How can the expansion of the rapid transit system affect travel behavior? evidence from the megacity of São Paulo

Luiz Pedro Couto Santos Silva[#]

Abstract

This study investigates the effects of the expansion of a rapid transit network on the distribution of bilateral trip flows in São Paulo Metropolitan Region (SPMR) in a 10 year period. The empirical analysis has three steps: It starts with estimates of differences in travel time between public transit and private modes over the period. This is followed by an investigation on the relationship between the increase in the proximity to new rapid transit stations and the shift in transport mode. I use granular data that observes the improvement of active accessibility to public transit through the lens of catchment areas of BRT, train, and subway stations. Later, the investigation draws on a gravity-based model also using these refined measures of station catchment areas. The results show that although the expansion of the rapid transit network had failed in promoting gains of speed of travel relative to cars, rail modes (train and subway) had a decline of -3.9% in travel time when compared to cars. I found that only rail transit stations are correlated with transport mode shift. The study also concludes that better access to new rapid transit stations promoted more intensity of travel flows between the areas with new rail than to the areas with new BRT stations. These push and pull effects on travel flows promoted by the new rapid transit stations had shown a distance decay effect in which rail users are more sensible to the walking time among the origin, rapid station, and destination. These findings suggest that a combination of an increase of average speed and better physical access to transit stations would enhance the bilateral travel flows between SPMR regions by public transit, which should contribute to alleviate traffic congestions.

Keywords: Station catchment area, Gravity model, São Paulo Metropolitan Region, Mobility patterns

Area ENABER: Infra-estrutura, transporte, energia, mobilidade e educação.

1 Introduction

The reduction in travel costs due to emergence of motorized vehicles in the XX century significantly influenced the spatial structure of urban areas (Anas et al., 1998; Broox and Deneoux, 2022). Urban transport systems are effective in fostering proximity through agglomeration externalities as long as they connect the population and the opportunities. This connectivity shapes the patterns of spatial interaction and travel flows (Ahlfeldt et al., 2015; Duranton and Puga, 2020; Fujita, Krugman and Venables, 2001; Glaeser, 2010). However, the rapid urban demographic growth in Latin America in recent decades has challenged the capacity of its cities to provide adequate transport infrastructure, which incentivized a shift towards private vehicle use (Bryan et al., 2020; Pojani and Stead, 2018; Vasconcellos, 2005). The lack of coordination between land use and transport infrastructure policies exacerbates

---

[#] PhD candidate at the Federal University of Juiz de Fora.

congestion in Latin American cities, particularly in large urban areas where coordinating transport policies across multiple municipalities is more complex (Pojani and Stead, 2018).

Agglomeration and dispersion forces in cities have been assessed by a growing body of literature that uses spatially granular data in gravity-based models to estimate how exogenous shocks on urban structure (e.g., transport network, land markets) affect the spatial distribution of trip flows (Ahlfeldt et al., 2015; Ahlfeldt and Wendland, 2016; Ahrens and Lyons, 2021; Dingel and Titelnot, 2020; Tsivanidis, 2023). Another recent use of gravity-based models aims to estimate the effects of improvements in the physical access to the public transit system on the distribution of urban trip flows (Gaduh et al., 2022 ; Severen, 2023). Despite the recent advances in this literature, there is still a large venue for understanding how such agglomeration and dispersion forces are consequences of incentives for trip flows.

This study examines how the expansion of rapid transit systems influence the spatial organization of trip flows in cities. Specifically, it investigates how improvements in physical access to the public transit stations affect the spatial pattern of trip flows using the São Paulo metropolitan region as a case study. I assess if the expansion of this rapid transit system between 2007 and 2017 reduced the difference in travel times between public transit modes and cars. The achievement of this reduction would be an incentive for mode shift from the use of cars towards more use of public transit modes. Then, I draw on the literature of station catchment areas to assess if the increase in the catchment areas promoted by the new BRT and rail stations was related to the increase the proportion of public transit users by zone of São Paulo metro. I use refined geographic data on the street's network and on the location of population and opportunities to compute their connectivity with the rapid transit system through the station catchment areas of the BRT and rail. Finally, I also combined these refined measures of station catchment areas with spatial interaction models to assess if the increases in the connectivity with the public transit system affected the pattern of trip flows in São Paulo metro.

The metro area of São Paulo, one of the largest megacities in Latin America, has witnessed a sharp increase in the use of motorized private vehicles in recent decades (Bocarejo, 2020; Carvalho and Pereira, 2013; Vasconcellos, 2005). This has burdened the transport infrastructure and contributed to increase average commuting times from 37 minutes in 1992 to 45 minutes in 2012 (Carvalho and Pereira, 2013), undermining individual labor productivity by -2.7% for each 10 minutes spent on commuting (Haddad et al., 2015). Nevertheless, promoting changes in travel behavior to encourage a shift from private to public transit modes is a significant challenge because it evolves improving the travel speed and spatial connectivity of transit systems as well as people’s access to transit stations (Brooks and Deneoux, 2022; Bocarejo et al., 2020).

During the period (2007-2017), São Paulo metro has expanded its high-speed transit system through 65 new train, subway, and BRT stations to incentivize a mode shift from private vehicles to public transit. Despite that, the influence of transit stations on the propensity of individuals to travel by public transit mode becomes weaker as potential users need to face longer walks (El-geneidy et al., 2014; Vale, 2021; Kamruzzaman, et al., 2014), which also depends on the transit mode (e.g., bus, rail, subway) (Estupiñan and Rodriguez, 2008; Murray et al., 1998). These qualitative dimensions of travel conditions influence passenger comfort, given that walking time is inherent to travel by public transit (Vale, 2021). Therefore, the understanding of how individuals are propense to use public transit depends on the station catchment area.

This study contributes simultaneously to two distinct but related fields of knowledge. The use of the concept of station catchment areas (Estupiñán and Rodriguez, 2008; El-geneidy et al., 2014; Murray et al., 1998; Kamruzzaman et al., 2014) explicitly measures how the changes in densities of population and opportunities covered by the transit system shape the decisions of intraurban trips (Ahlfeldt et al., 2015; Gaduh et al., 2022; Severen, 2023). From the best of our knowledge, this is the first time that the concepts of spatial interaction models on urban scale and station catchment areas are combined. Furthermore, the study makes a deep assessment on how recent public transit policies affected travel behavior in the São Paulo metro area, which brings useful results to guide policy implications.

The remainder of the study is organized as follows. The next section presents a literature review on station catchment areas and gravity-based interaction models. Section 3 describes the study area, followed by the methods section that details the data and the econometric models used in the paper. Sections 5 and 6 present the results and the final remarks of the study.

2 Literature review

Some of the dimensions of the built environment, such as population density, street design, and land use mix, determine the attractiveness of the public transit system (Cervero and Kockelman, 1997; Handy et al., 2002), which is moderated by local socioeconomic conditions (Ewing and Cervero, 2001; 2010). Therefore, the choice of residence and work places is governed by the spatial distribution of urban amenities, including travel conditions, by which the efficiency of the transit network system becomes an incentive for proximity between urban agents.

The level of speed of travel and the comfort promoted through physical access to the public transport system are understood as mechanisms for transport mode shifts (Brooks and Deneoux, 2022; Gaduh, 2022; Estupiñán and Rodriguez, 2008; Murray et al., 1998). These aspects contribute to delimit the spatial range from transit stations in which potential riders are drawn, defined as *station catchment areas*. Therefore, the influence of public transit stations on the usage of the system involves spatial decay (El-geneidy et al., 2014; Kamruzzaman et al., 2014). The utility level that a transport mode provides to the passenger also explains travel behavior, as the choice of mode for travel is assumed to be governed by a trade-off between his choices and the other available modes and routes (McFadden, 1974). On an aggregated geographic level, forces of attraction at the origin and destination and the efficiency of the transport network in promoting their connectivity can explain the intensity of bilateral travel flow between the pair of distinct areas. This sort of bilateral travel flow has been the object of study of a long strand of literature related to spatial interaction models (Wilson, 1971; Haynes and Fotheringham, 1985; Roy and Thill, 2004).

