

<https://doi.org/10.1038/s44333-024-00020-6>

The impact of the Bangkok metro expansion via Big Taxi GPS probe data



Eva Ayaragarnchanakul^{1,2}✉, Hannah Klauber³, Florian Nachtigall^{1,3}, Felix Wagner^{1,3}, Jia-wei Hu¹ & Felix Creutzig^{1,3,4}

Urban rail investments in megacities aim to reduce congestion and related externalities. This study investigates the impact of Bangkok's metro expansion on taxi trips using GPS-based data from 2017 to 2019. A rigorous approach to processing and analyzing the extensive traffic data in a difference-in-differences estimation framework shows that the metro expansion decreases taxi trips by approximately 32.2%, compared to areas not served by the metro system. This effect persists for catchment sizes of 300, 500, and 800 meters and increases over time. Bangkok's 19 new metro stations result in an annual climate benefit of approximately 0.17–0.34 MtCO₂-eq yearly or 0.21–0.42% of Thailand's transport sector emissions. Excluding private vehicle trips from the analysis, this is a conservative estimate of total climate savings. Multi-decadal CO₂ emission savings are valued at around \$1 billion, supporting the need for climate funding for transit systems in rapidly growing cities in developing countries.

Cities account for merely 2% of the earth's surface but release over 60% of the greenhouse gas emissions according to the United Nations. Cities are hence vital to reaching global climate targets. The transport sector is one of the highest emitting sectors. Global GHG emissions trends originating from human settlements are rising from 38 GtCO₂-eq in 1990 to 59 GtCO₂-eq in 2019, averaging 2.1% per annum growth during 2000–2010. Of the global trends, while net CO₂ from land use slightly diminished its share from 13% in 1990 to 11% in 2019, CO₂ from fossil fuel and industry escalated from 59% in 1990 to 64% in 2019¹. From 2010 to 2019, the GHG emissions in the transport sector rose faster than other sectors, averaging 1.8% annual growth², due to structural change³ arising from urban form, landscape, and infrastructure; travel behavior induced by preference, income growth, and transport mode choice; and new trends that affect demand concepts.

Apart from global warming, major transport externalities comprise air pollution, noise pollution, accidents, and congestion^{4–6}. Global transport externalities covering air, road, rail, and water modes are responsible for about \$13 trillion⁷ and the urban transport sector emitted 2.8 GtCO₂-eq in 2010⁸. Empirical evidence shows that new urban rail systems in mid-sized European cities with over 100,000 inhabitants can ease congestion, travel time, and pollution⁹. Similarly, America's densest metropolitan areas and those with mature railway networks are the lowest carbon emitters per capita¹⁰. From this point forward, we use the terms metro, urban rail, and heavy rail interchangeably.

Although unable to separate the effect of metro lines from buses, strikes have proven that removing public transit has increased congestion. In Los Angeles, for instance, a transit strike in 2003 led to a 47% increase in the pre-strike average in freeway travel times, linked to transit benefits of around \$1.2–\$4.1 per passenger-mile during peak-hour¹¹. Public transit strikes in Rome also show short-run effects of travel time increases of about 0.017 min/km and 0.0065 min/km during peak- and off-peak hours, respectively¹².

In coping with transport-related externalities, many studies state the relevance of vehicle kilometers traveled (VKT) or vehicle mileage traveled (VMT), especially in mitigating urban climate. Policies with the largest GHG reduction opportunities are typically those that reduce VMT¹³. From a 2001 household travel survey in Santiago, the elasticity of VKT due to the distance to the metro's effect on both car ownership and usage is positive¹⁴. A 10% increase in the density of rail miles reduced annual VMTs by 4.2 percentage points in the 26 US urbanized rail transit cities in 1990¹⁵. From the 2001 national household survey in the US, the predicted reduction of VMT per unit of rail availability is –10.9 VMT daily¹⁶. In addition¹⁶, estimated a reduction of 5.2 billion gallons/year in fuel use is associated with transit availability (rail and bus), implying a reduction of carbon emissions too.

Nonetheless, some studies do not support a reduction of VMT because of an increase in rail availability. Urban rail ridership and supply had a weak and indirect effect on VMT per capita in 370 US urbanized areas¹⁷. The

¹Technische Universität Berlin, Sustainability Economics of Human Settlements, Straße des 17. Juni 135, 10623 Berlin, Germany. ²Prince of Songkla University, Faculty of Economics, 15 Kanjanavanich, 90110 Songkhla, Thailand. ³Potsdam Institute for Climate Impact Research, Member of the Leibniz Association, EUREF Campus 19, Torgauer Str. 12–15, 10829 Berlin, Germany. ⁴Bennett Chair of Innovation and Policy Acceleration, University of Sussex Business School, Jubilee Building, Arts Rd, Falmer, Brighton BN1 9SL, UK. ✉e-mail: eva.a@win.tu-berlin.de

authors observed that the ability to capture a strong direct effect may be limited due to the use of cross-sectional data from the year 2003 only. However, another study (also using cross-sectional US data from 2003–2004) calls this into question by indicating that previous studies may suffer from omitted variable bias by not including network load density¹⁸. This study demonstrates that the density near rail stations induces walking and transit use. Introducing fixed effects to account for area- and time-specific differences in rail usage into the model will improve the reliability of results.

Fast-growing cities in Low-to-Medium Income Countries (LMICs) largely have less matured urban rail networks compared to High-Income Countries (HICs) and, therefore, have dissimilar commuting behavior. This raises the question of whether the rail system is worth investing in for LMICs when making investment decisions under limited resources. Existing literature on the impacts of urban rail on traffic congestion and distributional effects on other travel modes in these cities is inconclusive. Some studies mentioned that new metro stations do not necessarily alleviate road traffic congestion when measured in terms of traffic speed¹⁹, VKT²⁰, and cross-area trips²¹. On the contrary, Gu et al.²² examined 45 subway lines across 42 cities in China and found that nearby road speed increased by 4% after the first year of a line opening. Further, there are reductions in vehicle ownership and driving of residents in the metro catchment area^{20,23}. Also, changes take place in modal share. Metro trips replace walking²⁴, biking^{24,25}, bus rides^{19,24,25}, and car rides^{21,25}. Additionally²⁶, determined that new metro stations cause a decrease in ridership of old stations.

Studies in North and South America used cross-sectional data with analysis techniques such as Structured Equation Modeling (SEM)^{16,17}, negative binomial regression¹⁸, regression discontinuity¹¹, and Discrete Choice Modeling^{14,15,27}. Studies in LMICs are clustered in fast-growing Chinese cities like Nanchang, Xi'an, Shenzhen, and Beijing. They mostly use cross-sectional or repeated-panel surveys of either self-reported commuting trips in the travel diary^{20,21,24,25} or ridership reports^{19,26}. Various analysis techniques have been applied, for instance, Tukey's multiple comparison test²⁴, ordinal logit regression²⁰, Heckman two-step model^{20,23}, regression discontinuity design¹⁹, and difference-in-differences regression^{25,26}. However, only a few of them explored the temporal dynamics of the impact²².

A few studies focus on the relationship between the metro and taxis. The built environment factors such as the neighborhood size (in terms of population) and the availability of alternative transit modes play a crucial role in the metro-taxi correlation^{28–32}. Depending on the trip characteristics, taxis can be subway-competing, subway-extending, or subway complement^{33–35}. These studies circulate the prevention of taxis absorbing subway ridership. Our study is associated with this idea. Further^{32,35}, generated taxi trips from taxi light status with minor preprocessing by essentially removing missing data and outliers.

