A.M., 2014. Methods for evaluating changes in health care policy: the difference-in-differences approach. JAMA 312 (22), 2401–2402.

Engebretsen, Ø., Christiansen, P., Strand, A., 2017. Bergen light rail – Effects on travel behaviour. J. Transp. Geogr. 62, 111–121. https://doi.org/10.1016/j.jtrangeo.2017.05.013.

Ettema, D., Nieuwenhuis, R., 2017. Residential self-selection and travel behaviour: What are the effects of attitudes, reasons for location choice and the built environment? J. Transp. Geogr. 59, 146–155. https://doi.org/10.1016/j.jtrangeo.2017.01.009.

Ewing, R., Cervero, R., 2010. Travel and the built environment: A meta-analysis. J. Am. Plann. Assoc. 76 (3), 265–294.

Ewing, R., Hamidi, S., 2014. Longitudinal Analysis of Transit's Land Use Multiplier in Portland (OR). J. Am. Plann. Assoc. 80 (2), 123–137. https://doi.org/10.1080/01944363.2014.949506.

Feinstein, A.R., 1995. Meta-analysis: Statistical alchemy for the 21st century. J. Clin. Epidemiol. 48 (1), 71–79. https://doi.org/10.1016/0895-4356(94)00110-C.

Gopalakrishnan, S., Ganeshkumar, P., 2013. Systematic Reviews and Meta-analysis: Understanding the Best Evidence in Primary Healthcare. J. Family Med. Primary Care 2 (1), 9–14. https://doi.org/10.4103/2249-4863.109934.

Gurevitch, J., Koricheva, J., Nakagawa, S., Stewart, G., 2018. Meta-analysis and the science of research synthesis. Nature 555 (7695), 175–182. https://doi.org/10.1038/nature25753.

Henry L., Litman T. (2006). *Evaluating new start transit program performance: comparing rail and bus*: Victoria Transport Policy Institute Victoria, BC.

Higgins, J.P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A., 2019. Cochrane handbook for systematic reviews of interventions. John Wiley & Sons.

Higgins, J.P.T., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. Br. Med. J. 327 (7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557.

Hong, A., Boarnet, M.G., Houston, D., 2016. New light rail transit and active travel: A longitudinal study. Transp. Res. A Policy Pract. 92, 131–144. https://doi.org/10.1016/j.tra.2016.07.005.

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.

Ibraeva A., van Wee B., Correia G. H. d. A., Pais Antunes A. (2021). Longitudinal macro-analysis of car-use changes resulting from a TOD-type project: The case of Metro do Porto (Portugal). *Journal of Transport Geography, 92*, 103036. https://doi.org/10.1016/j.jtrangeo.2021.103036.

Ingvardson, J.B., Nielsen, O.A., 2018. Effects of new bus and rail rapid transit systems – an international review. Transp. Rev. 38 (1), 96–116. https://doi.org/10.1080/01441647.2017.1301594.