Recent literature has collapsed the strategies for identifying travel behavior in spatial interaction models that predict the probability of interaction between pairs of blocks of a city (Ahlfeldt, 2015). This framework assumes that individuals simultaneously choose their household and work places based on the urban amenities nearby (e.g., green area at household site and productivity at work site) and their level of connectivity through the transport network. The individual's choice is also governed by the idiosyncratic component of utility when he aims to maximize his utility level given his constraints (e.g., monetary budget, time available for commuting), the demand and supply of housing at origin and for workers at the destination place. Under these assumptions, individuals will choose the unique combination of household and work places that will maximize their respective utility levels. If land and job markets clear, the spatial equilibrium under the multiple demand and supplies for city blocks determine the commuting flows between them. Therefore, changes in travel cost could induce the demands for city blocks, both for purposes of living and working.

For example, the empirical analysis of Ahlfeldt et al. (2015) concluded that the reunification of western and eastern Berlin reduced commuting costs and intensified commuting flows between these locations that were formerly separated. Their model also computes agglomeration economies through wages and rental prices, which they argue increased at some blocks that benefited from the reduction of commuting costs and improvements in production and residential externalities. Tsivanidis (2023) estimates the impacts of a BRT expansion in Bogotá on the demand for land and the car ownership decision with a gravity-commuting model that computes the simultaneous gains in the access of firms to workers and residences to jobs through the public transit network. This cumulative market access (CMA) enhances the agglomeration economies promoted by the new BRT network through impacts in land and wage prices in the city blocks treated: those that had increases in CMA over the period.

When successful in terms of reducing travel costs, the expansion of the transit system and the improvement of the access to the new rapid transit stations may affect the demand for travel using such transport modes. Under these assumptions, Severen (2021) investigated the effects of the proximity to new rail stations on the patterns of commuting flows in the Greater Los Angeles area. He assigns treatment by whether a census tracts had received a new rail station, and alternatively, within Euclidean distance of 250 or 500 meters of the tracts' centroid. The study of Gaduh et al. (2022) analyzed the effects of BRT network expansion on travel flows in Jakarta. The authors assigned the treatment by access to the transit system through Euclidean distances of 1 km between the borders of subdistricts (*kecamatan*) and the BRT stations. Gaduh (2022) found no significant impact of better access to BRT stations on the probability of bilateral travel flows. This result is explained by the lack of gains of relative travel speed through the BRT network when compared to private vehicles, given the insufficient investment in infrastructure from this policy. However, these simple distance measures of Severen (2021) and Gaduh (2022) do not account for the streets' design and overestimate the real conditions required to reach the transit network system (El-geneidy et al., 2014).

Numerous empirical studies have shown the relationship between the catchment area of the transit system and the walking distance to its stations. The longer is the walking distance to the transit station, the lower the percentage of residents that use it (El-geneidy et al., 2014; García-Palomares et al., 2018; Murray et al., 1998). Because of this relationship, the circuity factor (a measurement of the amount of a street network that fits into a radius format) and the level of connectivity of the sidewalk network are built environment dimensions that influence the catchment of each transit station (Hsio et al., 1997; O'sulliavan et al., 1997; Kamruzzaman et al., 2014). The walking distance that individuals tolerate to use the transit system tends to be longer for rail users and shorter for bus users (Burke and Brown, 2007; Daniels and Mulley, 2013; O'sullivan et al., 1997; El-geneidy et al., 2014). This spatial heterogeneity in the influence of the station catchment area is also a consequence of the relationship between the speed of travel offered by different public transit modes and the incentive to use the public transit system.

When the distances from the origin and destination to a transit station are short enough to encourage the use of the public transport system, an incentive is created that prevents the use of private motorized transport. However, transport infrastructure is typically restricted to serve only part of populations in cities of developing countries because their institutional development rarely had the capacity to follow the demand derived from urbanization growth (Bryan et al., 2020; Pojani and Stead, 2018). Considering the relevance of land use within the built environment dimensions that influence travel behavior, an increase in population density near transit systems can improve regional accessibility (Pojani and Stead, 2018; Venter et al., 2019).

3 The recent expansion of rapid transit in São Paulo

With 22 million inhabitants, the São Paulo metropolitan region is one of the largest urban agglomerations in the world, and it concentrates approximately 19% of the Brazilian GDP (IBGE, 2021). The rapid population growth of São Paulo has increased the economic pressure in the residential market, has burdened its population with housing and transportation costs, and led this region to reach the second highest living cost in Brazil (Acolin and Green, 2017; Almeida and Azzoni, 2016). Figures 1 and 2 show that although the population is significantly dispersed throughout the metro area, there is a high spatial concentration of jobs near the city center. The spatial dispersion of population towards peripheral areas of SPMR challenges the provision of transport infrastructure for commuting, and inhabitants of its peripheral areas have significantly lower levels of accessibility to job opportunities (Boisjoly et al., 2020; Gianotti et al., 2021; Vieira and Haddad, 2015).

