Understanding the incentives for commuting flows: the impacts of a rapid transit expansion in the megacity of São Paulo

Luiz Pedro Couto Santos Silva[1]

Abstract

This paper examines the impacts of a rapid transit network expansion on travel flows in São Paulo over 10 years. Using spatial gravity-based models with granular data, it evaluates access improvements to BRT, train, and subway stations through the concept of spatial catchment area. Rail modes had a -3.9% travel time reduction over cars, although buses havent shown any reduction. Better access to rail stations increased travel flows more than BRT stations, with rail users being more sensitive to walking times. The findings suggest that combining faster transit with improved station access could boost public transit flows and reduce congestion.

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

Área: 4. POLÍTICAS PÚBLICAS: GÊNERO, RAÇA, INCLUSÃO

[1] Visiting student at UC Irvine and PhD candidate at Federal University of Juiz de Fora.

1. Introduction

The reduction in travel costs was a noticeable influence of motorized vehicles on the spatial structure of urban areas during the XX century (Anas et al., 1998). If the spatial distribution of land use attracts both economic activity and the population through agglomeration externalities, then urban transport systems are effective in fostering proximity (Ahlfeldt et al., 2015; Fujita, Krugman and Venables, 2001; Glaeser et al., 2001). The side effects of urban agglomeration include congestion costs, which are among the main challenges of large metropolitan areas, as they reduce workers' welfare due to more exposure to disamenities such as stressful commutes, likelihood of absence to work, and pollution, and have negative effects on worker productivity (Van Ommeren and Gutiérrez-I-Puigarnau, 2011; Zenou, 2002; Yriakopoulou and Xepapadeas, 2013). The rapid urban demographic growth in Latin America in recent decades has challenged the capacity of its cities to provide transport infrastructure (Bryan et al., 2020; Pojani and Stead, 2018; Vasconcellos, 2005). The lack of coordination between land use and transport infrastructure policies increases these congestion forces in Latin American cities, which are worse in large urban areas due to the greater complexity of coordinating transport policies between multiple municipalities (Pojani and Stead, 2018).

A growing body of literature on urban economics uses spatial granulated data in gravity-based models to estimate how exogenous schocks on urban structure (e.g., transport network, land markets) affect the spatial distribution of travel flows (Ahlfeldt et al., 2015; Ahlfeldt and Wendland, 2016; Ahrens and Lyons, 2021; Balboni et al., 2020; Dingel and Titelnot, 2020; Tsivanidis, 2023). A more recent use of gravity-based models aims to estimate the effects of improvements of the physical access to the public transit system on the distribution of urban travel flows (Gaduh et al., 2022). The objective of this paper is to contribute to this strand of literature by understanding how the improvement on the physical access to the public transit system affects the spatial pattern of travel flows in São Paulo metropolitan region (SPMR). The paper draws on the literature of station catchment area for the rapid transit modes of rail and BRT to observe incentives for the realization of trips between pair of districts of SPMR, which desintangles the incentives for travel using its public transport system. The novel empirical approach of this paper uses granular spatial information about the street network of SPMR to measure the changes on the population and jobs covered by the rapid transit system. It combines these measures of station catchment areas with data from Origin-Destination surveys of SPMR during the 2007-2017 period in econometric models to estimate whether such improvements enhanced gravity effects on bilateral travel flows.

SPMR is the megacity with the largest population of Americas, in which increases in household income in recent decades have led to a sharp increase in the usage of private vehicles for travel (Bocarejo, 2020; Carvalho and Pereira, 2013; Vasconcellos, 2005). This preference for private vehicles has burdened the transport infrastructure and increased the travel times of this metropolitan area in recent decades (Carvalho and Pereira, 2013). Evidence shows that it became an economic issue, since each 10 minutes spent on commuting reduces the potential earnings in its labor market by -2.7% (Haddad et al., 2015). Nevertheless, the promotion of mode-shift behavior in Global South megacities is complex. It evolves promoting speed of travel, spatial connectivity, and easy access to the transit system (Brooks and Deneoux, 2022; Bocarejo et al., 2020). During the period (2007-2017), SPMR 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 transport.

However, the geographic area of influence of transit stations on the propensity of individuals to travel by public transit transport mode is limited (El-geneidy et al., 2014; Vale, 2021; Kamruzzaman, et al., 2014). Their area of influence relate to the heterogeneous level of speed according to the public transit mode (e.g., bus, rail, subway) (Estupiñan and Rodriguez, 2008; Murray et al., 1998). These qualitative dimentions of travel conditions influence on passenger comfort, given that walking time is inherent to travel by public transit (Vale, 2021), but this aspect has received little attention among urban economists.