Järv, O., Ahas, R., Witlox, F., 2014. Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records. Transport. Res. Part C: Emerging Technol. 38, 122–135. https://doi.org/10.1016/j.trc.2013.11.003.
Jeihani M., Zhang L., Ardeshiri A., Amiri A., Arefeh N., Zamir K. R., Baghaei B. (2013). Development of a framework for transit-oriented development (TOD). Retrieved from https://rosap.ntl.bts.gov/view/dot/26799.
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.
Liu, C.Q., Li, L., 2020. How do subways affect urban passenger transport modes?-Evidence from China. Econ. Transp. 23 https://doi.org/10.1016/j.ecotra.2020.100181.
Loo, B.P.Y., Chen, C., Chan, E.T.H., 2010. Rail-based transit-oriented development: Lessons from New York City and Hong Kong. Landsc. Urban Plan. 97 (3), 202–212. https://doi.org/10.1016/j.landurbplan.2010.06.002.
Lopez, B.J., Cummins, S., Gasparrini, A., 2019. Difference in difference, controlled interrupted time series and synthetic controls. Int. J. Epidemiol. 48 (6), 2062–2063. https://doi.org/10.1093/ije/dyz050.
Lund, H., Willson, R.W., Cervero, R., 2006. A RE-EVALUATION OF TRAVEL BEHAVIOR IN CALIFORNIA TODS. Retrieved from J. Archit. Plann. Res. 23 (3), 247–263. http://www.jstor.org/stable/43030775.
Maziarz, M., 2022. Is meta-analysis of RCTs assessing the efficacy of interventions a reliable source of evidence for therapeutic decisions? Stud. Hist. Philos. Sci. 91, 159–167. https://doi.org/10.1016/j.shpsa.2021.11.007.
Mikolajewicz, N., Komarova, S.V., 2019. Meta-Analytic Methodology for Basic Research: A Practical Guide. Front. Physiol. 10 https://doi.org/10.3389/fphys.2019.00203.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. Ann. Intern. Med. 151 (4), 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135.
Moore, G., Audrey, S., Barker, M., Bond, L., Bonell, C., Cooper, C., Hardeman, W., Moore, L., Cathain, A., Tinati, T., Wight, D., Baird, J., 2014. Process evaluation in complex public health intervention studies: the need for guidance. J. Epidemiol. Community Health 68 (2), 101. https://doi.org/10.1136/jech-2013-202869.
Ng, T.T., McGory, M.L., Ko, C.Y., Maggard, M.A., 2006. Meta-analysis in Surgery: Methods and Limitations. Arch. Surg. 141 (11), 1125–1130. https://doi.org/10.1001/archsurg.141.11.1125.
Pan, H., Shen, Q., Zhao, T., 2013. Travel and Car Ownership of Residents near New Suburban Metro Stations in Shanghai, China. *Transport. Res. Rec.* 2394 (1), 63–69. https://doi.org/10.3141/2394-08.
Paulley, N., Balcombe, R., Mackett, R., Titheridge, H., Preston, J., Wardman, M., Shires, J., White, P., 2006. The demand for public transport: The effects of fares, quality of service, income and car ownership. Transp. Policy 13 (4), 295–306. https://doi.org/10.1016/j.tranpol.2005.12.004.
Pulido, D., Darido, G., Munoz-Raskin, R., Moody, J., 2018. The urban rail development handbook. World Bank Publications.
Senior, M.L., 2009. Impacts on travel behaviour of Greater Manchester's light rail investment (Metrolink Phase 1): evidence from household surveys and Census data. J. Transp. Geogr. 17 (3), 187–197. https://doi.org/10.1016/j.jtrangeo.2008.11.004.
Song Y., Liu Y., Qiu W., Qin Z., Tan C., Yang C., Zhang D. (2020). MIFF: Human Mobility Extractions with Cellular Signaling Data under Spatio-temporal Uncertainty. *4*(4 %J Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.), Article 159. doi:10.1145/3432238.
Spears, S., Boarnet, M.G., Houston, D., 2017. Driving reduction after the introduction of light rail transit: Evidence from an experimental-control group evaluation of the Los Angeles Expo Line. Urban Stud. 54 (12), 2780–2799. https://doi.org/10.1177/0042098016657261.
Sun, G., Zhao, J., Webster, C., Lin, H., 2020. New metro system and active travel: A natural experiment. Environ. Int. 138 https://doi.org/10.1016/j.envint.2020.105605.
Tao, T., Cao, J., Wu, X.Y., 2021. The Road Less Traveled: Does Rail Transit Matter? J. Plan. Educ. Res. https://doi.org/10.1177/0739456x211035825.
Tawfik, G.M., Dila, K.A.S., Mohamed, M.Y.F., Tam, D.N.H., Kien, N.D., Ahmed, A.M., Huy, N.T., 2019. A step by step guide for conducting a systematic review and meta-analysis with simulation data. Trop. Med. Health 47 (1), 46. https://doi.org/10.1186/s41182-019-0165-6.
Thomas, B.H., Ciliska, D., Dobbins, M., Micucci, S., 2004. A Process for Systematically Reviewing the Literature: Providing the Research Evidence for Public Health Nursing Interventions. Worldviews Evid. Based Nurs. 1 (3), 176–184. https://doi.org/10.1111/j.1524-475X.2004.04006.x.
Thornton, A., Lee, P., 2000. Publication bias in meta-analysis: its causes and consequences. J. Clin. Epidemiol. 53 (2), 207–216. https://doi.org/10.1016/S0895-4356(99)00161-4.
Transport Department of Hong Kong. (2011). *Travel Characteristics Survey 2011* Retrieved from https://www.td.gov.hk/en/publications_and_press_releases/publications/free_publications/travel_characteristics_survey_2011_final_report/index.html.
Valentine, J.C., Pigott, T.D., Rothstein, H.R., 2010. How many studies do you need? A primer on statistical power for meta-analysis. J. Educ. Behav. Stat. 35 (2), 215–247.
Verma, A., Dhingra, S.L., 2006. Developing Integrated Schedules for Urban Rail and Feeder Bus Operation. J. Urban Plann. Dev. 132 (3), 138–146. https://doi.org/10.1061/(ASCE)0733-9488(2006)132:3(138).
Wang, J., Lu, Y., Yang, Y., Peng, J., Liu, Y., Yang, L., 2023. 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.
Wang, D., Zhou, M., 2017. The built environment and travel behavior in urban China: A literature review. Transp. Res. Part D: Transp. Environ. 52, 574–585. https://doi.org/10.1016/j.trd.2016.10.031.
Werner, C.M., Brown, B.B., Tribby, C.P., Tharp, D., Flick, K., Miller, H.J., Smith, K.R., Jensen, W., 2016. Evaluating the attractiveness of a new light rail extension: Testing simple change and displacement change hypotheses. Transp. Policy 45, 15–23. https://doi.org/10.1016/j.tranpol.2015.09.003.
Xiao, C., Goryakin, Y., Cecchini, M., 2019. Physical Activity Levels and New Public Transit: A Systematic Review and Meta-analysis. Am. J. Prev. Med. 56 (3), 464–473. https://doi.org/10.1016/j.amepre.2018.10.022.
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.
Xie, B., Lu, Y., Wu, L., An, Z., 2021. Dose-response effect of a large-scale greenway intervention on physical activities: The first natural experimental study in China. Health Place 67, 102502. https://doi.org/10.1016/j.healthplace.2020.102502.
Zhao, P., Li, S., 2018. Suburbanization, land use of TOD and lifestyle mobility in the suburbs: An examination of passengers' choice to live, shop and entertain in the metro station areas of Beijing. Retrieved from J. Transp. Land Use 11 (1), 195–215. https://www.jstor.org/stable/26622399.
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: Urban Anal. City Sci. 2399808320982305 https://doi.org/10.1177/2399808320982305.