To the best of our knowledge, the study by³⁶ provided an aligned concept to ours. They investigated the impact of the first metro line in Wuxi

in 2014 on taxi ridership. They measured the in- and out-degree values of node linkages of the taxi networks distributions from taxi GPS trajectories four weeks before and after the metro opened. The impact on taxis is greater at locations closer to subway stations and farther from the CBD, but the relation can be both positive and negative depending on the location.

We present a case study of the heavy-rail expansion in Bangkok on taxi trips. To this end, we use panel data and a two-way fixed effects estimation that controls for weather conditions, area-specific factors such as local amenities, and time-specific events that may affect taxi trips and observe how the effect grows over time. Based on the big GPS-based traffic data, we meticulously filter relevant taxi trips and assign them to the treatment or control group. This study is the first of its kind to be applied in a Southeast Asian city context.

Our contributions to the related literature are two-fold. First, we demonstrate that in Bangkok, a monocentric megacity with polycentric subcenters belonging to an LMIC, extending the urban rail network to serve the areas outside of the inner ring road has the potential to reduce traffic, specifically by decreasing taxi trips. Second, we introduce a conservative cleaning process for taxi GPS data to minimize taxi signal light malfunctioning errors on top of the common method of filtering by relying on the changing signal lights and correcting for minor errors. In addition, we provide a procedure to overcome the methodological challenges from taxi trajectory malfunctioning cleaning, selection of the extent of the control area, dealing with overlapping of metro stations, to trimming out the COVID-19 shock. We also highlight that the treatment effect of the metro expansion takes approximately half a year to allow for mode-shifting behaviors from taxis to rail.

The remaining outline of this paper is organized as follows: results, discussion, and methodology.

Results

Two-way Fixed Effects (TWFE) regression model

According to Table 1, we find that after the new metro stations open, in the 800 m catchment full specification model, there is a statistically significant reduction of 1.59 taxi trips around new metro stations compared to the average taxi trips of 4.94 per week or a 32.2% reduction. We use catchment areas 500 m and 300 m as sensitivity analyses. Since they are not statistically significant in the full specification model, we will discuss more about this in the next section.

Dynamic treatment effect

We investigate further in Fig. 1, the dynamics of the treatment effect. Before opening new metro stations, the number of taxi trips was mostly insignificantly different between the treatment and control groups – substantiating that the parallel trend assumption holds. After new stations open, we can observe that the number of taxi trips step-wisely decreases right after and in the 6th month (or around the 23rd week, see Supplementary Tables 2–4)

Table 1 | Estimated coefficients from the two-way fixed effects regression (Eq. (1))

Variable	300m ^a			500 m			800 m		
Treat	−0.31 (0.203)	−0.25 (0.221)	−0.27 (0.196)	−0.94 ^c (0.486)	−0.96 ^c (0.489)	−0.8 (0.495)	−1.73 ^d (0.788)	−1.76 ^d (0.789)	−1.59 ^d (0.799)
95%CI	[−0.71, 0.09]	[−0.68, 0.19]	[−0.65, 0.12]	[−1.89, 0.02]	[−1.92, 0.003]	[−1.77, 0.17]	[−3.28, −0.19]	[−3.31, −0.22]	[−3.15, −0.02]
Weather variables	no	no	yes ^b	no	no	yes	no	no	yes
District-Month FE	no	yes	yes	no	yes	yes	no	yes	yes
Time FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
Unit FE	yes	yes	yes	yes	yes	yes	yes	yes	yes

Station-week clustered standard errors are in parentheses. 95% confidence intervals are in brackets. The dependent variable is taxi trip count in 800/500/300 m catchment of metro stations.

^aStation N13 is dropped from the analysis because of insufficient number of taxi trips.

^bHeavy rain is excluded due to many zeros.

^cIndicates significance level at 10%.

^dIndicates significance level at 5%. The 800 m model shows a statistically significant reduction of around 1.6 trips or 32.2% after the metro expansion took place.

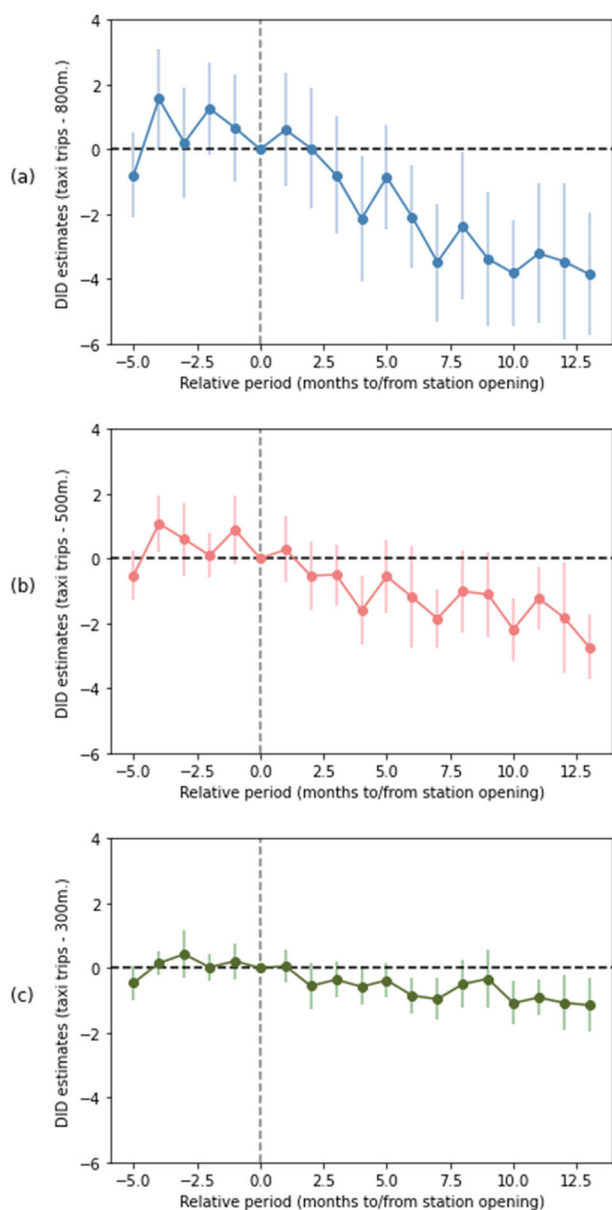


Fig. 1 | Dynamic difference-in-differences coefficient estimates of taxi trip count (Eq. (3)) and their corresponding 95% confidence interval in 3 catchment sizes around metro stations. a 800 m (blue plot), b 500 m (pink plot), and c 300 m (green plot). The treatment effects are immediate in all cases. See Supplementary Tables 2–4 for the estimated coefficient tables.

after the opening of new metro stations in the 800 m catchment model. Even though the effects are not as gradually decreasingly visible in the 500 m and 300 m catchment models, we can see immediate and statistically significant effects in both models too. The treatment effect is up to -3.9 , -2.7 , and -1.1 fewer trips in catchment sizes 800 m, 500 m, and 300 m, respectively. The results suggest that while the overall average treatment effect is insignificant for catchment sizes less than 800 m, the weekly dynamics of the treatment effect allow us to gain additional information that there are effects in the treatment over the weeks after receiving treatment. The effects are weaker for smaller buffer sizes because there are very few taxi trips thus potentially creating insufficient variation for the analysis.

Potential climate benefits

Back-of-the-envelope calculations suggest that if a 32.2% reduction in taxi trips occurs according to our estimation and 97% of trips happen during 6 a.m. to 12 a.m., then the climate benefits from this reduction in accordance

with the metro expansion is at most approximately 1,018.4 tCO₂-eq/day. This number is calculated using Eq. (4), assuming taxi trip reduction happens across Bangkok and taxis are affected evenly, given the average travel distance for a commute trip in Bangkok is 12.64 km/trip and the total taxi trips are 1.36 million daily³⁷ and other related data are given in Supplementary Table 5.