Thus, the contribution of the paper is threefold. First, it makes an empirical analysis to understand changes in the relative differences in travel times between different transport modes before and after the investiments in the public transport system of SPMR. This investigation follows the intuition of Gaduh et al. (2022) to understand if the expansion of the rapid transit system promoted relative gains in travel time, and therefore, could incentivize changes in travel behavior towards more usage of the public transport. Second, the spatial granular information about the streets network that connects households, rapid stations and jobs brings a novel approach to understand the relationship between station catchment areas and the spatial distribution of travel flows for this megacity. The goal of this approach is to compute the influence of the convenience to use the rapid transit system. Finally, the paper brings the first evidence about changes on the spatial pattern of travel flows of SPMR using a gravity-based spatial interaction model approach. This sort of analysis is scarce in developing countries due to the lack of multi-period spatial granulated data. The spatial interaction models developed on this paper allow disentangling the incentives for travel through the transportation system of SPMR and can be useful to guide policy implications.

The remainder of the paper is organized as follows. The next section presents a literature review about station catchment areas and gravity commuting models, followed by a section that describes the study area and another section that details the data and identification strategy used. Then, the last section presents the results of and final remarks about 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 transport 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 travel and the comfort promoted through physical access to the public transport system are understood as mechanisms for transport mode shift (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 presents 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 dimension, forces of attraction at the origin and destination and the efficiency of the transport network on promoting their connectivity can explain the intensity of bilateral travel flows between the pair of distinct areas. This sort of bilateral travel flows have been object of study of a long strand of literature related to the spatial interaction models (Wilson, 1971; Haynes and Fotheringham, 1985; Roy and Thill, 2004).

Recent literature has collapsed the strategies to identify travel behavior on 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 indidividual choice is also governed by the idiosyncratic 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 level. If these land and job markets clear, the spatial equilibrium under the multiple demand and supplies for city blocks draw the commuting flows between them. Therefore, changes in travel cost could induce the demands for city blocks, both in 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 the commuting costs and intensified the commuting flows between these locations that were formely separated. Their model also computes agglomeration economies through the wages and rental prices, which they argue that 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 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 transport network. This cumulative market access (CMA) detects the agglomeration economies promoted by the new BRT network through impacts in land and wage prices in the city blocks treated (those who had increases in CMA over the period).

When well succeeded in terms of reducing tavel 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 mode. Under these assumptions, Severen (2021) investigated the effects of the proximity to new rail stations on the patterns of commuting flows of the Great Los Angeles. He adopts three rules to assign the treatment: 1) by observing the census tracts that had received a new rail station; 2) considering an euclidean distance of 250 meters between census tracts' centroids and a new station; and 3) same as 2) with 500 meters. His models have found positive impacts of the proximity to the new rail stations on the number of bilateral commuting flows. The study of Gaduh et al. (2022) analyzed the effects of BRT network expansion on the travel flows of Jakarta. The authors assigned the treatment to observe the access to the transit system through euclidean distances of 1 km between the borders of his unit areas and the BRT stations. They found no signiticative impact of better access to BRT stations on the probability of bilateral travel flows. This result is justified due to the lack of gains of relative speed of travel through the BRT network when compared to private vehicles, given the insuficient investment in infrastructure from this policy. However, these simple measures of distance of Severen (2021) and Gaduh (2022) do not account for the streets' design and overestimates 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 with spatial refinement (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 transport modes and the incentive to use the public transport system.

Wheather the distance from the origin and destiny to the transit station is enough to encourage the use of the public transport system, there are incentives that prevent the use of private motorized transports. 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 commuting behavior, an increase in population density near transit systems can regional accessibility (Pojani and Stead, 2018; Venter et al., 2019). Furthermore, the land use policies that foment this transit-oriented densification in developing countries should focus on middle-poor income populations since they face more restrictions to better levels of accessibility due to the spatial distribution of intrastructure (Boisjoly et al., 2020; Venter et al., 2019; Vasconcellos, 2017).

3. The study area

With 22 million inhabitants, São Paulo metropolitan region is the largest urban agglomeration in American continent, in which approximately 19% of the Brazilian GDP is concentrated (IBGE, 2021). The rapid population growth of SPMR in recent decades 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 population is significantly dispersed throughough SPMR, there is a concentric spatial distribution of jobs towards São Paulo city. 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).

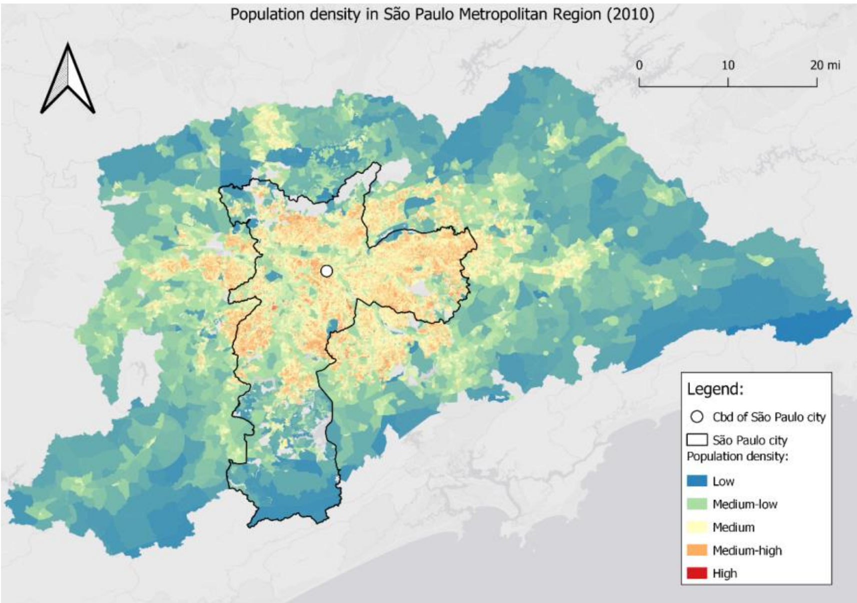

Figure 1 – Population density in São Paulo Metropolitan Region (2010)
Source: author's own, from Brazilian census of 2010.