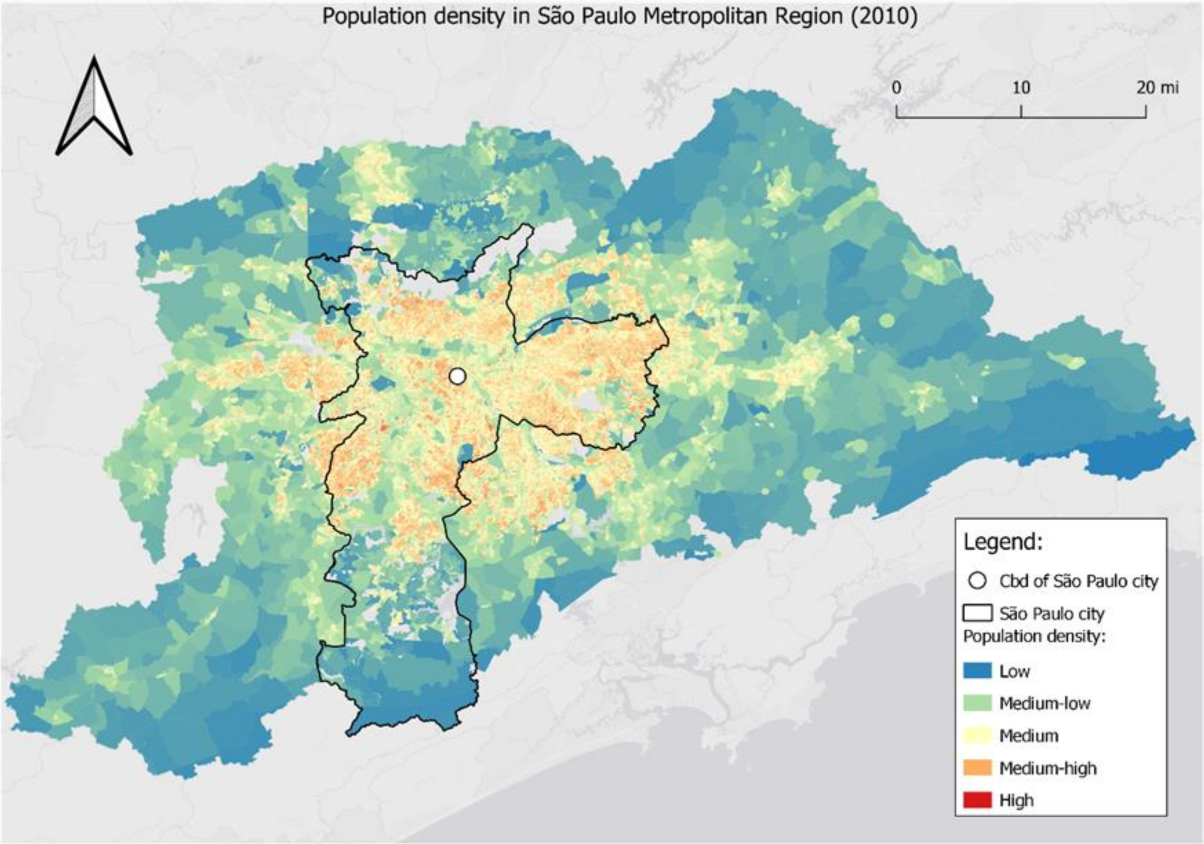

Figure 1 – Population density in São Paulo Metropolitan Region (2010)
Source: author’s own, from Brazilian census of 2010.

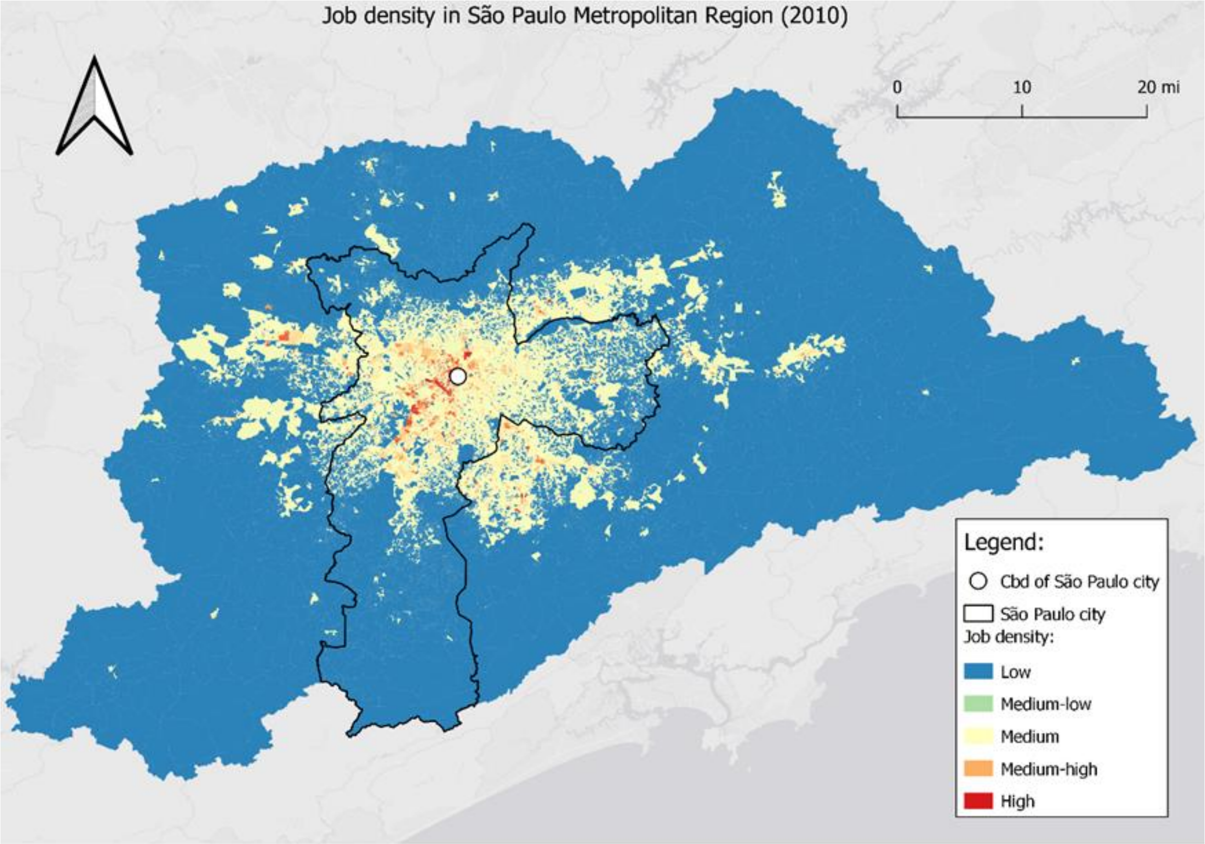

Figure 2 – Job density in São Paulo Metropolitan Region (2010)
Source: author's own, from the Annual Social Information Report of 2010.

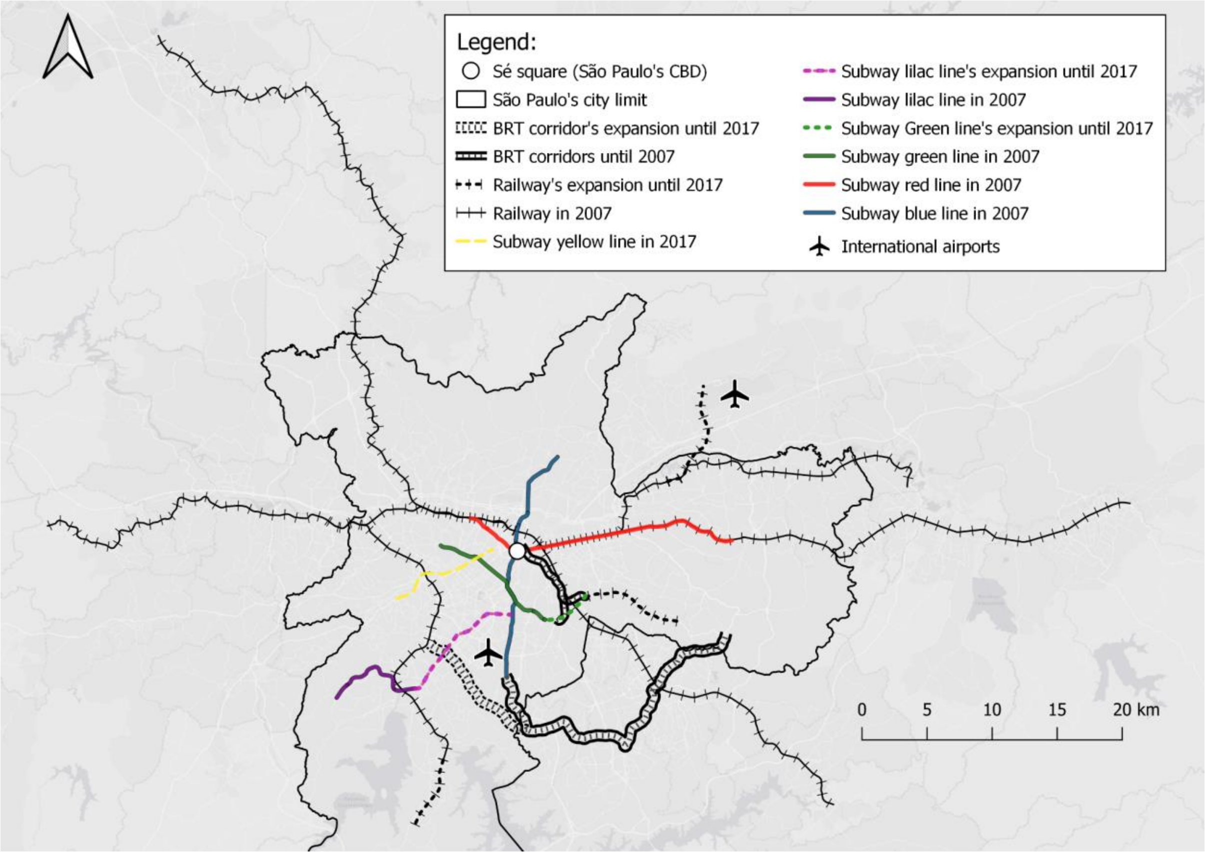

Figure 3 – Expansion of the rapid transit system of the metropolitan area of São Paulo between 2007 and 2017. Source: Author’s own elaboration, from OSM and IDTP data.