However, since the electricity consumption emission cost of riding the metro is 97.7 tCO₂-eq/day, the net benefit reduces to 921 tCO₂-eq/day. The emission cost is estimated from electricity consumption of 0.39 kWh/passenger, the number of increased metro ridership (since taxi occupancy is about 1 passenger, then we assume a 1:1 shift from a taxi to the metro) (i.e., 424,782 passengers), and emission factor of grid electricity of 0.5897 tCO₂/MWh, according to³⁸.

Due to the absence of information on the number of affected taxi trips from the metro expansion areas, a sensitivity analysis can be carried out by altering the proportion of affected taxis in Eq. (4), but in essence, the benefit of the metro expansion is about 10 times the cost. For instance, the climate benefit reduces to 645 tCO₂-eq/day and 460 tCO₂-eq/day if 70% and 50% of the taxi trips, respectively, are affected by the metro expansion. Assuming half of the taxi trips affected should be the lower bound of the analysis since our data supports that taxis are much more likely to drive around the metro system coverage zones, i.e., the more urbanized areas of Bangkok (see Supplementary Fig. 2).

The fact that private vehicles share 69% of all trips in Bangkok, the climate benefit from expanding the metro system of 921 tCO₂-eq/day is reasonably below road pricing of private vehicles where emissions reduction benefits are from 2484 to 9575 tCO₂/day, depending on commuter's elasticity of demand⁴.

Discussion

This study demonstrates that urban heavy rail expansion can reduce the number of taxi trips on the congested roads of Bangkok. The impact of the reduction is stepwise – magnifying after approximately 6 months. This duration is useful for policymakers to refer to when required to evaluate a policy effect. Some commuters' behavior may change immediately when an alternative presents itself while others may take more time to adjust. However, the impact is sensitive to the catchment size selection. This result is an advocate for transit-oriented developments. The local government should improve walkability around metro stations up to 800 m. The overall average number of taxis is insignificantly different between the control and treatment areas for smaller catchment sizes because the effect is rather dynamically gradual. Hence, dynamic treatment effects should be considered as well as the one-point-in-time treatment from the beginning of a policy intervention.

We contribute to academic knowledge by using a rich geodata processing protocol. We are also the first to evaluate the impact of the metro system on the number of taxi trips using the dynamic difference-in-differences method.

As Thailand's capital city, Bangkok's busy road traffic cost approximately 7–10.8% of the provincial GDP in 2017. Local government is putting effort into alleviating road traffic congestion by providing alternative transport modes. One prominent strategy is to improve the coverage of the metro system to drive people off roads and onto rails (Mass Rapid Transit Master Plan for the Bangkok Metropolitan Region 2010–2029: M-MAP). The metro system, which began operating in late 1999, now spans 211.94 km (see Supplementary Table 1). An additional 90.2 km is currently under construction, with a total planned expansion reaching 509 km (source: www.m-map2thailand.com). A simple linear regression shows an increasing ridership trend (1,000 people-trips per day) for both the MRT and BTS from January 2011 to December 2019: MRT ridership follows the equation $\text{Ridership} = 0.046t - 1665.9$ ($R^2 = 0.8455$) (data source: <https://investor.bemplc.co.th/ridership-report/ridership>), while BTS ridership follows $\text{Ridership} = 2.57t - 89087$ ($R^2 = 0.8116$) (data source: http://mistran.otp.go.th/mis/Interview_HIBTSVolume.aspx). Our results confirm that fewer taxi trips to the existing stations occurred after expanding the metro

system. However, metro ridership would gain from accessibility improvement in terms of safe access and cost-efficient feeder services, especially for vulnerable individuals³⁹.

The results should be used with careful interpretation. The impact cannot be transferred to all modes of road traffic. Taxis accounted for 1.36 million trips/day or 4.2% of all trips in 2017 reported by the Bangkok Household Travel Survey³⁷. Further research on the impact on private vehicles would substantiate the benefit of the metro expansion. This, however, is not possible under current conditions regarding data availability. There is potential demand for shared pooled mobility and other modes of transportation such as the metro, but the shift away from private vehicles can happen only if private vehicles are charged for their social costs^{18,40}. It was calculated that this cost should be internalized by an inner-city road charge of 2.7 USD in combination with a 3.2 USD hourly parking fee⁴.

Although our results cannot be used to directly draw inferences about road traffic congestion in general, we illustrate that the expansion of the metro system leads to fewer taxi trips in the catchment area of new metro stations. This finding relieves the issue concerning taxis winning the metro-taxi competition.

Nonetheless, back-of-the-envelope calculations suggest that the potential climate benefit associated with Bangkok's metro expansion is approximately 0.17–0.24 MtCO₂-eq/year or at most 0.34 MtCO₂-eq/year. Given that the transport sector in Thailand and Bangkok, respectively, emit 79.6 MtCO₂-eq/year in 2022⁴¹ and 12.42 Mt-CO₂eq/year in 2016⁴², this climate benefit accounts for around 0.21–0.3% (at most 0.42%) of Thailand's and 1.35–1.89% (at most 2.71%) of Bangkok's annual transport GHG emissions.

This research is not without limitations. First of all, like most transport GPS data, the representation of the data may be biased. The number of registered taxis in Bangkok was around 71,000 vehicles⁴³. Our dataset contains around 1000–3000 taxis daily. Further, we initially included more treatments during and after COVID-19. However, the result from the treatment is mixed and sensitive to the selection of control and treatment areas. This suggests that there may be some shifts in the commuting behavior of Bangkokians after COVID-19. Also, ride-hailing services like Grab which owns 70% of the market share, were legalized in May 2021, although it was well-recognized before that. The legalization allows taxi drivers to simultaneously join ride-hailing services, so we cannot predict how this legalization would affect the studied taxi trips. Conservatively, removing the data before COVID-19 removes any effect that ride-hailing services may have on taxis. Additionally, we only estimate the direct substitution of taxi usage with the metro, excluding the indirect effect through land use patterns. We also excluded many overlapping stations in the 800 m catchment area from the analysis and applied this logic to all catchment sizes. This may result in an underestimated treatment effect. To be able to accurately calculate climate benefits, we are extending this research by looking at the impact of the metro expansion directly on taxi vehicle-kilometers traveled. We also plan to estimate the change in road link speed or vehicle speed to further evaluate the congestion cost. This research sheds light on the importance of private vehicle GPS data collection (conditioned on privacy and security) in gaining greater benefits when assessing the impact of transport policies on overall road traffic.

Methodology

We use the two-way fixed effects model and dynamic difference-in-difference method to measure the effect of metro expansion on taxi trips and explore how effects evolve over time. Our outcome variable, the number of taxi trips, is processed using a conservative filtering and trip selection framework.

Model assumption

We assume that the metro system expansion reduces the number of taxi trips from areas with a newly opened station to the old metro stations' catchment. This assumption is built on the minimum distance to the nearest

metro station criteria. The introduction of a new metro station reduces the distance to the metro system connection. Therefore, commuters starting trips from the new metro catchment have the opportunity to shift their travel mode from taxis to the metro. Simultaneously, commuters starting trips from elsewhere will remain using taxis because there is no change in the availability of transportation modes in their neighborhood. We describe this graphically in Fig. 2. In addition, there are no car promotional policies launched during the study period and the fuel price is quite stable during the same time. As an example, the fuel price excluding local tax can be found here: <https://www.bangchak.co.th/th/oilprice/historical>. Accordingly, the opening of new stations does not collide with other major incidences affecting traveling behavior.