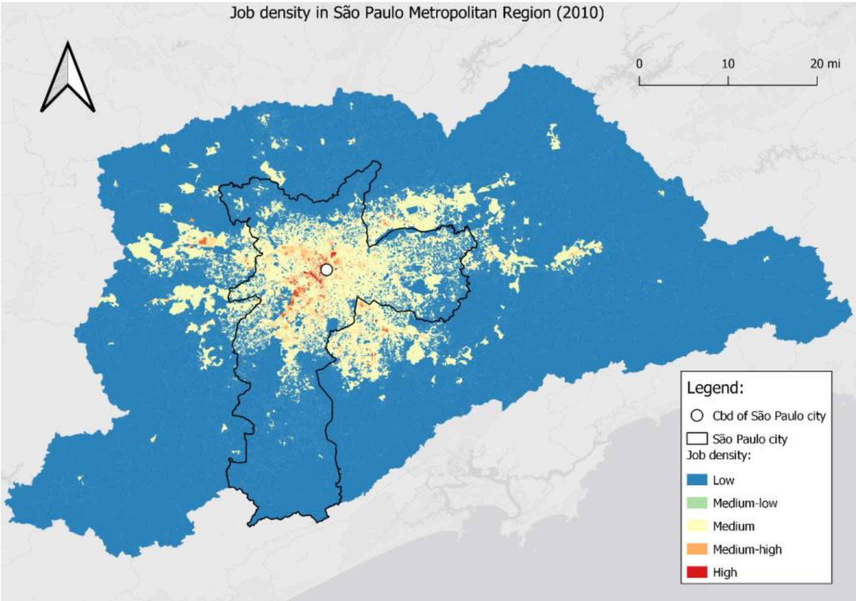

Figure 2 – Job density in São Paulo Metropolitan Region (2010)
Source: author's own, from the Annual Social Information Report of 2010.

In this context of challenges in the transport infrastructure, 65 new rapid transit stations were built in the SPMR between 2007 and 2017. Figure 3 shows that by 2017, the subway system expanded from the central region to the western region of São Paulo city on the yellow line, toward the eastern region by the green line, and from the southwestern region to the central-southern zone through the purple line. The rail system expanded its existing network system in the city of São Paulo toward its eastern and southern regions and from São Paulo's northeast to the city of Guarulhos. These new rail lines totaled 30 km of expansion in 10 years.

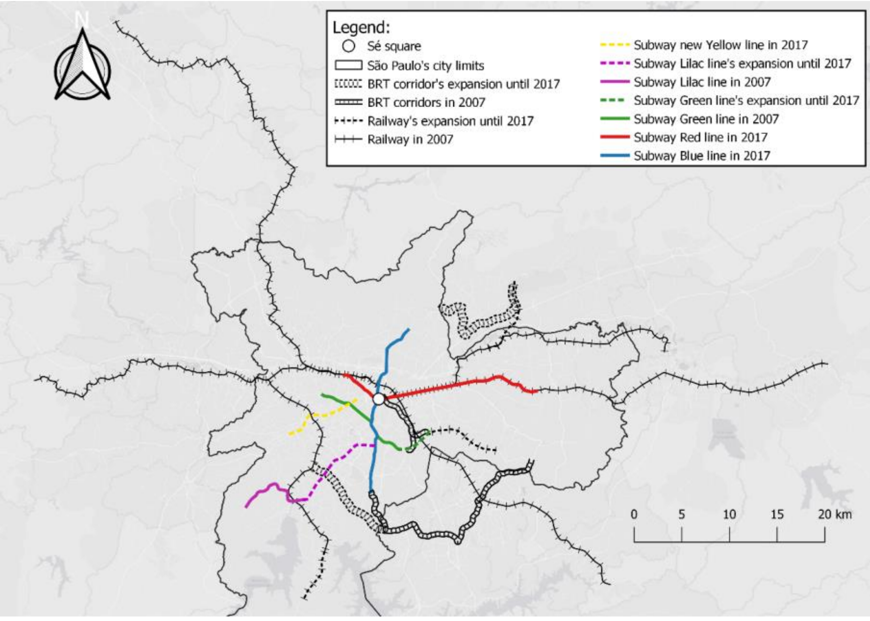

Figure 3 – The rapid transit system of the SPMR in 2017.
Source: Author's own elaboration, from OSM and IDTP data.