In recent years, there has been a substantial expansion of public transit infrastructure in São Paulo, which opened 65 new rapid transit stations between 2007 and 2017. Figure 3 shows that by 2017, the subway system expanded from the central to the western region of São Paulo city on the Yellow line, towards the east with the Green line, and from the southwestern region to the central-southern zone through the Purple line. In the same period, the rail network was expanded towards the east and south regions and from São Paulo’s northeast to the city of Guarulhos. Figure 3 also shows that the BRT network was expanded by 26 Km with new stations opened between the south and the southwest regions of the city of São Paulo, and a new corridor opened in the city of Guarulhos.

This expansion of the rapid transit system of São Paulo has substantially expanded people’s access to opportunities (ref), especially after the city government of São Paulo launched the “single fare” scheme in 2005 (Rolnik and Klintowitz, 2011). This new fare rule allowed passengers to use the whole public transport system and to transfer between modes (buses, subway and rails) paying a single fare and which was cheaper than the fares that one would pay to take each public transport mode singly (São Paulo, 2008; Sptrans, 2013). In 2014, this single fare rule was expanded to the remaining municipalities of the metropolitan region, along with a monthly plan fare (Santiago, 2013). However, even after the expansion of the rapid transit network between 2007 and 2017, many areas in the SPMR still lack quick access to the rapid transit system. According to Mobilidados (2023), only 12% of the total SPMR population lived within 1 kilometer of a rapid transit station in 2017. This percentage increases to 31% for those in the highest income quintile and decreases to 10% in the lowest income quintile.

4 Data and methods

4.1 Data

This study uses repeated cross-sectional data covering travel behavior, population characteristics and public transit station networks. Information on observed trips and the characteristics of individuals and their home and work locations are drawn from the 2007 and 2017 household travel surveys (Origin-Destination) conducted in the São Paulo metro area (Metro, 2007; 2017). These surveys are sampled across 423 geographic zones in 2007 and 470 zones in 2017, and the expanded sample represents 29 and 31 thousand households, respectively. To allow for the spatial comparability of the results for both years, most analyses conducted in this paper are aggregated in the 134 districts of the metropolitan region, displayed in Figure 1. Information on respondents' household and workplace locations are available in point coordinates. The household travel survey was used to quantify the number and duration of commuting trips between districts by transportation mode.

Data on rapid transit stations is from the MOBILIDADOS data portal[1]. This data includes information on rail and BRT transit stations, including spatial coordinates and dates

---

[1] https://mobilidados.org.br/rms/rmsp.

of inauguration. Data from OpenStreetMap (OSM) was used to obtain information on the street network of São Paulo. for the year 2019, which requires the assumption that the street network has not changed significantly between 2007 and 2019.

The OSM street network was used to calculate the shortest route by walking from the location of each household to every rapid transit station using the r5r package in R (Pereira et al., 2022). Average walking speed is assumed to be 3.6 km/h, following Fitzpatric and Brewer (2006). I computed for 2007 and 2017 the total number of jobs and residential population within the catchment areas of rapid transit stations considering different walking time thresholds to reach the stations of (10, 20 and 30 minutes) . This allowed us to calculate how the number of people and jobs that were covered by the rapid transit system in each district changed between the two household surveys of 2007 and 2017. Therefore, the catchment areas tend to have smaller areas than the districts.

The descriptive statistics for the household travel surveys data by transport mode are presented in Table 1. It shows that the mean commuting times decreased for bus, car and rail users, as well as the mean commuting distances. The mean Euclidean distance from the household location to the closest BRT and rail stations also decreased for individuals who commute by transportation modes (except for car users), which reflects how significant the expansion of the rapid transit system has been.

Table 1 – Descriptive statistics of trips and travelers between districts for all purposes. São Paulo metropolitan region, 2017 and 2017.

Transportation mode	**Bus**				**Car**				**Rail**			
Year	**2007**		**2017**		**2007**		**2017**		**2007**		**2017**	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Travel time (Minutes)	53	31	44	25	30	26	26	21	47	28	43	22
Euclidean Distance (meters)	6,356	5,609	5,201	4,508	5,551	6,241	5,438	6,278	7,746	6,735	6,727	5,556
Euclidean Distance to the nearest BRT station (meters)	10,669	9,055	8,399	8,140	7,811	7,122	7,198	7,456	6,214	6,233	4,657	4,421
Euclidean Distance to the nearest Rail station (meters)	4,848	5,144	4,752	5,068	3,215	3,812	3,769	4,521	1,922	2,901	1,460	1,827
Household Income (R$)	2,534	2,254	3,946	3,368	5,777	4,575	7,833	6,863	3,738	3,127	6,293	5,378
Male	0.44	0.5	0.41	0.49	0.54	0.5	0.54	0.5	0.5	0.5	0.51	0.5
Age	36	17	39	19	40	18	43	19	37	17	40	17
Cars per household	0.6	0.74	0.55	0.65	1.6	0.93	1.4	0.77	0.78	0.82	0.72	0.73
Number of trips	25,343		20,837		67,578		53,551		5,407		4,916	
% of total trips	21.19		25.77		68,73		54.46		5.5		5	

Source: Author's own elaboration, from the OD surveys of SPMR of 2007 and 2017.
Notes: Individuals who used more than one transport mode are not considered on these statistics. Household income information is in nominal values. Euclidean distance is the total linear distance from the origin to the destination of the trip.

4.2 Econometric models

In this paper, I used three different but related econometric models to examine the extent to which the expansion of the rapid transit system has (1) affected the travel time differences between transit and cars by zone, (2) increased the share of transit trips by zone, and (3) affected the spatial pattern of trip flows through the probability of bilateral flows between districts.

Model 1) Linear model to estimate differences in travel times between transportation modes

The aim with this first model is to understand whether the expansion of São Paulo's rapid transit network has contributed to making its transit system more attractive relative to cars. Following Gaduh et al. (2022), I used OLS models to estimate how the differences in travel times between different transit modes and cars changed over the period of the analysis (2007-2017). The Linear regression to be estimated is:

$$LogTime_i = \beta_0 + \beta_1 PT_i + \beta_2 X_i + \varepsilon_{ij} \qquad (1)$$

Where *Time* is the travel time for individual *i*. The sample is restricted for public transport and car trips, and *PT* is a dummy variable for individuals who used a public transportation mode. The analysis used different regression models to compare *PT* with cars. In each of these regressions, *PT* is represented by bus (all types) or rail (subway and train). Another regression compares the aggregated public transit modes with cars. *X* is a set of control variables: log of Euclidean distance of the trip, hour of departure time, weekday, origin zone, destination zone, travel purpose, and an interaction dummy between hour and weekday. Moreover, ε_{ij} is an error term. I also followed the framework of Gaduh et al., (2022) to compare the travel time of *potential BRT* users with other modes: bus, rail and cars. It consists of assigning a dummy variable in equation (1) as BRT user for individuals who commuted by bus and household place located within a radius distance of 1 km from a BRT station.