Data description and processing

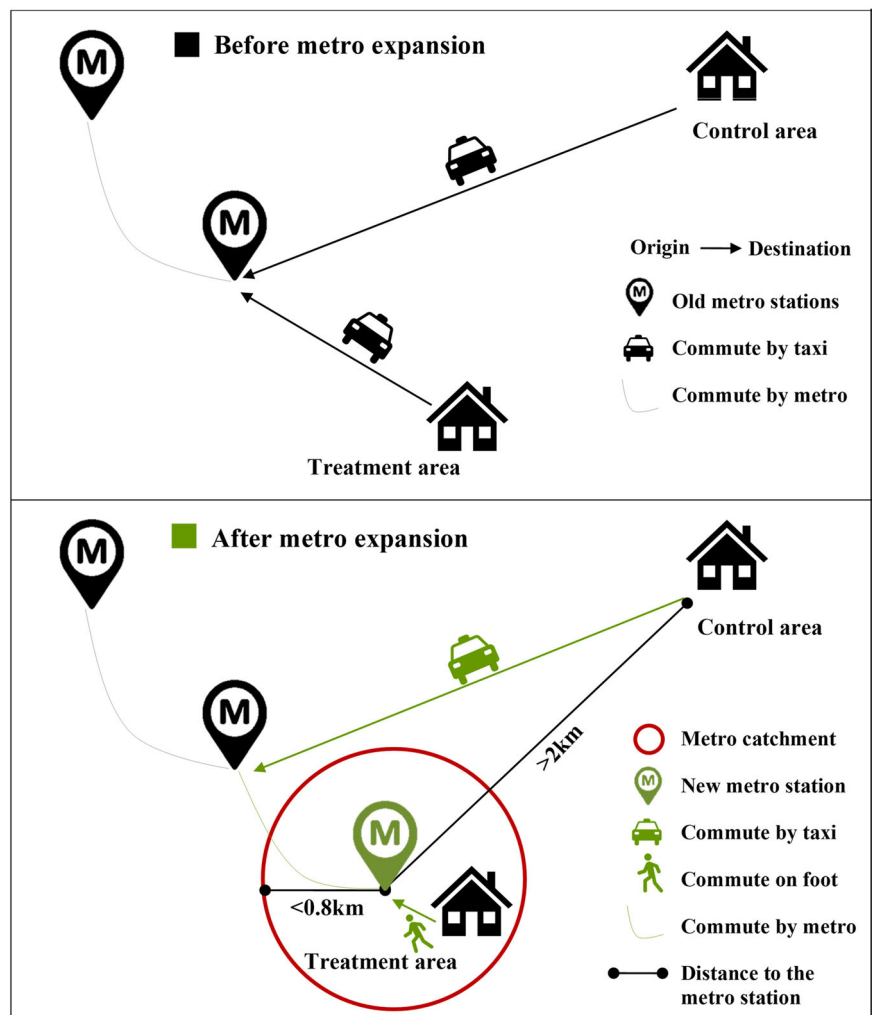
The Intelligent Traffic Information Center (iTIC) Foundation of Thailand provides open access traffic data which was gathered from various sources such as intelligent traffic signs, traffic surveillance cameras at intersections, mobile probes from taxis, logistic trucks, buses, private vehicles, and Longdo traffic map application⁴⁴ voluntary reports. According to the Bangkok Household Travel Survey in 2017, taxis accounted for 4.2% of all commuting trips and the metro accounted for 1.3% of all commuting trips. We use the GPS trajectories of each vehicle with a temporal resolution of every minute if the vehicle engine is active and every three minutes if it is inactive (see further details in⁴⁵). Available data now spans from 1 January 2017 to 31 December 2023. However, we trimmed the dataset to pre-COVID-19, i.e., 1 January 2017 to 12 January 2020 since the first COVID-19 case in Thailand was reported on 13 January 2020. We also analyzed by including the post-COVID-19 period but found mixed effects. We suspect that commuters altered their behavior due to COVID-19. This is supported by the significant drop in MRT ridership from March 2020 until December 2022. MRT ridership surpassed its pre-COVID-19 number from January 2023 (source: <https://investor.bemplc.co.th/th/ridership-report/ridership>).

The key variable is the number of taxi trips that start in the buffered areas of new metro stations (stations opened from 6 December 2018 onward) or in hexagons outside of the 2 km radius (and within 3 km) from the metro system and end in the buffered areas of the old metro stations (stations opened before 6 December 2018) as shown in Fig. 3. We initially used the hexagons 2–5 km criteria from the metro system, but results were consistent with hexagons 2–3 km from the metro system but with increased running time. This essentially means that the hexagons located farther than 3 km from the metro system do not provide more information to the analysis, i.e., the data contains many zeros. The spatial distribution of the average taxi trips by spatial unit in the 800 m buffer model is shown in Fig. 4.

The 2–3 km criteria selection is motivated by the desire to ensure that the distance is sufficiently large (>2 km) such that hexagons are not affected by new stations while sufficiently small (<3 km) such that the treatment and control areas are comparable in terms of, for example, population composition. We refer to the trips starting in the buffered areas around new metro stations as the “treatment” group and trips starting in the hexagons within the 2–3 km radius from the metro system as the “control” group. In the 800 m catchment of all stations, if new stations overlap with old stations, they are removed from the analysis. Due to this station selection criteria, 19 new stations opening on 4 different dates during the period of study were included in the analysis, which can be seen in Table 2.

The geoprocessing protocol in Fig. 5 ensures that all trips in the treatment group can be made by using the metro. Thus, carving out the true effect of the metro expansion on taxi trips, i.e., commuters possibly shifting from riding taxis to the metro after the metro expansion takes place. For the sensitivity analysis, we chose 3 alternative buffer sizes: 300, 500, and 800 m to ensure that all bands are within 15-minute walking distance^{24,26}. Although Deng & Zhao (2022) raised the issue of treatment effect underestimation from selecting too restricted catchment area, we limit the catchment area to 800 m in our main analysis for three reasons: (1) it is the average farthest walking distance for Bangkokians from the travel survey in 2021 from⁴⁰, (2) metro station catchment areas are not pedestrian-friendly⁴⁶, and (3) the

Fig. 2 | Model assumption. In the new metro station catchment (i.e., treatment group), commuters have the opportunity to shift from starting a trip by riding taxis to using the metro to reach their destination in the old metro catchment after the expansion. Commuters starting trips farther than 2 km from any metro station (i.e., control group) will likely continue riding taxis to reach their destination in the old metro catchment since no mode availability changes there.



weather in Bangkok is hotter than in most case studies. If the 800 m buffer zones of different simultaneously introduced new stations overlap, the trips starting in the overlapping areas are assigned evenly to the corresponding stations. We use the “H3” library in Python (<https://h3geo.org/>) to create the hexagons in size 8 (of approximately 600 m sides). Theoretically, the selection of hexagons and circular buffers would not affect the outcome of our analysis because we also included unit fixed effects in our model to control for the difference between these units in the regression model. We generated the taxi trip count from taxi probe trajectories by using the filtering process as described in Fig. 5. This cleaning process guarantees (conservatively) that trips are valid, and not wronged by defective taxi status lights.

The data processing rule from the literature relies mainly on trips filtered by taxi status lights. However, with careful observation while cleaning the data, we found several unreasonable trips. Hence, we tune the preprocessing rules as follows and demonstrate how our preprocessing method is crucial by providing the remaining trips in each process in parentheses. We start preprocessing taxi trip validity shown in Fig. 5 (left) by dropping unidentifiable trajectories – those without light status indication. Then if trajectories are not changing their light/occupancy status or are shorter than 5 meters apart, we drop these duplicated trajectories (100% of trips generated). We identify trips with their status light (50% of trips remaining). But, if the taxi is unoccupied, so-called deadheading, we drop these trips. If taxis are not moving for at least 6 min, we consider them as interrupted trips and remove them from the sample (43% of trips remaining). At the occupancy status change

trajectory, we drop trips starting with a speed exceeding 40 km/hour (33% of trips remaining). Lastly, we filter out trips shorter than 1.5 km (22% of trips remaining).