Figure 3 also shows that there was 26 km of expansion for the BRT network from the south region to the southwest region of the city of São Paulo and a new corridor opened in the city of Guarulhos. Besides that, the administration of the city of São Paulo launched the "single fare" rule in 2005 (Rolnik and Klintowitz, 2011). This new fare rule allowed the use of the whole public transport system (buses, subway and rails) after paying one single fare that was lower than the fares that one would pay to take each public transport mode singly (São Paulo, 2008; Sptrans, 2013). In 2014, this fare rule was expanded to the remaining municipalities of the metropolitan region, along with a monthly plan fare (Santiago, 2013). Therefore, the effects of these incentives to use the public transport system in SPMR must be assessed with the aims to .

4. Material and methods

4.1 Data

This study uses pseudo panel data of SPMR about travel, income, population and public transit station networks. The travel, household, and work place information are drawn from household travel surveys (Origin-Destination) conducted in SPMR in the years of 2007 and 2017 (Metro, 2007; 2017). These OD surveys are sampled considering geographic zones and districts that are based on census tracts. Household OD information is available in point coordinates. The OD information was used to quantify the number and duration of commuting trips between districts by transport mode. The spatial structure of the database consists of the OD districts. Data from OpenStreetMap (OSM) was also used to obtain information about the street network of the SPMR, which is a cross sectional data for the year of 2023. Because of that, we hold on the assumption that the street's network havent had significative changes between 2007 and 2017. The data of rapid transit stations is from MOBILIDADOS[2], which provides information of coordinates of the station, mode type, as well as date of inauguration and closure.

The street network data from the OSM was used to calculate the shortest route by walking from the centroid of the location of each household to every rapid transit station using the r5r package in R, developed by Pereira et al. (2022). I adopt an average walking speed of 3.6 km/h, following Fitzpatric and Brewer (2006), to build this travel time matrix. The study computed different walking time thresholds to reach the rapid transit stations (10, 20 and 30 minutes) considering their spatial distribution for each year (2007 and 2017). The combinations between these walking time information and the spatial distribution of rapid transit stations allows observe the changes of the percentage of population and jobs that were covered by the rapid transit system at each district on the OD surveys of 2007 and 2017.

The descriptive statistics for the OD surveys data by transport mode are presented on 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 station also decreased for all of the transport modes, which relates to its network expansion. Although the mean household euclidean distance from household to the rail network (Train or subway) remained stable for bus users, it increased for car users and decreased for rail users. The data in table 1 is also suggestive of sorting for access to high speed modes, by which the car users have the higher household incomes, followed by rail users and then by bus users.

Table 1 – Descriptive statistics.

Commuting transport mode	Bus	Car	Rail

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

Year	2007		2017		2007		2017		2007		2017	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Commuting time (Minutes)	53	31	44	25	30	26	26	21	47	28	43	22
Commuting 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
Linear 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
Educational level:												
Illiterate	9%		9%		8%		10%		3%		2%	
Primary school	17%		12%		9%		7%		8%		3%	
Elementary school	19%		18%		9%		7%		13%		7%	
High school	42%		44%		29%		29%		44%		40%	
Bachelor degree	12%		18%		45%		47%		32%		48%	
Number of observations	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. Commuting transport mode is based on the only transport mode declared by the individual on the survey. Household income are in nominal values.

4.2 Econometric models

Linear models to estimate differences in travel times between transport modes

Following Gaduh et al. (2022), we used OLS models to estimate how the differences of travel times between different public transport modes and private vehicles (cars) in SPMR changed over the period of the analysis (2007-2017).

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

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 cars with PT: bus, rail (subway and train) and the aggregated public transport modes. *X* is a set of control variables: log of linear distance, hour departing time, weekday, origin zone, destination zone, travel purpose, and interaction dummy between hour and weekday. Moreover, ε are robust standart errors clustered by an interaction between the origin and destination zones.

We also followed the framweork 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 lived within a radious distance of 1 km from a BRT station. The aim of these models is to estimate if the public transport modes had reduced their relative difference of travel time compared to cars in the years of 2007 and 2017, which could be related to the transport policies of SPMR for the period.

Gravity model to estimate incentives and desincentives for travel by public transport

This study used poisson models to estimate the relationship between the percentage of districit's population and jobs within rapid stations' catchment areas and the distribution of travel flows. The models follow the framework proposed by Ahlfeldt et al. (2015) and Gaduh et al. (2022) to estimate spatial interaction for intraurban areas based on economic incentives. It assumes that the observed quantity of bilateral travel flows from home to the destiny place (with the purposes to work or study) reflects a spatial equlibrium determined by demands and supplies of amenities located in the districts. The probability of bilateral travel flows drawn from this equilibrium is:

$$\pi_{ij} = \frac{\sum_{i=1}^{n} W_{ij}}{\sum_{i=1}^{n} W_{i}}$$

(2)

where π is the probability of interaction between the district of origin *i* and destination *j* explained by the numer of residents *W* in *i* who traveled to *j*. The sample is restricted to individuals who traveled by public transport to work or study. The variable π is balanced by push and pull factors, such as the amount and quality of opportunities in *i* and *j*, which will determine the gravity forces for travel. The transport network plays a role on these gravity forces by intermediating their spatial connectivity, in which the travel cost reduces the utility level achieveble through the interaction between *i* and *j*. To explore the dimensions of the (dis)incentives for bilateral travels, we extend the framework of Ahlfeldt et al. (2015) and Gaduh et al. (2022) by using station catchment areas as a measure of access to the transit system. The reduced form of the poisson models follow:

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

(3)

where the ln of the probability of travel between the origin and the destination districts at year *t* is explained by %*Covered*, the percentage of population and jobs 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 for the change in population and jobs covered for the districts at the origin and destination over the years 2007 and 2017. *Time* is the average travel time by public transit between *i* and *j*. Moreover, δ and γ are fixed effects for the origin and destination, respectively, as well as φ, υ, and *T* are time-origin, time-destination and time-year fixed effects, respectively, and ε is an origin and destination clustered error term. Therefore, because $\%Covered_{it} * \%Covered_{jt}$ is time variant over the period (2007-20017), β_1 in equation (3) estimates how the increases in the proportion of population and jobs covered by the rapid transit stations affected the distribution of bilateral travel flows of SPMR.

This strategy follows the same intuition of Gaduh et al., (2022), which used the distance from the boarders of each district's polygon to BRT stations to calculate their measure of access to the transit system. However, the measure of access used here has two advantages: 1) the use of latitude and longitude coordinates of households and job or study locations. It brings more geographic precision about the starting and ending points of the travel; 2) the use the walking distances along the transport network, which is more precise and realistic than euclidean distances and is better able to capture the influence of urban form on walking access to the transit system. Moreover, the approach allows us to estimate the spatial decay of the station catchment given varying walking times. Finally, equation (3) estimates both the incentives through the ease of access to the transit system and the disincentives for longer travels, which allows us observing the two sides of the coins of the travel through public transport (Vale, 2021).

5. Results

5.1 Relative differences in travel time between transport modes

The results through the differences in travel time between the public and private transport system were analyzed, as we assume it is a potential mechanism for changes in commuting behavior. Figure 4 shows the average difference between travel times by public transport (bus and heavy rail) and private (car) by income quintile. It suggests that the median travel time difference between these transport modes (travel time by Public transport minus travel time by car) had only reduced slightly for the third quintile and increased slightly in the fourth quintile. Although there were decreases in the average travel times of each transport mode at each income quintile (METRO, 2017), it seem that when we compare travel time between public transport and cars, these relative differences in time were stable from 2007 to 2017. To better understand this, formal tests were run in linear regression analyzes, with the aim to estimate such relative differences in travel time between public and private transport.

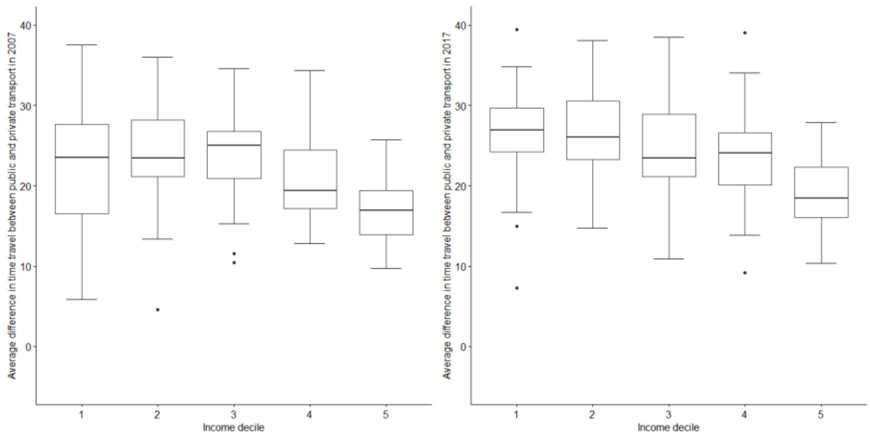

Figure 4 – Box plots for average difference in travel time between public and private transport travels by income quintile.

Source: Author's own, from Origin Destination surveys of SPMR of 2007 and 2017.

The results of the OLS regressions that test the differences in travel times between public transport and cars are reported in Table 2. The coefficients in columns (1-2) suggest a slight increase on the relative differences of travel times between buses and cars of 1.7% percentage point over the years of 2007 and 2017. In the year 2007, the mean travel time by bus was 57% higher than by cars, and 59% in the year 2017. In an opposite trend, the relative difference in travel time between rail (subway + train) and cars in columns (3-4) had decrease from the year 2007 to the year 2017 in -3.9% percentage points. It suggests some improvement in the efficiency of the rail service relative to cars through the SPMR, thus, travels made through the heavy rail system reduced its relative costs in terms of time of travel in comparison with travels made by car.

Overall, the results of the columns (5-6) in table 2 evidence that the aggregated travels through the public transport (bus and rail) improved only slightly its relative travel time in comparison to cars. In fact, the better results for the rail transport system shown in columns (3-4) are expected due to the higher magnitude of investment in the expansion of the rail than of the buses infrastructures in the study area. Although there was an expansion in the RMSP of 26 km in new BRT corridors during the 10 year period, it seem that such new infrastructure still did not generate reasonable gains of speed of commutation to compete with the speed of travels made by cars.