Model 2) Linear models to estimate transit mode share

In the second model, I investigate the extent to which the opening of new rapid transit stations and consequent increase in the share of the population and opportunities (number of for work and study) locating within the catchment area of the rapid transit system affected the share of trips by public transit in each zone between 2007 and 2017. Considering catchment areas of different sizes (5, 10, 15 and 20 minutes), I run a cross-sectional linear regression model with the following specification:

$$\Delta PT_i = \beta_0 + \beta_1 \Delta\%Covered_i + \beta_2 \Delta\%Covered_j + \beta_3 \Delta X_i + \varepsilon_i \qquad (2)$$

where ΔPT_i is the difference in the percentage of public transit users between 2007 and 2017 in the OD zone i. $\Delta\%Covered_i$ and $\Delta\%Covered_j$ are respectively the differences between 2007 and 2017 of the percentage of population and opportunities that fall within the walking catchment area of rapid transit stations in OD zones i and j respectively. I conducted distinct analyzes considering as the dependent variable the share of trips by all public transit modes, and the share of trips by bus and by rail (train and subway), separately. In the latter case, the types of rapid stations considered in $\Delta\%Covered_i$ and $\Delta\%Covered_j$ were selected accordingly. I only considered users who used a single transport mode to conclude their trips. ΔX_i is a set of control variables at the zone level, which includes the changes in the zone averages of: household income, travel time, number of cars per household, number of motorcycles per household, and population age. Finally, ε_i is an error term. I expect that a positive value in β_1 or β_2 indicate positive correlation between the ease of access to the public transit system and transport mode shifts.

Model 3) Gravity model to estimate incentives and disincentives for travel by public transit

Finally, I used a spatial interaction model to estimate how the spatial organization of trips between origins and destinations might have been affected by the opening of new rapid transit stations, considering the expansion in the share of people and opportunities within the catchment area of rapid transit stations. The model follows the framework proposed by Ahlfeldt et al. (2015) and Gaduh et al. (2022) to estimate spatial interaction between places based on economic incentives. It assumes that the observed quantity of bilateral trip flows from home to the opportunity place (with the purposes to work or study) reflects a spatial equilibrium determined by demands and supplies of amenities located at the origins and destinations of trips. The probability of bilateral trips drawn from this equilibrium is:

$$\pi_{ij} = \frac{W_{ij}}{\sum_{k=1}^{n} W_{ik}} \tag{3}$$

where π_{ij} is the probability of interaction between the district of origin *i* and destination *j*, is equal to the number of residents *W* in *i* who traveled to *j* among the *k* possible destinations.

The probability π is balanced by push and pull factors, such as the quantity and quality of opportunities in *i* and *j*, which determine the gravity decay between origins and destinations. The transport network also shapes these gravity forces by determining the connectivity and travel cost between *i* and *j*, in which the travel cost reduces the utility level achievable through the interaction between *i* and *j*. To explore the dimensions of the (dis)incentives for bilateral travels, I extend the framework of Ahlfeldt et al. (2015) and Gaduh et al. (2022) by using transit station catchment areas as a measure of access to the transit system. The spatial interaction model is implemented using the following log-linear Poisson regression:

$$log\pi_{ijt} = \beta_0 + \beta_1(\%Covered_{it} * \%Covered_{jt}) + \beta_2 Time_{ijt} + \delta_i + \gamma_j + \varphi_{it} + \upsilon_{jt} + T_t + \varepsilon_{ij}$$

(4)

where π_{ijt} is the probability of a trip between origin district *i* and destination district *j* at year *t*, where the log of this probability is explained by *%Covered*, the share of population and destinations within the catchment area of transit stations in that year at origin and destination. Thus, the interaction between $\%Covered_{it}$ (origin) and $\%Covered_{jt}$ (destination) is a continuous variable equalizing the product of the coverage rations at the origin and destination over the years 2007 and 2017. $Time_{ijt}$ is the average travel time by public transit between districts i and j *in year* t. Moreover, δ_i and γ_j are fixed effects for the origin and destination districts, respectively, and φ_{it}, υ_{jt}, and T_t are time-origin, time-destination and time-year fixed effects, respectively, and ε is an error term. Therefore, because the share of households and opportunies (for work or study) covered by the catchment area of rapid transit station vary over the period (2007-20017), β_1 in equation (4) estimates how the increase in population and opportunities within the catchment areas affected the spatial distribution of bilateral trip flows.

In this model, the sample is restricted to trips to work and study by public transit. Additionally, due to data sparsity, in this regression model I used the data for trips aggregated the district levels as origins and destinations. I treated sample noise by only considering those bilateral trips that had at least 10 occurrences in the raw sample. This is a common procedure in the literature (Ahfeldt et al., 2015; Ahlfeldt and Wendland, 2016; Gaduh et al., 2022), given that small sample representing bilateral trips may reduce the precision of the estimates (Dingel and Titelnot, 2020). To understand the heterogeneous effects of new rapid transit stations on travel flows according to transit mode, this study conducts empirical analyses using different regression models for rail and BRT station catchment areas.

This model 4 is similar to the one used by Gaduh et al., (2022), but it differs in how access to the rapid transit system is calculated. Gaduh et al., (2022) used the distance from the borders of each district's polygon to each BRT station as a proxy of access to the transit network. The measure of access to transit used here has a few advantages: First, I use the latitude and longitude coordinates of households and job or study locations, which provides more geographically precise information about the starting and ending points of trips; Second, I count the number of people and jobs or study destinations within the catchment area of transit stations based on walking times along the road network, which is more precise and realistic than Euclidean distances because it captures the influence of urban form on walking access to the transit system. Moreover, this approach allows me to conduct sensitivity analysis considering different sizes of the catchment areas given varying walking times. Finally, equation (4) estimates both the incentives through the ease of access to the transit system and the disincentives for longer trips, which allows me to observe the two sides of the coin of travel through public transport (Vale, 2021).

5 Results

5.1 Relative differences in travel time between transportation modes

The results of the OLS regressions testing the differences in travel times between public transit and cars are reported in Table 2. The coefficients in columns (1-2) show that the mean travel time by bus was 57% higher than by car in 2007 and increased to 59% higher in 2017, indicating that bus trips became slightly slower by 1.7 percentage points compared to cars. Conversely, the relative difference in travel time between rail (subway and train) and cars shown in columns (3-4) decreased by -3.9 percentage points between 2007 to 2017, suggesting some improvement in rail service relative to cars in São Paulo.

The results of columns (5-6) of table 2 indicate that the aggregated travel times for public transit (bus and rail) relative to cars improved marginally. The better performance of the rail transport system shown in columns (3-4) is expected due to the significantly higher investment in rail infrastructure compared to bus infrastructure in the study area and because of how rail services are immune to road traffic congestion. Further investigation of the performance of BRT lines compared to cars is presented in columns (7-8) Table 2. It is assumed that bus travels starting and ending within 1 km of BRT stations were made using the BRT mode. Comparing BRT with the rail system (train and subway), they show a travel time difference of 25% between BRT and rail users in 2007, along with a significant reduction of -9.4 by the year 2017.

Table 2 – OLS models for the estimates of the differences in travel times between public and private transport.

Modes compared	Bus (including BRTs) vs Car		Rail vs Car		Transit (all modes) vs Car		BRT vs Car	
Year	2007	2017	2007	2017	2007	2017	2007	2017
Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	0.57***	0.59***	0.32***	0.28***	0.54***	0.51***	0.53***	0.58***
	(0.012)	(0.005)	(0.021)	(0.018)	(0.010)	(0.005)	(0.029)	(0.017)
Adjusted R^2	0.59	0.78	0.58	0.78	0.59	0.79	0.58	0.78
Sample (N)	73,305	74,391	53,392	58,468	106,178	100,672	49,396	55,463

Author's own elaboration.