In Fig. 5 (right), the geoprocessing step begins with filtering out the trips outside the metro operating hours of six in the morning to mid-night. After identifying the trip origin and destination locations, if the trips lie partially outside of the boundary of the BMR, we leave them out of the sample. If trips originate within the 2 km circular buffer of new metro stations and end in the 2 km of the old metro stations, they are considered the treatment group. For trips originating in hexagons within a 2–3 km radius from any metro station and ending in the 2 km of the old metro stations, they are considered the control group. The basis of selecting the control and treatment groups is to calculate the differences in the number of taxi trips between captive commuters (i.e., controlled areas unserved by the metro system for the entire period of study) and non-captive commuters (i.e., treated areas that were previously not served by the metro system and became metro-available after a station opening).

The Air Quality and Noise Management Bureau, Pollution Control Department of Thailand provides hourly statistics on weather conditions. There are 23 weather monitoring stations located in the BMR as shown in Fig. 3, although not all weather variables are always available from all stations every day. Hence, we use the inverse-distance (squared) weighted (IDW) spatial extrapolation to account for missing data. We use two weather variables: average temperature (deg. Celsius) and average rainfall height (mm).

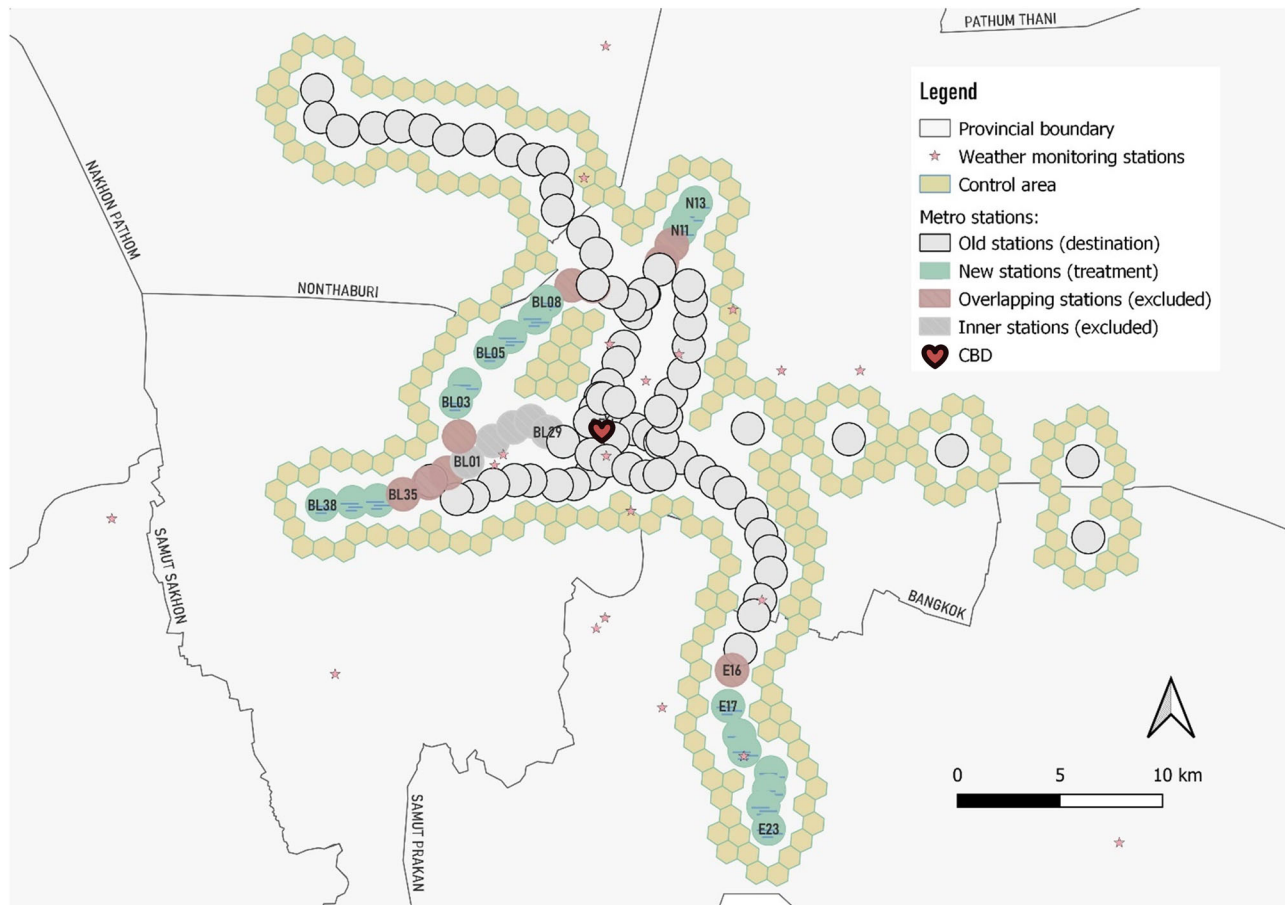


Fig. 3 | Study area and quasi-experimental design. We select all trips ending in the catchment area of old stations (black-outlined circles). Of those trips, we determine those that start in the catchment area of new stations (green circles) as the *treatment group*. The 800 m buffer is represented in this figure. Trips starting in areas not covered by the metro system (yellow hexagons) refer to the *control group*. Stations

overlapping between old and new stations (pink circles) and stations that are inside of the inner ring road of Bangkok (gray circles) are left out of the analysis as they may cause biased results. The CBD, where the “Siam” station is located, is marked with a red heart. See Supplementary Table 1 for details of station names and opening dates.

In addition, the iTIC foundation also provides access to traffic incident reports. This information is in extensible markup language (XML) format, geocoded, timestamped, and updated every 30 min. We extracted the number of reports of heavy rain as this may induce commuting by taxis and the metro.

Estimation method: difference-in-differences

The difference-in-differences (DiD) method is a common method used in causal inference, especially to measure the outcome of policy interventions⁴⁷. We use the two-way fixed effects model, also referred to as the generalized difference-in-difference model, to estimate the treatment effect and explore its temporal dynamic. We tested that the data does not violate the parallel trend assumption and checked the robustness of the treatment effect by using three catchment areas with 300, 500, and 800 m buffer sizes for the treatment group. We also found no multicollinearity in the model after testing the control variables with the variance inflation factor (VIF).