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

	2007	2017	2007	2017	2007	2017
Model	(1)	(2)	(3)	(4)	(5)	(6)
Bus x Car	0.576***	0.593***				
	(0.0123)	(0.0059)				
Rail x Car			0.321***	0.282***		
			(0.0216)	(0.0128)		
PT (all) x Car					0.542***	0.516***
					(0.0104)	(0.0058)
Adjusted R^2	0.59	0.78	0.58	0.78	0.59	0.79
Sample (N)	73,305	74,391	53,392	58,468	106,178	100,672

Notes: This table reports linear regression models that have the log of individual travel time as dependent variable. Each model has a dummy variable indicating a travel made by a public transport mode, compared with travels made by car. The additional controls in the regressions are: log of linear 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, clustered by an interaction of origin and destination zones, are reported in parentheses. * / ** / *** denotes significant at the 10% / 5% / 1%, respectively.

Further investigation about the performance of BRT lines in comparison to other transport modes for the RMSP following similar approach of Gaduh et al., (2022) was made and presented on Table 3. It is assumed that travels by bus that started and ended within the distance of 1 km from BRT stations were made by the BRT mode. The results show statistical difference for the difference in travel time between BRT and bus in the year 2007, although it changed in the year 2017. Thus, Table 3 shows a marginal difference of -7.5% on the time travel between BRT and buses for the year 2007, and no statistical difference between the travel time of bus and BRT commuters for the year 2017. However, the columns (3-4) suggest that the higher times of BRT travels in comparison with cars had a marginal increase in percentage points of 5.4%. In the comparison between the BRT with the rail system (train and subway), columns (5-6) in Table 3 show that there was a difference in travel time of 25% between BRT and rail commuters for the year 2007. However, Table 3 shows a significative reduction on that difference in travel times between BRT and Rail commuters for the year 2017, in which I found a reduction of -9.4% percentage points in comparison to the year 2007.

Table 3 – OLS models for the estimates of the differences of time travel between BRT buses and other transport modes.

	2007	2017	2007	2017	2007	2017
Model	(1)	(2)	(3)	(4)	(5)	(6)
BRT x Bus	-0.075**	-0.0006				
	(0.0295)	(0.0193)				
BRT x Car			0.530***	0.584***		
			(0.0294)	(0.0172)		
BRT x Rail					0.255***	0.161***

					(0.0405)	(0.0236)
Adjusted R^2	0.49	0.61	0.58	0.78	0.64	0.70
Sample (N)	25,309	20,839	49,396	55,463	6,796	6,827

Notes: BRT commuters were defined as the individuals who commuted by bus and lived within a distance of 1 kilometer of a BRT station. The dependent variable is the log of individual travel time. The additional controls in the regressions are: log of linear 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, clustered by an interaction of origin and destination zones, are reported in parentheses. * / ** / *** denotes significant at the 10% / 5% / 1%, respectively.

Even after the expansion of the rapid transit network between the years 2007 and 2017, there are still many areas in the SPMR without quick access to the rapid transit network. According to Mobilidados (2023), only 12% of the total population of SPMR lived within 1 kilometer from a rapid transit station in the year 2017. However, that percentage increases to 31% for those on the highest income quintile, and decreases to 10% on the lowest income quintile. This spatial concentration of improvement of rapid transit infrastructure relates to the reductions in the relative travel time in the rail system shown in the results in Figure 3, in which it is possible to observe a very punctual reduction of relative mean travel time for the third income quintile. Given that the subway system had the better improvements in commute times in comparison to cars and that the new subway lines are mostly located in areas with middle-higher income population, it is possible to affirm that this population was the most benefited by the improvements in speed through the public transport.

5.2 Results for gravity commuting in the SPMR

In this section, the results of the models following equation 3 are presented using different specifications to observe how the consideration of different variables to observe the commuting costs affect the results. Table 4 for reports the poisson models with the coefficients for the variable that measure the station catchment areas for BRT mode. It shows that each 10 percentage increase on the number of residences or jobs covered

Comentado [RP1]: para testar se os resultados são sensiveis quando se considera catchment area de diferentes tamanhos. Alias, essa decisão de fazer analise de sensiilidade para tamanho da catchment area tinha q ter sido explicada / justificada na secao de metodos within 10 minutes from a BRT station increases the probability of commuting between the pair of districts in 1.08%. This positive relationship decreases to 0.2% when considering a walking time threshold of 20 minutes, and to 0.1% for a longer threshold of 30 minutes.

Table 4 - Gravity models with station catchment areas for BRT stations

Dependent Var.:	Log prob of commuting		
Walk time to/from nearest station	10 minutes	20 minutes	30 minutes
%ResCovBRT xJobsCovBRT	0.0108***	0.0020***	0.0010***
	(0.0032)	(0.0004)	(0.0002)
Log of travel time	-1.679***	-1.662***	-1.650***
	(0.1616)	(0.1592)	(0.1560)
Fixed-Effects:			
Origin	Yes	Yes	Yes
Destination	Yes	Yes	Yes
Origin & year	Yes	Yes	Yes
Destination & year	Yes	Yes	Yes
Year	Yes	Yes	Yes
S.E.: Clustered	Origin & Destiny	Origin & Destiny	Origin & Destiny
Observations	3,508	3,508	3,508
Squared Cor.	0.48091	0.49295	0.49589
Pseudo R2	0.09210	0.09351	0.09398
BIC	7,133.9	7,133.0	7,132.7

Source: author's own.