Notes: This table reports linear regression models that have the log of individual travel time as the dependent variable. Each model has a dummy variable indicating whether each trip observation was made by a public transport mode, compared with trips made by car. BRT (potential) users were defined as the individuals whose the travel by bus started within a distance of 1 kilometer of a BRT station. The additional controls in the regressions are: log of Euclidean distance and dummies of hour time, week day, interaction dummy of hour and week day, origin zone, destination zone, and travel purpose. Robust standard errors are clustered by an interaction of origin and destination zones and reported in parentheses. * / ** / *** denotes significant at the 10% / 5% / 1%, respectively.

5.2 Results for mode share

The results in Table 4 indicate that there is a positive and statistically relationship between the proportion of households within the catchment area of transit stations at the use of rail to travel to work or study. Nonetheless, our results show that the magnitude of this association becomes smaller as I consider larger catchment areas. When considering a catchment area of 5- minutes walking, each unitary increase in the difference between the percentage of households that could reach a heavy rail station within 5 minutes of walking increased the percentage of rail transit users by 0.46. The models also show that this relationship reduces to 0.19 when considering the walking time threshold of 10 minutes. By flexibilizing the walking distance threshold to 15 minutes, the model indicates that both the walking distance from the households to a station and from the station to the destination place (to work or study) are related to the increase in mode share for rail transit. When considering the threshold of 20 minutes of walk, the household place doesn't show to have any further relationship with the mode share for rail. However, the relationship between the walking time from the station to the destination place shows to be significant.

Table 4 - Results of OLS models for the effects of new rail stations on the share of trips made by rail.

Walking time threshold of catchment areas (in minutes)	5	10	15	20
Dependent Variable:	Delta percentage of travels made by rail, subway or train			
Constant	-0.271	-0.216	-0.249	-0.260
	(0.285)	(0.270)	(0.237)	(0.221)
delta %houses covered	0.462***	0.198***	0.078**	-0.042
	(0.121)	(0.048)	(0.028)	(0.031)
delta%opportunities covered	0.499	0.008	0.124*	0.241***
	(0.369)	(0.106)	(0.052)	(0.053)
delta travel time	-0.027	-0.013	0.011	0.027
	(0.025)	(0.023)	(0.021)	(0.020)
N	423	423	423	423
R2	0.17	0.29	0.39	0.42
Adj. R2	0.16	0.27	0.38	0.41

Source: Author's own elaboration.

Notes: Dependent variable is the difference in the percentage of trips made by rail between 2017 and 2007. Additional control variables are average income, average number of cars by household, and average number of motorcycles by household. Unit of analysis is OD zones. Standard errors are clustered by zones and reported in parentheses. * / ** / *** denotes significant at the 10% / 5% / 1%, respectively.

The analysis for BRT stations is shown in Table 5. The increase in the coverage of percentage of population and destination only had a strong association to changes in share of bus trips when considering catchment areas of 10 minutes from the station to the workplace. Although it also shows a negative relationship with the walking distance from the household place and the BRT stations, the statistical significance is 0.06, which suggests a negative weak relationship.These results indicate that the easy access to BRT stations were weak in promoting mode shift towards more usage of buses in São Paulo. This set of results for transport mode share are in line with the former analysis about the changes in relative travel time between trips by transit and cars. Given that only the rail modes were consistent in reducing the difference in travel time when compared to cars, this gain of speed seems to have become an incentive for mode shift to rail transit.

Table 5 - Results of OLS models for the effects of new BRT stations on the share of trips made by bus.

Walking time threshold of catchment areas (in minutes)	5	10	15	20
Dependent Variable:	Delta percentage of travels made by subway of bus			
Constant	1.06**	1.01**	1.06	1.059
	(0.448)	(0.445)	(0.445)	(0.444)
delta %houses covered	0.104	-0.129*	-0.104	-0.008
	(0.074)	(0.0703)	(0.057)	(0.057)
delta%opportunities covered	0.133	0.467**	0.072	0.067
	(0.446)	(0.234)	(0.145)	(0.113)
delta travel time	0.061	0.058	0.061	0.066
	(0.056)	(0.056)	(0.056)	(0.057)
N	423	423	423	423
R2	0.04	0.04	0.04	0.04
Adj. R2	0.02	0.03	0.02	0.03

Source: Author’s own elaboration. Notes: Dependent variable is the difference in the percentage of travels made by bus between 2017 and 2007. Additional control variables are average income, average quantity of cars by household, and average quantity of motorcycles by household. Unit areas are OD zones. Standard errors are clustered by Zones. * / ** / *** denotes significant at the 10% / 5% / 1%, respectively.

5.3 Results for gravity travel in the SPMR

The results of the spatial interaction models are presented next considering different sizes of the catchment area of BRT (Table 6) and rail stations (Table 7). Table 6 reports show that a unitary increase in the share of residences or opportunities within 5 minutes of a BRT station increases the probability trips between the pair of districts by 5.6%. The magnitude of this relationship decreases to 0.83% with a walking time threshold of 10 minutes, and further to 0.25% for a threshold of 15 minutes. The relationship between access to transit stations and travel flows diminishes as larger walking time thresholds are adopted in the models of Table 6. This is expected since increased walking time to reach a station raises disutility, acting as a disincentive to use the public transit system. As reported in Table 6, the travel time between origins and destinations is negatively associated with the probability of travel flows between pairs of districts. According to the specification in equation 4, these coefficients for the travel time variable can be interpreted as semi-elasticities, which indicate that additional each minute of travel between a pair of districts reduces the probability of commuting tripos between them by -1.4%.

These elasticities are lower magnitude than the results found by Ahfeldt et al. (2015), who reported -7% for Berlin, and by Gaduh et al. (2022), who reported between -6% and 13% for Jakarta. This difference might occur due to the larger territory of São Paulo[2].

Table 6 - Gravity models with station catchment areas for BRT stations

Dependent Variable:	Log $\boldsymbol{\pi}$		
Walk time to/from nearest station	5 minutes	10 minutes	15 minutes
%ResCov BRT x OppCov BRT	0.056*** (0.015)	0.008*** (0.001)	0.002*** (0.0004)
Log of travel time	-0.014*** (0.001)	-0.014*** (0.001)	-0.014*** (0.001)
Observations	3,508	3,508	3,508
Squared Cor.	0.46	0.47	0.47
Pseudo R2	0.072	0.073	0.073
BIC	7,696	7,695	7,695

Source: Author's own elaboration.

Notes: "%ResCov" and "%OppCov" are the percentages of residences and the percentage of Destinations (jobs or workplaces) within the station catchment areas, respectively. Additional variables are fixed effects for: origin, destination, origin and year, destination and year, and year. Standard errors are clustered by origin and destination in parenthesis. * / ** / *** denotes significant at the 10% / 5% / 1%, respectively.

The results of models for the effects of rail stations (subway and trains) on the travel flows through the rapid transit system are reported in Table 7. They show that the opening of rail stations has a higher effect than BRT stations when considering catchment areas of 5 minutes. Here, for each unitary increase in the share of residences or opportunities, the probability of a trip between pairs of districts increases by 6%.

Table 7 – Gravity models with station catchment areas for Rail stations

Dependent Variable:	Log $\boldsymbol{\pi}$		
Walk time to/from nearest station	5 minutes	10 minutes	15 minutes
%ResCovRail x %OppCov Rail	0.060*** (0.013)	0.004*** (0.0009)	0.001*** (0.0002)
Log travel time	-0.014*** (0.001)	-0.013*** (0.001)	-0.013*** (0.001)
Observations	3,508	3,508	3,508
Squared Cor.	0.466	0.469	0.474

---

[2] While Berlin has as area of 891.3 km² and Jakarta of 661.5 km² , the metro region of Sao Paulo has an area of 7,946 km². This large area of the São Paulo results in longer trip distances on average, and might explain the larger propensity of traveling further distances than in Berlin and Jakarta.