We are interested in estimating τ from Eq. (1) to measure the treatment effect of the treatment group after they receive the treatment.

$$\text{Origincount}_{it} = \tau \text{Treat}_{it} + x_{it}\gamma + \lambda_t + c_i + \theta_{dm} + u_{it} \quad (1)$$

$$x_{it}\gamma = \gamma_1 \text{temperature}_{it} + \gamma_2 \text{rainfall}_{it} + \gamma_3 \text{heavyrain}_{it} \quad (2)$$

where $i = \text{unit}(\text{buffer/hexagon})$; $t = \text{week}$; $m = \text{month}$; $d = \text{district}$,

Origincount_{it} is the number of taxi trips starting in unit i in week t ,

$$\text{Treat}_{it} = \begin{cases} 1 & \text{for unit } i \text{ if week } t \text{ is after station was opened} \\ 0 & \text{otherwise} \end{cases},$$

$x_{it}\gamma$ is a vector of control variables that account for factors that vary at the unit level over time, such as local weather (Eq. (2)),

temperature_{it} is the weekly average temperature (deg. Celsius),

rainfall_{it} is the weekly average rainfall (mm),

heavyrain_{it} is the number of weekly reports of heavy rain,

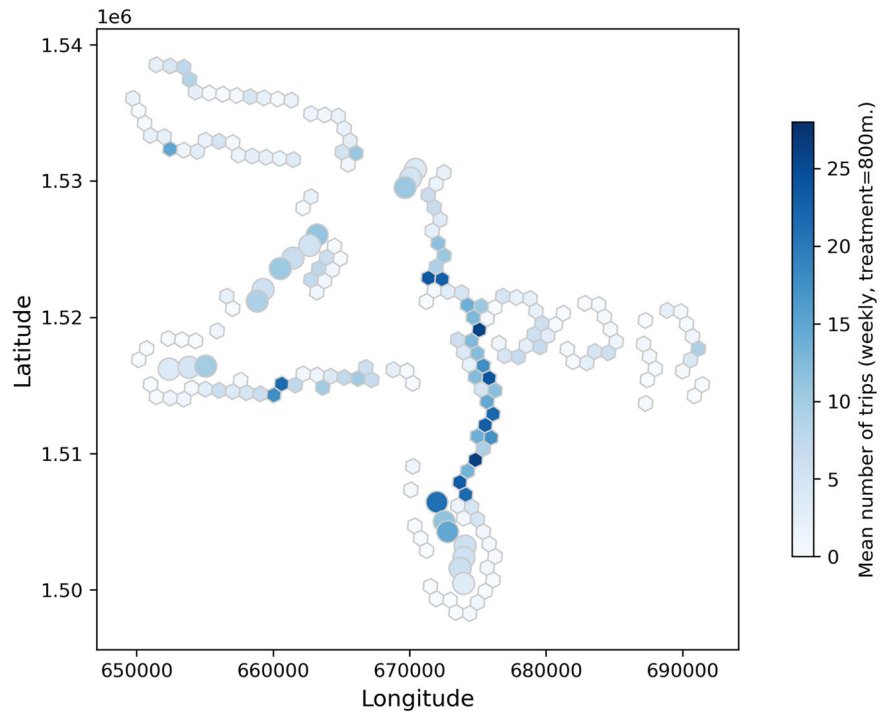
λ_t is a time FE that controls for time-varying factors at the weekly level (e.g. holidays, fuel prices),

c_i is a unit FE that controls for buffer or hexagon area-specific differences that are time-invariant (e.g. point of interests, walkability, station amenities), and θ_{dm} is a district-month fixed effect that controls for district-specific time-varying factors, thereby restricting identifying variation to weeks before and after treatment within the same month in the same district (e.g. special events). This captures any monthly changes common to all units within a district.

In addition to the pooled post-treatment effect, we estimate the treatment effects in each month after treatment over a period of about one year according to Eq. (3). This method is called dynamic DiD or event study. The relative periods measured in each month are done by pooling weeks within each month together. The aim of this approach is to increase the statistical power of the estimation and thereby improve the precision of the

Fig. 4 | The spatial distribution of the mean taxi trips by spatial units in the 800 m buffer model.

The average number of weekly taxi trips during the study period (1 January 2017 to 12 January 2020 or 158 weeks) in the control group (180 hexagons) and treatment group (800 m radius circles around 19 stations) is up to almost 30 trips/week/unit of area.

**Table 2 | Timeline of station openings in the observational period**

Opening date	Line name: station names (IDs)	No. of stations
06.12.2018	BTS Sukhumvit: Chang Erawan museum (E17) – Kheha (E23)	7
21.09.2019	MRT Blue: Phasi Charoen (BL36) – Lak Song (BL38)	3
04.12.2019	BTS Sukhumvit: Ratchayothin (N11) – Kasetsart University (N13)	3
04.12.2019	MRT Blue: Bang O (BL08) – Sirinthorn (BL06)	3
23.12.2019	MRT Blue: Bangyikan (BL05) – Fai Chai (BL03)	3
Total		19

Out of all the new stations that started operating during the period of study (1 January 2017 to 12 January 2020) (see Supplementary Table 1 for further details on all the metro station opening dates since the first metro line in 1999), 19 stations opening on 4 different dates were included in the analysis since they did not overlap with any old station in the 800 m catchment model.

estimated coefficients. The treatment effects measured in weeks after treatment (see Supplementary Note 1) are reported in Supplementary Figure 1. We show that the data follows the parallel trend assumption prior to treatment. The pre-treatment coefficients should not be statistically significantly different from zero, i.e., there should be no difference in the taxi trip count between the control and treatment groups. The post-treatment coefficients indicate the effect of the treatment over time which provides more information on the time dimension than τ obtained from the TWFE model (Eq. (1)).

$$\text{Origincount}_{it} = \sum_{j=0}^{T_{bf}-2} \tau_{-(T_{bf}-j)} \text{Treatment}_{it} + \sum_{k=0}^{T_{af}} \tau_{(T_{af}-k)} \text{Treatment}_{it} + x_{it}\gamma + \lambda_t + c_i + \theta_{dm} + u_{it} \quad (3)$$

where τ_{-1} is used as the reference period (i.e., the month prior to the opening month) and dropped out of the model,

$T_{bf} = 5$ (the 5th pre-treatment period), and $T_{af} = 13$ (the 13th post-treatment period).

The potential climate benefit could be calculated after the treatment effect estimation using the following equation.

$$\Delta \text{Emission} = \left(\Delta \text{trips} \times \text{metro\_open} \times \text{taxis\_affected} \times \text{no\_trips} \times \sum_i p_i f_{e_i} \times \text{ave\_dist} \right) \times 10^{-3} \quad (4)$$

where $\Delta \text{Emission}$ is the GHG emission change from expanding the metro stations (tCO₂-eq/day),

Δtrips is the trip change from commuters shifting from taxis to the metro (%),

metro_open is the proportion of taxi trips during the metro opening hours (%),

taxis_affected is the proportion of taxi trips affected by the new metro opening (%),

no_trips is the baseline number of daily trips made by taxis (trips/day),

p_i is the proportion of taxis using fuel type i (%),

f_{e_i} is the fuel efficiency of the fuel type i (kgCO₂-eq/km),

ave_dist is the average travel distance (km/trip/day), and i stands for fuel types of LPG, CNG, and Benzene.