Notes: "%ResCovBRT" and "%JobsCovBRT" are the percentages of residences and the percentage of jobs within the station catchment areas of BRT, respectively. These variables follow the explanation given in equation (3).

This relationship of lower effects of access to transit stations on travel flows as we adopt larger threshold of walking time in the models of Table 4 is expected, since the increase in walk time to reach a station increases the disutility and reflects a disincentive to use to public transit system. The travel time is the other component of commuting cost.

Its coefficients are also reported in the models of Table 4 and show a negative relationship with the probability of commuting flows between pairs of districts. Given the specification in equation 2, these coefficients for travel time can be interpreted as elasticities. They are significantly higher than Ahfeldt et al. (2015), which found -0.07 for Berlin, and Gaduh et al. (2022), with -0.059 for Jakarta. Either the coefficients for the access to transit station and the elasticities of travel time are very marginally sensible to the inclusion of one of them in the model.

The models for the effects of rail stations (subway and trains) on the commuting flows through the rapid transit system are reported in Table 5. They show that the effects of these transit modes are much higher in the threshold of 10 minutes walk than of BRT stations, and therefore, at each increase of 10% of jobs or population covered, the probability of commuting between pairs of districts increase by 2.1%. However, the decay effect of walking time to a rail station on the probability of commuting is stronger than to BRT stations. Finally, the travel elasticities on commuting flows do not differ significantly for the estimates of the models on Table 4. This is expected because all the models have the same sample, and the differences are on which kind of rapid transit station affect the commuting flows for the aggregated rapid transit system.

Table 5 – Gravity models with station catchment areas for Rail stations

Dependent Var:	Log prob of commuting		
Walk time to/from nearest station	10 minutes	20 minutes	30 minutes
%ResCov Rail x %JobsCov Rail	0.0219** (0.0067)	0.0016*** (0.0003)	0.0009*** (8.99e-5)
Log travel time	-1.673*** (0.1610)	-1.609*** (0.1605)	-1.579*** (0.1602)
Fixed-Effects:			
Origin	Yes	Yes	Yes

Destiny	Yes	Yes	Yes
Origem & year	Yes	Yes	Yes
Destiny & year	Yes	Yes	Yes
Year	Yes	Yes	Yes
S.E.: Clustered	Origin & Destiny	Origin & Destiny	Origin & Destiny
Observations	3,508	3,508	3,508
Squared Cor.	0.47493	0.47870	0.49535
Pseudo R2	0.09184	0.09261	0.09492
BIC	7,134.0	7,133.5	7,132.1

Source: Author's own.

Notes: "%ResCovRail" and "%JobsCovRail" are the percentages of residences and the percentage of jobs within the station catchment areas of rail, respectively. These variables follow the explanation given in equation (3).

These results dialogue with the literature of 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 the rail system is more efficient on promoting speed than the bus system, the users are more sensible 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 commuted by bus, it is also expected that the rail transit users are less propense to walk longer distances to reach the transit system than BRT users.

A further step of the analysis, models the influence of rail and BRT stations together. This is a sensitivity analysis, given that every gravity model shown at this study consider both commuting trips made by rail or bus[3]. However, Table 6 shows that the inclusion of both variables that measure the level of coverage of BRT and rail stations in the same models don't change their magnitudes neither standard error significantly. This is an indicative that there might not be high collinearity between the variables in the models.

Table 6 – Gravity models with station catchment areas for BRT and Rail

Dependent Variable:	Log prob of commuting

[3] Considering only rail or bus on the gravity models result in small sample, which reduces the reliability of the estimates.

Walk time to/from nearest station	10 minutes	20 minutes	30 minutes
%ResCovBRT x %JobsCovBRT	0.0108***	0.0020***	0.0009***
	(0.0032)	(0.0004)	(0.0002)
%ResCovRail x %JobsCovRail	0.0217***	0.0015***	0.0009***
	(0.0065)	(0.0003)	(8.48e-5)
Log travel time	-1.667***	-1.583***	-1.542***
	(0.1601)	(0.1570)	(0.1535)
Fixed-Effects:			
Origin	Yes	Yes	Yes
Destiny	Yes	Yes	Yes
Origin&year	Yes	Yes	Yes
Destiny&year	Yes	Yes	Yes
Year	Yes	Yes	Yes
S.E.: Clustered	Origin & Destin	Origin & Destin	Origin & Destin
Observations	3,508	3,508	3,508
Squared Cor.	0.48873	0.50443	0.52332
Pseudo R2	0.09360	0.09574	0.09838
BIC	7,141.1	7,139.7	7,138.1

Source: Author's own.