Pseudo R2	0.072	0.073	0.074
BIC	7,696	7,695	7,694

Source: Author’s own elaboration.
Notes: “%ResCov” and “%OppCov” are the percentages of residences and the percentage of Destinations (jobs or workplaces) within the station catchment areas, respectively. Additional variables are fixed effects for: origin, destination, origin and year, destination and year, and year. Standard errors are clustered by origin and destination in parenthesis. * / ** / *** denotes significant at the 10% / 5% / 1%, respectively.

However, the marginal effect of walking time to a rail station on the probability of travel is stronger than in the BRT case. Figure 4 shows the change in magnitude of effects by considering catchment areas of different sizes around rail and BRT stations. The opening of new rail stations exert stronger positive influence on the probability of trip flows than BRT stations as the share of households or opportunities increase when the walking time is shorter.

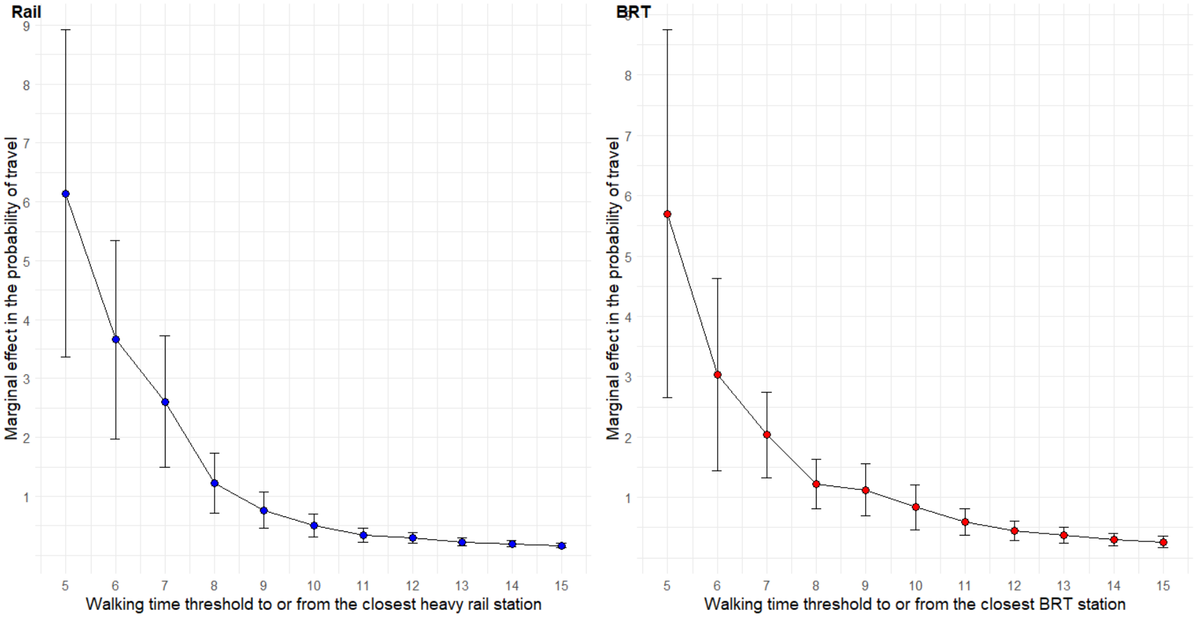

Figure 4 - Marginal effects of share of population and opportunties to rapid transit station by walking minute.

Source: Author’s own elaboration. Notes: The figure presents the coefficient results of 22 different regressions following equation 4. Coefficients are multiplied by 100 to ease interpretation. Each regression considers a different walking time threshold for the refined station catchment variable. Vertical bars represent the confidence intervals from each regression. Standard errors are clustered by origin and destination districts.

Trip flows by public transit are slightly more influenced by rail than by BRT stations until the walking time threshold of 8 minutes. From the 9 walking minutes threshold, the rail stations show a stronger effect than BRT stations on their capacity of incentivizing the public transit trips within SPMR. Although small in magnitude, these differences between rapid transit modes are aligned with the literature on station catchment areas and with what the data shows about the users of the transit system and the differences in travel time between transit modes and cars. Since rail services are faster and more reliable than road-based transit modes, passengers are more sensitive to the distance from a rail station than a BRT station when they decide to use the transit system or not. Also, given that the income level of those who used the rail system was between 47% (in 2007) and 59% (in 2017) higher than those who traveled by bus, it is also expected that the rail transit users are less likely to walk longer distances to reach the transit system than BRT users.

Regarding the specifications of the spatial interaction models used in this study, they show some very important features to control for unobserved effects and sample noise. The first is the use of 5 fixed effects at district and time level. In particular, the interactions between origin & year and destination & year aim to control for changes such as population or job increases between 2007 and 2017, which may have occurred because of the new zoning rules in São Paulo city (see section 3).

6 Final remarks

This study aims to understand the effects of changes in the rapid transit network on travel behavior in the São Paulo Metropolitan Region (SPMR) between 2007 and 2017. The empirical approach is grounded in the literatures on spatial interaction models and station catchment areas, with the objective of observing changes in incentives and disincentives for travel via public transportation within the São Paulo metropolitan area using multi year data with high spatial resolution.

I found that there were no substantive changes in the relative difference in travel times between buses and cars between 2007 and 2017, although there was a -3.9% decrease in travel times for the rail system (trains and subways) within the SPMR. Although 26 km of new BRT corridors were added in the RMSP over the 10-year period, this new infrastructure does not yet appear to have generated significant speed gains to compete effectively with car travel.

I found a positive relationship between the increase in the catchment areas of the rail system and the mode shift to rail transit, but a much weaker relationship for BRT stations. These results can be explained through the incentive promoted by the gain of speed due to the expansion of the rail system. The analysis of travel flows suggests that it became more intense in areas with access to the rail than those to the BRT system within the total walking distances of 16 minutes from the house place and to the destination place, which may relate to higher rail speed. The relationship between travel flows and walking times to and from the stations aligns with the literature on station catchment areas, indicating a stronger decay in the attraction to make a travel by rail compared to BRT.

The study brings an approach that decomposes the incentives for travel by public transit into the ease of reaching a rapid station and the travel time, with the aims to deepen the understanding of how different components of public transit travel can incentivize individuals to use the system. This combines insights from urban planning and urban economics to further understand what encourages people to use public transit. It concludes that travel time is only one of many factors influencing this complex decision. Future research could use these findings to guide public policies on where to build new transit stations, using hypothetical scenarios to predict how different locations of the stations might affect public transit use. Another possibility for future agenda is to investigate how the expansion in the rapid transit system of São Paulo metro affected the spatial distribution of population and opportunities.

References

ALMEIDA, A. N.; AZZONI, C. R. Custo de vida comparativo das regiões metropolitanas brasileiras: 1996-2014. **Estudos Econômicos (São Paulo)**, v. 46, p. 253-276, 2016.

AHLFELDT, G. M. et al. The economics of density: Evidence from the Berlin Wall. **Econometrica**, v. 83, n. 6, p. 2127–2189, 2015.

AHLFELDT, G. M.; WENDLAND, N. The spatial decay in commuting probabilities: Employment potential vs. commuting gravity. **Economics Letters**, v. 143, p. 125–129, 2016.

AHRENS, A.; LYONS, S. Do rising rents lead to longer commutes? A gravity model of commuting flows in Ireland. **Urban Studies**, v. 58, n. 2, p. 264–279, 2021.