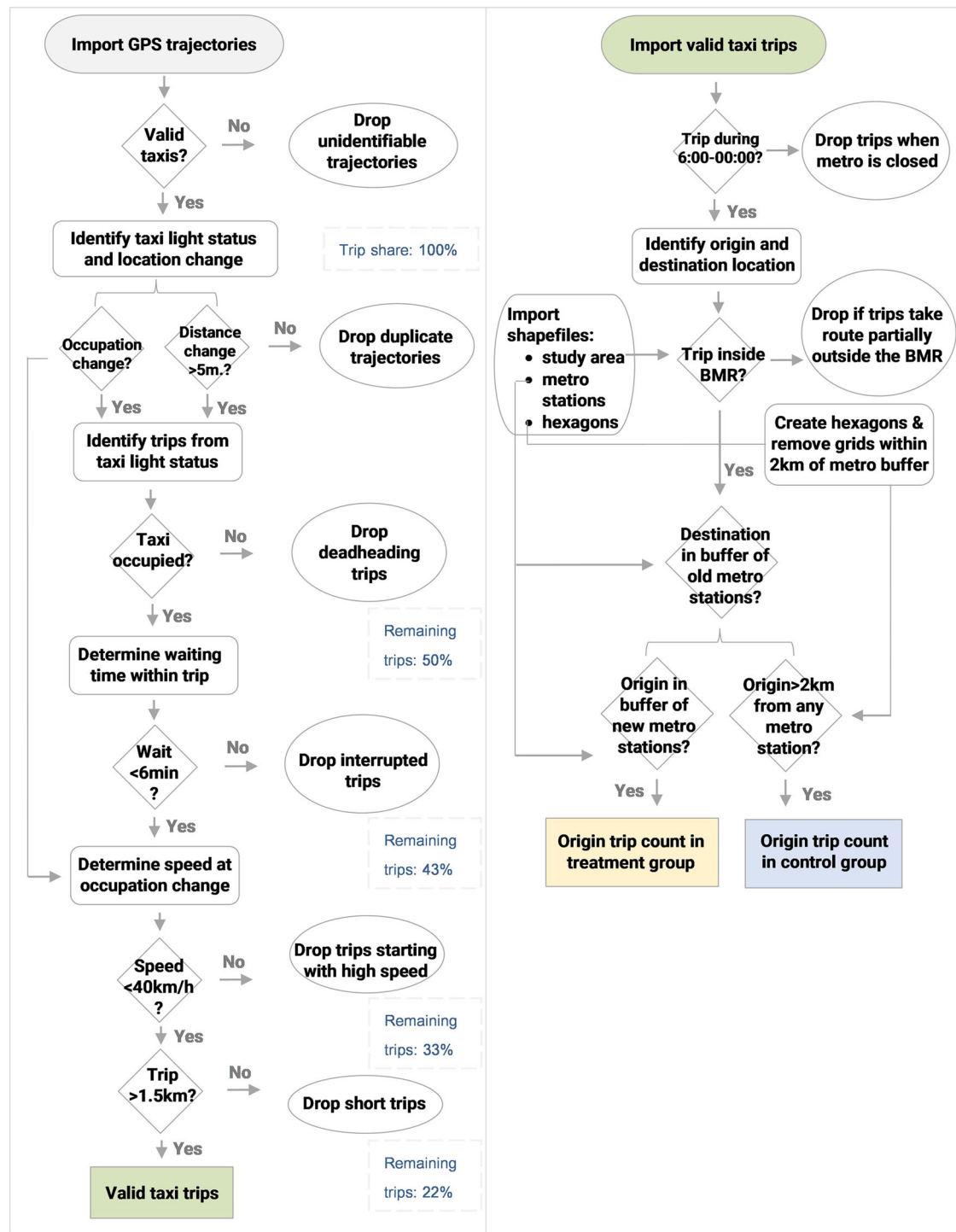


Fig. 5 | Data processing framework for taxi trip count: preprocessing (left) and geoprocessing (right). Data preprocessing involves trajectories filtering and trip generation, following these steps: (1) drop trajectories without light status indication, (2) drop trajectories not changing light/occupancy status or are shorter than 5 meters apart, (3) use taxi light status to identify trips, (4) drop unoccupied trips, (5) drop taxis not moving for at least 6 min, (6) drop trips starting with a speed

exceeding 40 km/hour, and (7) filter out trips shorter than 1.5 km. To provide the importance of this process, the proportion of trips left in each step is also given. Data geoprocessing ensures that the taxi trips between the control and treatment groups are comparable by the following steps: (1) drop trips when the metro is not operating, (2) drop trips partially outside the BMR, and (3) identify the control group and treatment group.

Conclusion

Urban transport planners need evidence in deciding on the policies and investments that are worth wasting valuable resources. Our results support urban rail investments by stating the case of its benefit in reducing the number of taxi trips in the megacity of Bangkok, a representative of LMICs with limited resources and rapidly increasing demand for high-quality

transit. We find that by expanding the urban rail system, taxi trips along the 19 new urban rail stations are reduced by at most approximately 32.2% trips weekly. The effect is dynamic and persists in different buffer sizes. The resulting climate benefit from expanding 19 metro stations is around 0.17–0.24 or up to 0.34 MtCO₂-eq yearly which accounts for 0.21–0.3 or up to 0.42% of Thailand's GHG emissions from the transport sector. Other

climate benefits from the reduction of private vehicle use could be an order of magnitude larger. Even monetizing this very conservative lower bound of CO₂ emissions savings at a low carbon price of \$100/tCO₂ translates into about \$1 billion worth of carbon credits over 30 years. This suggests that financing public transit in developing countries, and in particular, in rapidly growing cities, is a key area for international climate finance, e.g., by the Green Climate Fund.

For completeness, we also note that transit access leads to better traffic flow, better air quality^{13,16}, and is a necessary component of social inclusion (though not sufficient for social cohesion⁴⁸). The results are beneficial directly to the Department of Land Transport and the Office of Transport and Traffic Policy and Planning of Thailand and indirectly to other urban land transport planning departments in assessing the impact of a transport policy intervention.

Data availability

Taxi GPS data used to analyze during the current study were retrieved from <https://iticfoundation.org>.

Code availability

The code of this work is accessible at the git repository here <https://github.com/ai4up/bangkok-metro>.

Received: 15 June 2024; Accepted: 5 December 2024;