Notes: "%ResCov" and "%JobsCov" are the percentages of residences and the percentage of jobs within the station catchment areas, respectively. These variables follow the explanation given in equation (3).

Therefore, the ranking in the magnitudes of the effects of access to the transit stations remains: rail stations have stronger positive effects on gravity commuting flows than BRT stations within the 10 minutes threshold, but weaker effects from the 20 minutes threshold. The coefficients that change at most on the models of Table 6 are the elasticities of travel time on the model of 20 minutes of walk threshold from or to the nearest station, which reduce from the range of -1.6 to -1.58. Therefore, all of the models suggest that the sensitivity to commute relative to the travel time tends to reduce for those who are further from a rapid transit station, which may reflect a lack of options to commute.

Regarding the specifications of my gravity models, 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 destiny & 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). The sample noise was treated by only considering those bilateral commuting that had at least 10 occurrences in the raw sample. This is a common procedure on the literature of gravity models for urban areas (Ahfeldt et al., 2015; Ahlfeldt and Wendland, 2016; Gaduh et al., 2022), given that small sample representing bilateral commuting trips may reduce the precision of these kind of estimates (Dingel and Titelnot, 2020). In fact, then the models were estimated with the total sample, the squared correlation is smaller than when restricting the number of bilateral commutings.

Finally, a model with all of the possible interactions between the available information about commuting costs was implemented as an exercise in Table 7. The intuition behind such strategy is to estimate how the interaction between the station catchment areas and the travel time affect the probability of bilateral commuting, by checking if these specifications could better measure the decisions of individuals to commute.

The models in Table 7 suggest the same ranking in magnitude for the rail and BRT station coverages on the commuting flows of the models without interaction between station coverage and travel time. The new information here is that the interaction between station coverage level on residences and jobs and the travel time results in negative coefficients. Although negative, these new coefficients with interaction have lower magnitude than the coefficients for the coverage level of BRT and rail stations, as well as the elasticities of travel time. It seems to be a compensating adjustment, given that an increase in the value of the variable that measure the coverage level of stations decreases the probability of commuting flows when interacting with the travel time. Also notice that the interaction variable of station coverage and travel time in the model with threshold of 20 minutes of walk from or to the stations is not statistically significant. Therefore, the interpretation of such interactive coefficients between the level of coverage of stations and travel times needs further insights.

Table 7 – Gravity models with interactions station catchment areas and travel time

Dependent Var:	Log of prob commuting		
	10 minutes	20 minutes	30 minutes

%ResCovBRT x %JobsCovBRT	0.0325***	0.0049***	0.0018***
	(0.0090)	(0.0014)	(0.0005)
%ResCovRail x %JobsCovRail	0.0944*	0.0020*	0.0017***
	(0.0442)	(0.0012)	(0.0004)
%ResCov BRT x %JobsCovBRT x log travel time	-0.0068***	-0.0009*	-0.0003*
	(0.0020)	(0.0004)	(0.0002)
%ResCovRail x %JobsCovRail x log travel time	-0.0221*	-0.0002	-0.0002*
	(0.0118)	(0.0004)	(0.0001)
Log travel time	-1.622***	-1.543***	-1.438***
	(0.1678)	(0.1701)	(0.1792)
Fixed-Effects:			
Origin	Yes	Yes	Yes
Destiny	Yes	Yes	Yes
Origin & year	Yes	Yes	Yes
Destiny & year	Yes	Yes	Yes
Year	Yes	Yes	Yes
S.E.: Clustered	Origin & Destiny	Origin & Destiny	Origin & Destiny
Observations	3,508	3,508	3,508
Squared Cor.	0.49205	0.50626	0.52581
Pseudo R2	0.09391	0.09599	0.09889
BIC	7,157.2	7,155.9	7,154.1

Source: Author’s own.

6. Final remarks

This study aims to understand the effects of changes on the rapid transit network on the commuting behavior for the São Paulo Metropolitan Region between the years of

2007-2017. The empirical approach was based on the gravity commuting and station catchment area literatures, in the aims to observe changes on incentives and disincentives for commuting through public transportation within SPMR using longitudinal data in high spatial resolution. The exploratory results, shows no changes in the average relative difference on commuting times between buses and cars but decreases of 3.9% for the heavy rail system of SPMR (trains and subways). The analysis for the commuting flows suggests that the coverage level of rail stations are more attractive for commuters than BRT stations, which might relate to the better capacity that the first must promote speed. The exercise of flexing the walk time to or from the stations of my models are accordingly to the literature of station catchment areas, and suggest that rail stations have a stronger decay than BRT users.

The approach released on this empirical essay aims to further understand how different components of the travel by public transit can incentivize individuals to use the system. The main contribution is to conciliate an approach from the urban planning literature (station catchment areas) with an approach of urban economics literature (urban gravity models) with the aims to disentangle the incentives for commuting flows within the SPMR. The models suggest that the commuting time is only one aspect of the decision. Further investigations could use the approach developed in this essay to guide public policies about the implementation of transit stations by using counterfactual analysis, with the aims to estimate how different locations would use the public transit given the proximity to new transit stations.

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. **Transportation 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.