ANAS, A.; ARNOTT, R.; SMALL, K. A. Urban spatial structure. **Journal of economic literature**, v. 36, n. 3, p. 1426–1464, 1998.

BAGLEY, M. N.; MOKHTARIAN, P. L. The impact of residential neighborhood type on travel behavior: A structural equations modeling approach. **The Annals of regional science**, v. 36, n. 2, p. 279–297, 2002.

BOISJOLY, G. et al. Accessibility measurements in São Paulo, Rio de Janeiro, Curitiba and Recife, Brazil. **Journal of Transport Geography**, v. 82, p. 102551, 2020.

BOWES, D. R.; IHLANFELDT, K. R. Identifying the impacts of rail transit stations on residential property values. **Journal of urban Economics**, v. 50, n. 1, p. 1–25, 2001.

BROOKS, L.; DENOEUX, G. What if you build it and they don't come? How the ghost of transit past haunts transit present. **Regional Science and Urban Economics**, v. 94, p. 103671, 2022.

CARVALHO, C. H. R. DE; PEREIRA, R. H. M. Indicadores de mobilidade urbana da PNAD 2012. 2013.

CERVERO, R.; KANG, C. D. Bus rapid transit impacts on land uses and land values in Seoul, Korea. **Transport policy**, v. 18, n. 1, p. 102–116, 2011.

CERVERO, R.; KOCKELMAN, K. Travel demand and the 3Ds: Density, diversity, and design. **Transportation research part D: Transport and environment**, v. 2, n. 3, p. 199–219, 1997.

SEVEREN, Christopher. Commuting, labor, and housing market effects of mass transportation: Welfare and identification. **Review of Economics and Statistics, v**. 105, n. 5, p. 1073-1091, 2023.

DINGEL, J; TINTELNOT, F. **Spatial economics for granular settings**. National Bureau of Economic Research, 2020.

EL-GENEIDY, A. et al. New evidence on walking distances to transit stops: Identifying redundancies and gaps using variable service areas. **Transportation**, v. 41, n. 1, p. 193–210, 2014.

ESTUPIÑÁN, N.; RODRÍGUEZ, D. A. The relationship between urban form and station boardings for Bogota's BRT. **Transportation Research Part A: Policy and Practice**, v. 42, n. 2, p. 296–306, 2008.

EWING, R.; CERVERO, R. Travel and the built environment: a synthesis. **Transportation research record**, v. 1780, n. 1, p. 87–114, 2001.

EWING, R.; CERVERO, R. Travel and the built environment: A meta-analysis. **Journal of the American planning association**, v. 76, n. 3, p. 265–294, 2010.

FITZPATRICK, K.; BREWER, M. A.; TURNER, S. Another look at pedestrian walking speed. **Transpor-tation Research Record**, v. 1982, n. 1, p. 21-29, 2006. Disponível em: <https://bit.ly/3XulLcQ>.

GADUH, A.; GRAČNER, T.; ROTHENBERG, A. D. Life in the slow lane: Unintended consequences of public transit in Jakarta. **Journal of Urban Economics**, v. 128, p. 103411, 2022.

GIANNOTTI, M. et al. Inequalities in transit accessibility: Contributions from a comparative study between Global South and North metropolitan regions. **Cities**, v. 109, p. 103016, 2021.

HADDAD, E. A. et al. The underground economy: tracking the higher-order economic impacts of the São Paulo subway system. **Transportation Research Part A: Policy and Practice**, v. 73, p. 18–30, 2015.

HADDAD, E. A.; BARUFI, A. M. B. From rivers to roads: Spatial mismatch and inequality of opportunity in urban labor markets of a megacity. **Habitat International**, v. 68, p. 3–14, 2017.

HADDAD, E. A.; VIEIRA, R. Mobilidade, acessibilidade e produtividade: nota sobre a valoração econômica do tempo de viagem na região metropolitana de São Paulo. **Revista de Economia Contemporânea**, v. 19, p. 343–365, 2015.

HANDY, S. L. et al. How the built environment affects physical activity: views from urban planning. **American journal of preventive medicine**, v. 23, n. 2, p. 64–73, 2002.

HAYNES, KINGLSEY E., and A. Stewart Fotheringham. "Gravity and spatial interaction models." (1985).

KAMRUZZAMAN, M. et al. Advance transit oriented development typology: case study in Brisbane, Australia. **Journal of Transport Geography**, v. 34, p. 54–70, 2014.

MCFADDEN, D. The measurement of urban travel demand. **Journal of public economics**, v. 3, n. 4, p. 303–328, 1974.

MOBILIDADOS, https://mobilidados.org.br/rms/rmsp.

MONTGOMERY, J. M.; NYHAN, B.; TORRES, M. How conditioning on posttreatment variables can ruin your experiment and what to do about it. **American Journal of Political Science**, v. 62, n. 3, p. 760–775, 2018.

MURRAY, A. T. et al. Public transportation access. **Transportation Research Part D: Transport and Environment**, v. 3, n. 5, p. 319–328, 1998.

O'SULLIVAN, S.; MORRALL, J. Walking distances to and from light-rail transit stations. **Transportation Research Record**, v. 1538, n. 1, p. 19–26, 1996.

PEREIRA, R. H. et al. Forma urbana e mobilidade sustentável : evidências de cidades brasileiras. **Texto para Discussão IPEA**, v. preliminar, 2022.

PEREIRA, R. H. M.; SCHWANEN, T. **Tempo de deslocamento casa-trabalho no Brasil (1992-2009): diferenças entre regiões metropolitanas, níveis de renda e sexo**. [s.l.] Texto para Discussão, 2013.

PEREIRA, R. H.; SCHWANEN, T.; BANISTER, D. Distributive justice and equity in transportation. **Transport reviews**, v. 37, n. 2, p. 170–191, 2017.

POJANI, D.; STEAD, D. Policy design for sustainable urban transport in the global south. **Policy Design and Practice**, v. 1, n. 2, p. 90–102, 2018.

ROY, J. R.; THILL, J. C. Spatial interaction modelling. **Papers in Regional Science, v**. 83, n. 1, p. 339–361, 2004.

TSIVANIDIS, N. **Evaluating the impact of urban transit infrastructure: Evidence from bogota's transmilenio**. UC Berkeley (mimeo), 2020.[Google Scholar], , 2019.

VALE, D. S. Active accessibility and transit-oriented development: Connecting two sides of the same coin. Em: **Urban Form and Accessibility**. [s.l.] Elsevier, 2021. p. 123–140.

VAN OMMEREN, J. N.; GUTIÉRREZ-I-PUIGARNAU, E. Are workers with a long commute less productive? An empirical analysis of absenteeism. **Regional Science and Urban Economics**, v. 41, n. 1, p. 1–8, 2011.

VENTER, C.; MAHENDRA, A.; HIDALGO, D. From mobility to access for all: Expanding urban transportation choices in the global south. **World Resources Institute, Washington, DC**, p. 1–48, 2019.

VIEIRA, R. S.; HADDAD, E. A. An accessibility index for the metropolitan region of São Paulo. Em: **The Rise of the City**. [s.l.] Edward Elgar Publishing, 2015. p. 242–258.

WILSON, Alan Geoffrey. A family of spatial interaction models, and associated developments. **Environment and Planning A**, v. 3, n. 1, p. 1-32, 1971.

WREDE, M. A continuous spatial choice logit model of a polycentric city. **Regional Science and Urban Economics**, v. 53, p. 68–73, 2015.

ZENOU, Y. How do firms redline workers? **Journal of urban Economics**, v. 52, n. 3, p. 391–408, 2002.