Published online: 19 January 2025

References

- Emissions Trends and Drivers. in *Climate Change 2022 - Mitigation of Climate Change* (ed. IPCC) 215–294 (Cambridge University Press, 2023). <https://doi.org/10.1017/9781009157926.004>.
- Transport. in *Climate Change 2022 - Mitigation of Climate Change* (ed. IPCC) 1049–1160 (Cambridge University Press, 2023). <https://doi.org/10.1017/9781009157926.012>.
- Creutzig, F. Evolving Narratives of Low-Carbon Futures in Transportation. *Transp. Res. Rev.* **36**, 341–360 (2016).
- Ayaragarnchanakul, E. & Creutzig, F. Bangkok's locked-in traffic jam: Price congestion or regulate parking? *Case Stud. Transp. Policy* **10**, 365–378 (2022).
- Creutzig, F. & He, D. Climate change mitigation and co-benefits of feasible transport demand policies in Beijing. *Transp. Res. Part Transp. Environ.* **14**, 120–131 (2009).
- Quinet, E. A meta-analysis of Western European external costs estimates. *Transp. Res. Part Transp. Environ.* **9**, 465–476 (2004).
- Sovacool, B. K., Kim, J. & Yang, M. The hidden costs of energy and mobility: A global meta-analysis and research synthesis of electricity and transport externalities. *Energy Res. Soc. Sci.* **72**, 101885 (2021).
- Creutzig, F. et al. Urban infrastructure choices structure climate solutions. *Nat. Clim. Change* **6**, 1054–1056 (2016).
- Fageda, X. Do light rail systems reduce traffic externalities? Empirical evidence from mid-size European cities. *Transp. Res. Part Transp. Environ.* **92**, 102731 (2021).
- Brown, M., Southworth, F. & Sarzynski, A. *Shrinking the Carbon Footprint of Metropolitan America*. (Brookings Institution, 2008).
- Anderson, M. L. Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion. *Am. Econ. Rev.* **104**, 2763–2796 (2014).
- Adler, M. W., Liberini, F., Russo, A. & van Ommeren, J. N. The congestion relief benefit of public transit: evidence from Rome. *J. Econ. Geogr.* **21**, 397–431 (2021).
- Gallivan, F., Ang-Olson, J., Liban, C. B. & Kusumoto, A. Cost-Effective Approaches to Reduce Greenhouse Gas Emissions through Public Transportation in Los Angeles, California. *Transp. Res. Rec.* **2217**, 19–29 (2011).
- Zegras, C. The Built Environment and Motor Vehicle Ownership and Use: Evidence from Santiago de Chile. *Urban Stud.* **47**, 1793–1817 (2010).
- Bento, A. M., Cropper, M. L., Mobarak, A. M. & Vinha, K. The Effects of Urban Spatial Structure on Travel Demand in the United States. *Rev. Econ. Stat.* **87**, 466–478 (2005).
- Bailey, L., Mokhtarian, P. & Little, A. The Broader Connection Between Public Transportation, Energy Conservation and Greenhouse Gas Reduction (ICF International, 2008).
- Cervero, R. & Murakami, J. Effects of Built Environments on Vehicle Miles Traveled: Evidence from 370 US Urbanized Areas. *Environ. Plan. Econ. Space* **42**, 400–418 (2010).
- Chatman, D. G. Deconstructing development density: Quality, quantity and price effects on household non-work travel. *Transp. Res. Part Policy Pract.* **42**, 1008–1030 (2008).
- Yang, J., Chen, S., Qin, P. & Lu, F. *The Effects of Subway Expansion on Traffic Conditions: Evidence from Beijing*. <https://www.jstor.org/stable/resrep15032> (2015).
- Huang, X., Jason Cao, X., Yin, J. & Cao, X. Can metro transit reduce driving? Evidence from Xi'an, China. *Transp. Policy* **81**, 350–359 (2019).
- Xie, L. Automobile usage and urban rail transit expansion: evidence from a natural experiment in Beijing, China. *Environ. Dev. Econ.* **21**, 557–580 (2016).
- Gu, Y., Jiang, C., Zhang, J. & Zou, B. Subways and Road Congestion. *Am. Econ. J. Appl. Econ.* **13**, 83–115 (2021).
- Zhang, Y., Zheng, S., Sun, C. & Wang, R. Does subway proximity discourage automobility? Evidence from Beijing. *Transp. Res. Part Transp. Environ.* **52**, 506–517 (2017).
- Sun, G., Zhao, J., Webster, C. & Lin, H. New metro system and active travel: A natural experiment. *Environ. Int.* **138**, 105605 (2020).
- Deng, Y. & Zhao, P. The impact of new metro on travel behavior: Panel analysis using mobile phone data. *Transp. Res. Part Policy Pract.* **162**, 46–57 (2022).
- Du, Q., Huang, Y., Zhou, Y., Guo, X. & Bai, L. Impacts of a new urban rail transit line and its interactions with land use on the ridership of existing stations. *Cities* **141**, 104506 (2023).
- Salon, D. Neighborhoods, cars, and commuting in New York City: A discrete choice approach. *Transp. Res. Part Policy Pract.* **43**, 180–196 (2009).
- Kim, K. Exploring the difference between ridership patterns of subway and taxi: Case study in Seoul. *J. Transp. Geogr.* **66**, 213–223 (2018).
- Yue, M. et al. Understanding the interplay between bus, metro, and cab ridership dynamics in Shenzhen, China. *Trans. GIS* **22**, 855–871 (2018).
- Austin, D. & Zegras, P. C. Taxicabs as Public Transportation in Boston, Massachusetts. *Transp. Res. Rec.* **2277**, 65–74 (2012).
- Hochmair, H. H. Spatiotemporal Pattern Analysis of Taxi Trips in New York City. *Transp. Res. Rec.* **2542**, 45–56 (2016).
- Chen, C., Feng, T., Ding, C., Yu, B. & Yao, B. Examining the spatial-temporal relationship between urban built environment and taxi ridership: Results of a semi-parametric GWPR model. *J. Transp. Geogr.* **96**, 103172 (2021).
- Wang, R., Chen, F., Liu, X. & Fujiyama, T. Spatiotemporal Analysis of Competition Between Subways and Taxis Based on Multi-Source Data. *IEEE Access* **8**, 225792–225804 (2020).
- Jiang, S., Guan, W., He, Z. & Yang, L. Exploring the Intermodal Relationship between Taxi and Subway in Beijing, China. *J. Adv. Transp.* **2018**, 3981845 (2018).
- Zhang, H., Cui, Y. & Jia, J. Mining Multimodal Travel Mobilities with Big Ridership Data: Comparative Analysis of Subways and Taxis. *Sustainability* **16**, 4305 (2024).
- Li, M., Dong, L., Shen, Z., Lang, W. & Ye, X. Examining the Interaction of Taxi and Subway Ridership for Sustainable Urbanization. *Sustainability* **9**, 242 (2017).
- Office of Transport and Traffic Policy Planning. *Travel Demand Freight Movement Survey For National Transport Planning*. <https://www.otp.go.th/index.php/edureport/view?id=147> (2017).

38. Yoosuk, P. *Preliminary Evaluation of Greenhouse Gas Emission on MRT Green Line Project: Mo Chit – Saphan Mai – Khu Kgot Section* (Thammasat University, 2016).
39. Prasertsubpakij, D. & Nitivattananon, V. Evaluating accessibility to Bangkok Metro Systems using multi-dimensional criteria across user groups. *IATSS Res.* **36**, 56–65 (2012).
40. Ayaragamchanakul, E., Creutzig, F., Javaid, A. & Puttanapong, N. Choosing a Mode in Bangkok: Room for Shared Mobility? *Sustainability* **14**, 9127 (2022).
41. A. Verkehrswende. *Towards Decarbonising Transport - Thailand 2024*, https://www.agora-verkehrswende.de/fileadmin/Projekte/2024/Thailand/Factsheet_Thailand_lay4_final.pdf (2024).
42. Bangkok Metropolitan Administration. *The Bangkok Master Plan on Climate Change 2013–2023*, https://webportal.bangkok.go.th/upload/user/00000231/web_link/air/MP%20Full%20Report%20_Kiso%20clean%20-11.pdf (2012).
43. Department of Land Transport. Registered vehicles by fuel type. *Transport Statistics, Department of Land Transport*, <https://web.dlt.go.th/statistics/> (2017).
44. Metamedia Technology Co., Ltd. *Longdo Map API* (Metamedia Technology Co., Ltd., 2005).
45. Siangsuebchart, S., Ninsawat, S., Witayangkurn, A. & Pravinvongvuth, S. Public Transport GPS Probe and Rail Gate Data for Assessing the Pattern of Human Mobility in the Bangkok Metropolitan Region, Thailand. *Sustainability* **13**, 2178 (2021).
46. Vichiensan, V., Fukuda, A. & Malaitam, S. Perceived Walkability Evaluation in Transit Station Catchment Areas in Bangkok. In *2022 Moratuwa Engineering Research Conference (MERCon)* 1–6, <https://doi.org/10.1109/MERCon55799.2022.9906209> (2022).
47. Wooldridge, J. M. *Econometric Analysis of Cross Section and Panel Data*, 2nd ed. (The MIT Press, Cambridge, Mass, 2010).
48. Wickham, J. Public transport systems: the sinews of european urban citizenship? *Eur. Soc.* **8**, 3–26 (2006).

Acknowledgements

We thank the Royal Thai Government for providing E.A. with a Ph.D. scholarship; the Air Quality and Noise Management Bureau, Pollution Control Department of Thailand for authorizing weather condition data usage; and the Intelligent Traffic Information Center Foundation of Thailand for supplying open access traffic data. We also thank the reviewers for their constructive and helpful comments, which have significantly improved the quality and clarity of our manuscript. This research received no funding.

Author contributions

conceptualization, F.C., J.H., F.W., F.N., H.K., E.A.; methodology, F.C., J.H., F.W., F.N., H.K., E.A.; formal analysis, F.N., H.K., E.A.; writing—original draft preparation, E.A.; writing—review and editing, F.C., J.H., F.W., F.N., H.K., E.A.; supervision: F.C.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44333-024-00020-6>.

Correspondence and requests for materials should be addressed to Eva Ayaragamchanakul.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025